UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE AUGUSTIN COURNOT ED 221

BUREAU D'ÉCONOMIE THÉORIQUE ET APPLIQUÉE UMR 7522

THÈSE

pour l'obtention du titre de Docteur en Sciences Économiques

Présentée et soutenue le 8 Septembre 2023 par

Kenza BOUSEDRA

CHANGE IN COMMERCIAL SECTORS EXPLOITING SPACE INFRASTRUCTURE: ANALYSIS AND ECONOMIC INDICATORS

Préparée sous la direction de Sandrine WOLFF

Membres du jury :

Mireille MATT	Directrice de Recherche, INRAE	Présidente
Victor DOS SANTOS PAULINO	Professeur associé HDR, Toulouse Business School	Rapporteur
Vasilis ZERVOS	Professeur associé, University of Thessaly	Rapporteur
Patrick COHENDET	Professeur, HEC Montréal	Examinateur
Henry R. HERTZFELD	Professeur, George Washington University	Examinateur
Sandrine WOLFF	Maître de conférences HDR, Université de Strasbourg	Directrice

Membres invités :

Membre INVITÉ	Laurent Bach, Maître de conférences, Université de Strasbourg
Membre INVITÉ	Murielle Lafaye, Sous-directrice Observatoire du Spatial, CNES

À ma famille,

Remerciements

Nous y voilà. L'expérience toute particulière de la préparation d'une thèse touche à sa fin pour moi et le moment est venu de remercier les personnes qui m'ont apporté leur soutien d'une manière ou d'une autre pendant ces quelques années.

J'adresse mes premiers remerciements à Sandrine Wolff, ma directrice de thèse. Merci Sandrine pour votre confiance et votre infinie patience. Vous avez su me laisser la liberté de développer mes intuitions tout en m'apportant des conseils scientifiques et stratégiques avisés aux moments où il fallait. J'ai beaucoup appris de nos nombreux et longs échanges, et j'espère avoir acquis grâce à ceux-ci un peu de votre rigueur scientifique.

Je veux ensuite remercier Laurent Bach, mon co-encadrant. Merci beaucoup Laurent, car même si cela n'apparaît pas explicitement, vous avez beaucoup contribué à ce travail de recherche. J'ai eu beaucoup de chance de pouvoir profiter de votre expertise et de la richesse de vos connaissances, ainsi que de vos traits d'humour et votre bonne humeur à toute épreuve.

Je souhaite remercier le CNES et Telespazio France, les organismes financeurs de ce travail de recherche. Merci à Murielle Lafaye, mon encadrante côté CNES. Je vous remercie de m'avoir accordé votre confiance pour mener un projet d'une telle envergure. Merci aussi pour votre gentillesse. Vous m'avez permis de découvrir le domaine passionnant de l'espace et m'avez fait bénéficier de votre expertise sur ce milieu. Merci également à tous les membres du CNES que j'ai pu rencontrer au cours de ces années de thèse.

Je remercie Nicolas Vincent (ex-Telespazio) qui a fortement marqué mon début de thèse avec ses connaissances aiguisées sur l'industrie spatiale et ses précieux conseils.

I sincerely thank Mireille Matt, Victor Dos Santos Paulino, Vasilis Zervos, Patrick Cohendet, and Henry Hertzfeld for agreeing to serve on my thesis jury. I am highly honored to have my work assessed by such distinguished researchers.

Mes pensées vont ensuite à mes collègues et camarades de galère doctorant.e.s du BETA. Merci aux "ancien.nne.s" – Cyrielle, Thomas, Rémy, Nuria, Debora, Yann, Laetitia – vous m'avez beaucoup aidée pendant mes premiers pas en tant que doctorante. Une pensée pour Maxime L., merci beaucoup pour ton aide avec les données. Je remercie ensuite celles et ceux qui étaient présent.e.s jusqu'à la fin de cette thèse : Agathe, Anne-Gaëlle, Romane, Guillaume, Emilien, Jérôme, Antoine, mes collègues du bureau 133 Louis et Théo. Merci infiniment pour tous ces moments de rigolade. Vous avez sans le savoir égayé beaucoup de mes journées. Vous êtes le sang. Je n'oublie pas mon amie Kéa, merci pour ta générosité et ton écoute. J'ai une pensée particulière pour Pierre, avec qui j'ai rédigé un chapitre de cette thèse. Merci pour ton aide précieuse sur ce chapitre, je garderai toujours en mémoire nos discussions matinales interminables au bureau qui nous laissaient l'impression de n'avoir rien fait de notre journée.

Vient ensuite le moment de remercier ceux qui m'ont apporté le soutien le plus précieux, ma famille. Merci d'abord à mes parents de m'avoir accordé le meilleur pour que je puisse arriver là où je suis aujourd'hui. Je vous en suis très reconnaissante. Merci à mon frère Amine et ma soeur Nawel, j'espère que vous êtes fiers de votre grande soeur. Merci à ma grand-mère qui m'a tout appris sans le savoir. J'espère lui rendre un jour un millième de ce qu'elle m'a donné. Merci à ma tante Leïla de m'avoir écoutée et supportée pendant mes moments de doute et de difficulté. Merci à mon chat Tigri, mon fils, de rendre mes journées plus douces. J'aimerais pour finir citer mon papi Omar, je t'aime et il ne passe pas un seul jour sans que je pense à toi.

Enfin, Al HamdouliLah. Je remercie Dieu pour tout, toujours.

Table of Contents

List of Acronyms

APE Activité Principale Exercée.

ASE Agence Spatiale Européenne.

ASI Agenzia Spaziale Italiana.

BEA Bureau of Economic Analysis.

CBA Cost-benefit analysis.

CNES Centre National d'Etudes Spatiales.

CSA Canadian Space Agency.

DLR Deutsches Zentrum für Luft.

DoD Department of Defense.

EO Earth Observation.

ESA European Space Agency.

ESPI European Space Policy Institute.

EU European Union.

EUMETSAT European Organisation for the Exploitation of Meteorological Satellites.

EUSPA European Union Agency for the Space Programme.

GEO Geostationary orbit.

GNSS Global Navigation Satellite System.

INSEE Institut National de la Statistique et des Etudes Economiques.

ISS International Space Station.

KARI Korea Aerospace Research Institute.

LEO Low-Earth orbit.

NAF Nomenclature des Activités Françaises.

NASA National Aeronautics and Space Administration.

NER Named Entity Recognition.

NLP Natural Language Processing.

NOSA Norwegian Space Agency.

NSO Netherlands Space Office.

RCS Registre du Commerce et des Sociétés.

SA Société Anonyme.

SARL Société à Responsabilité Limitée.

SAS Société par Actions Simplifiée.

SataaS Satellite-as-a-Service.

Sirene Système Informatique pour le Répertoire des ENtreprises et des Etablissements.

SpaaS Space-as-a-Service.

SSO Swiss Space Office.

UKSA United Kingdom Space Agency.

General Introduction

In July 2023, the European Space Agency (ESA) awarded a contract of 16 million euros to Spire Global group to develop an independent aviation surveillance system based from space[1]. The company, founded in 2012, specializes in the collection and analysis of space data and the provision of space services. It develops its own constellation of Earth-observation nano-satellites for monitoring maritime activities, aviation, and climate from space. The company operates under an emerging business model in the space sector known as *Space-as-a-Service (SpaaS)*. This new trend, also referred to by the term *Satellite-as-a-Service (SataaS)*, marks a shift in the space industry, where private companies deliver access to services and uses from space infrastructures rather than the space products themselves [Hein and Bruce Rosete, 2022]. It allows clients, both institutional and civil, to benefit from the services offered by satellite technology without having to invest in the physical infrastructure[2]. Thus, the SpaaS model opens up significant opportunities for the development of commercial civil uses of space, aligning with the transformation of the space sector initiated with *New Space*.

The *New Space* era refers to a set of developments in the space sector initiated in the early 2000s in the United States [Denis et al., 2020, Pasco, 2017, Pekkanen, 2019]. This new context constitutes the backdrop of the thesis, which is why we will analyze it in depth in our development. *New Space* refers to opening access to space to new private players, including large companies and startups in the digital sector. The newcomers invest in the space sector with an entrepreneurial mindset and new business models based on expanding space commercial markets. Illustratively, Blue Origin, founded in 2000 by American billionaire Jeff Bezos, also the founder of the

[1] `https://www.esa.int/Applications/Connectivity_and_Secure_Communications/Advanced_aircraft_tracking_will_come_live_from_space`

[2] `https://www.polytechnique-insights.com/dossiers/espace/comment-les-satellites-low-cost-transforment-le-spatial/des-satellites-souverains-aux-satellites-as-a-service/`

online retail site Amazon, is positioned in two main space markets: space tourism and launch services [Weinzierl, 2018]. Another famous example is the founder of the online payment service Paypal and automotive company Tesla, Elon Musk, who entered the space sector in 2002 with SpaceX. The company produces space vehicles and provides commercial space transport services. In addition to being involved in space exploration projects towards the Moon and planet Mars, SpaceX is also developing a satellite constellation called Starlink to provide global high-speed Internet access service. These two cases illustrate the convergence between the digital and space sectors in New Space [Nardon, 2017].
Commercial considerations are central to recent technological advances in the space sector [Paulino, 2020]. Supported by a political will to reduce the costs of access to space, *New Space* players aim to cut the launch and production costs of space systems [Gudmundsson, 2018, Robinson and Mazzucato, 2019]. This results in the development of reusable launchers, the adoption of mass production models with standardized components, and the miniaturization of satellites [OECD, 2023, Quintana, 2017].

Thus, New Space symbolizes the growth of space activities guided by economic profitability and growth. Public decision-makers now prioritize expanding commercial civil markets in their space policy [Besha and MacDonald, 2016, Robinson and Mazzucato, 2019]. Space agencies are shifting their operational model, acting as anchor clients for newcomers, enabling their growth by formulating demand for space services, as highlighted at the beginning of this introduction with the example of ESA and Spire Global. Moreover, in 2021, the U.S. government published its 'United States Space Priorities Framework,' detailing the sector's future development directions. One of the priorities outlined is fostering a "policy and regulatory environment that enables a competitive and burgeoning U.S. commercial space sector."[3]. Among the commercial activities mentioned, the document emphasizes space technologies, space applications, and services leveraging space technology and infrastructure. Similarly, the European space program for 2021-2027 emphasizes maximizing "socioeconomic benefits to enable growth and job creation" as one of its primary goals. One of the clearly identified levers is the adoption and use of data, information, and services from European Earth observation programs such as Copernicus and the Galileo satellite positioning system[4].

[3] https://www.whitehouse.gov/wp-content/uploads/2021/12/united-states-space-priorities-framework-_-december-1-2021.pdf

[4] https://eur-lex.europa.eu/EN/legal-content/summary/eu-space-programme-2021-2027-european-union-agency-for-the-space-programme.html#

Consequently, the growing emphasis on commercial considerations leads to questioning the contribution of the space sector to the economy. To evaluate this, it is essential to collect economic data and implement suitable methods to gauge the significance of the space sector in economic activities [Hertzfeld, 2002]. In the particular context of New Space, where the intensity of commercial space activities is accelerating, and their nature is evolving, having such tools is crucial. Assessing the commercial space sector benefits both public decision-makers and commercial space stakeholders. The former can better target and apply effective space policies, while the latter can identify dynamic space markets and guide their investment decisions.

In this regard, the OECD Space Forum[5] plays an important role. Comprising space experts and a steering committee that includes representatives from space agencies of various countries[6], it works to harmonize the definitions of space activities and to formulate recommendations in terms of evaluation and statistical collection. At the national level, various initiatives stand out. The U.S. Bureau of Economic Analysis (BEA) has been developing the "Space Economy Satellite Account" since 2020, a statistical tool that estimates the contribution of the space sector to the American economy. Space agencies in Canada, the United Kingdom, and Australia publish annual reports on their countries' commercial space activities [Australian Space Agency, 2021, CSA, 2020, UK Space Agency, 2022]. For France, the National Institute of Statistics and Economic Studies (INSEE) recently published a study in partnership with CNES, mainly focusing on the space manufacturing segment[7].

It is in the context of these reflections on evaluating the space sector, in transformation since the advent of New Space, that this thesis is situated. It aims to enhance the understanding of developments within this sector and identify the main challenges related to defining the activities that compose it. Specifically, the thesis aspires to contribute to ongoing work on the methodological developments necessary to measure commercial space activities by proposing an evaluation tool tailored to a particular set of activities within the space sector.

[5] https://www.oecd.org/innovation/inno/space-forum/

[6] The committee includes: Agenzia Spaziale Italiana (ASI) (Italy), Centre National d'Etudes Spatiales (CNES) (France), Canadian Space Agency (CSA) (Canada), Deutsches Zentrum für Luft (DLR) (Germany), Korea Aerospace Research Institute (KARI) (Korea), National Aeronautics and Space Administration (NASA) (United States), Norwegian Space Agency (NOSA) (Norway), Netherlands Space Office (NSO) (Netherlands), Swiss Space Office (SSO) (Switzerland), United Kingdom Space Agency (UKSA) (United Kingdom), and ESA (Europe).

[7] https://www.insee.fr/fr/statistiques/6525061#consulter

In this thesis, we will focus on a set of activities particularly dynamic in the era of New Space, commonly referred to as the *downstream space segment* or *downstream space sector*. As the thesis title indicates, this segment refers to all commercial sectors that utilize space infrastructures, data, or signals in their activity. More precisely, OECD [2022b] defines the downstream space segment as follows: "Downstream space activities comprise the provision of products and services that rely on satellite signals or data, aimed at consumer and business markets." ESA provides the following definition: "Downstream means all those activities based on space technology, or using a space-derived system in a space or non-space environment, that may result in an application, product or service to the benefit of the (...) economy or society."[8] The downstream segment can also be defined relative to the upstream part of the space sector, which includes R&D activities, production of space systems, and launch services [OECD, 2022b].
Three main application areas are generally distinguished from characterizing the space services developed downstream. The first is the domain of satellite telecommunications, which includes, for example, satellite television broadcasting, satellite telephony services, and high-speed satellite services. The second area is satellite navigation, encompassing geopositioning services and navigation equipment. Finally, the last area is Earth observation, referring to activities measuring and monitoring the Earth through satellites. This includes activities such as selling images and developing services for natural resource and disaster management, climate monitoring, etc. [CSA, 2020].

This thesis aims to develop a methodology for assessing the economic weight of the downstream space segment and propose associated economic indicators. To do this, we develop an evaluation tool that considers the specificities of the downstream space segment and recent developments in the space sector as a whole, which we define in advance. We also implement an initial application of this tool to the French downstream space segment in 2021, presenting the main economic indicators developed.

Our research is structured into three main parts. Before any empirical and methodological approach, it is essential to analyze the subject of study by situating it in its context and examining the issues surrounding its evaluation. The **First**

[8] https://www.esa.int/About_Us/Corporate_news/Downstream_Gateway_bringing_space_down_to_Earth

Part, comprising Chapters 1 and 2, will examine in detail the characteristics of New Space and highlight the main motivations to develop a new evaluation method for the downstream space segment based on the analysis of existing studies. The **Second Part**, including Chapters 3 and 4, will be devoted to presenting our evaluation tool for the downstream space sector and the first application of the method to the case of the downstream space sector in France in 2021. Finally, in the **Third Part**, composed of Chapter 5, we will adopt a theoretical approach to delve further into the challenges surrounding the evaluation of the downstream space segment. Here, we will explore the specificities of the economic goods produced in the downstream space, namely information-based goods, to analyze the consequences on defining the downstream space value chain.

More precisely, we will organize our research in the following way.

Main objectives and contents of the thesis

In this thesis, we will associate the construction of a measurement method for the downstream space sector with a reflection on the activities that compose it and the nature of the goods produced therein. Indeed, any ambition to develop an evaluation tool must necessarily be accompanied by a conceptual framework that justifies the need for a new method and extensively analyzes the subject of the study. The first aspect will be addressed in the **First Part** of the thesis, while the second aspect will be further explored in the **Third Part** of the thesis. The **Second Part** presents the empirical aspect of the thesis, a central element of our contribution.

Part 1: Why evaluate the downstream space segment? Contextual Background and Methodological Challenges

In this first part, we will define the context in which our methodological work is integrated, describing it as one of a structural change in the space sector (**Chapter 1**). Then, we will present a literature review of a set of studies that evaluate the space sector to define the challenges of evaluating the downstream part of the sector (**Chapter 2**).

Chapter 1 – New Space and the Structural Change of the Space Sector

This first chapter aims to depict the context in which our thesis work occurs. We offer an analysis of New Space and argue that it corresponds to a structural change in the space sector where downstream activities play a fundamental role.

We rely on specialized literature in space and various institutions' publications on sector trends to describe New Space according to three major characteristics. First, we highlight the change in orientation of space policies [Robinson and Mazzucato, 2019] and the new role of agencies [Heracleous et al., 2019, Zervos and Siegel, 2008]. The second characteristic identified in this chapter is the entry of new players from the digital sector [Nardon, 2017]. Finally, the third characteristic of New Space we analyze is the emergence of new industrial models focused on space profitability [Davidian, 2020]. The description of these developments shows that New Space corresponds to accelerating commercial space development, the primary sources of which are the new market opportunities downstream of the value chain. The possibilities for commercial use of space infrastructure are extended with technological advances in the digital field.

We conclude this chapter by questioning the nature of the presented evolutions and proposing the hypothesis of an techno-economic paradigm shift in the space sector. We rely on several examples to illustrate this hypothesis and conclude by outlining the implications of this paradigm shift in terms of the industrial dynamics of the space sector. This initial reflection allows us to lay the foundations for the thesis's conceptual framework and thus guide the methodological development undertaken in the second part.

Chapter 2 – Estimating the Downstream Space Sector: Purpose and Literature Review

Space activities are not recognized as a distinct and unique sector of activity in the official classifications of economic activities [Hertzfeld, 2002, OECD, 2022b]. Although it is still possible to identify sectors that include manufacturing activities (e.g., aerospace and space construction) and launching services in reference nomenclatures, activities related to the operation of space infrastructure and data are more difficult to detect within these classifications. As mentioned earlier in the introduction, space agencies and organizations frequently publish evaluation reports to provide statistics on the space economy. Chapter 2 thus proposes a literature review where we focus on existing estimates of the downstream space segment. Given that the thesis aims to develop a methodology to measure this segment of activities, we explore whether existing studies might already offer suitable methodological tools for this assessment.

The analysis covers 25 studies published by 15 different sources. The publication dates range from 2009 to 2022. We select studies based on two criteria: 1) the scope of the evaluation includes the downstream space segment, and 2) the study provides an explanation, even if brief, of the measurement method used.
Our review has two distinct objectives. First, it aims to identify the main figures of the downstream space sector advanced in studies to have a preliminary representation of this segment. Results concerning the downstream segment, such as revenue by application areas, allow us to understand the importance of its evaluation better. Indeed, in cases where the assessment is for the entire space sector, revenues generated by downstream activities are higher than those from manufacturing [Booz & Company, 2014, UK Space Agency, 2021].

The second aspect of our review focuses on the methodological tools used to measure the downstream space segment. A detailed analysis of the studies allows us to identify significant limitations in how the downstream space segment is specifically evaluated. A primary methodological challenge concerns the very definition of the downstream space segment. Studies use the value chain tool to represent successive space activities. We identify differences in defining the perimeter of the downstream segment. These differences are particularly marked for the downstream limit of the value chain and lead to difficulties in establishing comparisons between countries for the same segment. A second major methodological challenge concerns the evaluation methods employed to measure the downstream space segment. Through analyzing this set of studies, we arrive at conclusions that we will take into account for our methodological development. Firstly, an evaluation of the downstream segment based solely on data collection from *known* actors by space experts does not allow the integration of the sector's recent developments and potential new entrants. Secondly, an evaluation of the downstream space segment based solely on an approach to identify the sectors involved in the downstream (as is the case for those using the input-output framework) leads to a partial identification of downstream space activities.

The lessons learned from this review of studies on the space economy serve as justification and a starting point for developing a methodology dedicated to estimating the downstream space segment.

Part 2: Methodology for Assessing the Downstream Space Sector and First Application for France in 2021

The second part of the thesis will be dedicated to developing the method for measuring downstream space activities and proposing associated economic indicators. The proposed evaluation approach will consist of two steps: a procedure for identifying the actors involved in downstream space activities (Chapter 3) and measuring the revenues generated by these actors (Chapter 4). Each chapter will include a detailed presentation of the approach and an application to the French downstream space segment in 2021.

Chapter 3 – Identification of French downstream space companies: a text-based approach

In this chapter, we propose identifying companies engaged in downstream space activities based on their mentions in the press. Specifically, we use a technique called Named-Entity Recognition to extract company names from their context of mention in digital articles dealing with downstream space activities.

Named-Entity Recognition (NER) is a text-mining task that involves detecting and classifying proper names in a text automatically [Liu et al., 2022, Maurya et al., 2022]. These proper names can be places, people, or organizations. Thus, we establish a dictionary of French company names and formulate a set of instructions (or ”rules”) to apply to press text to identify company names involved in downstream activities. Our NER approach is called the ”rule-based approach” [Chiticariu et al., 2010]. It is particularly suited to our subject of study since it relies on transparent rules that can easily be adjusted by the expert in charge of the evaluation, depending on the specifics of the sector and its future developments.

Since we are proposing a methodological tool, we are careful to describe each step of the method so that it can be easily reproduced at regular intervals and in different geographical areas. Then, to demonstrate the method's efficiency, we apply it to the French downstream space sector. From the initial results we will present in detail in the chapter, we describe the calibration steps and evaluate the identification method to ensure its effectiveness. We conclude the chapter with a set of considerations to take into account for the renewal of the method (reproduction intervals, adaptation of rules to the evaluation context).

Lastly, the proposed method is not an *ex nihilo* approach in that it can enrich existing databases of downstream space companies by identifying previously unknown companies. We propose a detection tool with some elements adjustable to the evaluation context. That is why its implementation requires the involvement of experts with knowledge of the trends that mark the space sector at the time of the evaluation.

Chapter 4 – Measuring the downstream space sector in France in 2021: Evaluation method and economic indicators

Chapter 4 addresses the measurement of revenues generated by the downstream space segment. The goal is to provide a strategy for assessing the economic weight of the downstream space sector based on the downstream space companies identified in the previous chapter.

We adopt a two-step procedure. Firstly, we elaborate a survey of downstream space companies to collect data on their revenue, the downstream activity segments to which they belong, and the portion of their revenue corresponding to these activities. We use this information to create a new categorization of downstream space companies *by circles of actors* and to develop a series of economic indicators on the revenues generated by these companies. The second step in the evaluation procedure consists of generalizing the survey results to all identified downstream space companies. We propose imputation methods to estimate missing revenues and assign a circle of actors to companies that would not participate in the survey.

As with the identification part, this chapter contains a detailed explanation section on the steps of the evaluation and the necessary indicators for reproduction, followed by a section devoted to application. We thus present the first indicators for measuring the downstream space segment in France for 2021.

Part 3: Theoretical Analysis of Value Creation Dynamics in the Downstream Space Sector

In this third and final section, we enrich our research on the economic assessment of the downstream space segment with a theoretical reflection on its value creation dynamics. Downstream space activities focus on the valuation of data, signals, and information either generated or transmitted by satellites. As a result, a significant part of the value produced in this sector is linked to data and information. We thus propose a discussion on the distinct properties of data and information as economic commodities in the age of digital transformation. This analysis provides us with avenues to refine the proposed method, incorporating the dynamics of data and information to understand the future shifts in the downstream space sector.

Chapter 5 – Data, Information, and Value Creation in the Digital Age

Downstream space activities involve exploiting space infrastructures or their outputs (e.g., data and signals) to produce information-based goods and services. In economic theory, information is a unique type of good characterized by non-rivalry and non-excludability [Arrow, 1962, Varian, 1999]. This chapter delves deeper into concepts from the economics of information. Firstly, we detail the primary properties of information and assess to what extent these can be applied to data as an economic commodity.

Secondly, we examine data and information within the new techno-economic paradigm dominated by digital technologies. We explore how the unique properties of data and information are amplified in this new context of producing and using increasingly varied and large-scale data.

Next, we consider data and information from a dynamic standpoint to understand the effects of their unique properties on value-creation processes. To achieve this, we borrow theoretical elements developed by Nicholas Georgescu-Roegen in his bio-economic approach to productive activities [Georgescu-Roegen, 1971]. We define data and information as particular flow factors within economic processes, which we term *conservative flow factors*. We then illustrate economic processes incorporating data and information, drawing from the fund-flow approach [Georgescu-Roegen, 1970, 1984].

Lastly, we discuss the implications of this theoretical development on our subject of study, the downstream space segment. Based on the survey results from downstream companies about how they produce or exploit space and non-space data in their operations and the uses of their products and services by their clients, we suggest avenues for future evaluations. These insights focus on considering the specificities of this sector and the evolution of the scope of actors to take into account.

Introduction générale

En juillet 2023, l'Agence Spatiale Européenne (ASE) a accordé un contrat de 16 millions d'euros au groupe Spire Global pour développer un système de surveillance indépendant de l'aviation civile basé depuis l'espace[9]. La société, créée en 2012, est spécialisée dans la collecte et l'analyse des données spatiales et dans la fourniture de services spatiaux. Elle développe sa propre constellation de nano-satellites d'observation de la Terre pour la surveillance des activités maritimes, de l'aviation et du climat à partir de l'espace. L'activité de cette société repose sur un modèle d'affaires émergent dans le secteur spatial appelé *SpaaS*. Cette nouvelle tendance, aussi désignée par le terme *SataaS*, marque une évolution dans l'industrie spatiale, où les entreprises privées délivrent un accès à des services et des usages à partir des infrastructures spatiales plutôt que les produits spatiaux eux-mêmes [Hein and Bruce Rosete, 2022]. Elle permet aux clients, institutionnels comme civils, de bénéficier des services offerts par la technologie satellite sans avoir à investir dans l'infrastructure physique[10]. Ainsi, le modèle SpaaS ouvre des opportunités de développement des usages civils commerciaux de l'espace considérables, s'alignant avec la transformation du secteur spatial initiée avec le *New Space*.

L'ère du *New Space* fait référence à un ensemble d'évolutions dans le secteur spatial initiées au début des années 2000 aux États-Unis [Denis et al., 2020, Pasco, 2017, Pekkanen, 2019]. Ce contexte nouveau constitue la toile de fond de la thèse, c'est pourquoi nous aurons l'occasion de l'analyser de manière approfondie dans notre développement. Le *New Space* renvoie à l'ouverture de l'accès à l'espace à de nouveaux acteurs privés, notamment des grandes entreprises et start-ups du secteur numérique. Les nouveaux entrants investissent le secteur spatial avec une logique

[9] `https://www.esa.int/Applications/Connectivity_and_Secure_Communications/Advanced_aircraft_tracking_will_come_live_from_space`

[10] `https://www.polytechnique-insights.com/dossiers/espace/comment-les-satellites-low-cost-transforment-le-spatial/des-satellites-souverains-aux-satellites-as-a-service/`

entrepreneuriale et de nouveaux modèles d'affaires reposant sur l'idée d'étendre les marchés commerciaux du spatial. À titre illustratif, la société Blue Origin créée en 2000 par le milliardaire américain Jeff Bezos, aussi fondateur du site de vente en ligne Amazon, est positionnée sur deux marchés principaux du spatial que sont le tourisme spatial et les services de lancement [Weinzierl, 2018]. Un autre exemple célèbre est le fondateur du service de paiement en ligne Paypal et de la société automobile Tesla, Elon Musk, qui entre dans le secteur spatial en 2002 avec l'entreprise SpaceX. L'entreprise produit des engins spatiaux et fournit des services commerciaux de transport spatial. En plus d'être impliquée dans des projets d'exploration spatiale vers la Lune et la planète Mars, SpaceX développe également une constellation de satellites appelée Starlink qui vise à fournir un service mondial d'accès à l'Internet haut débit. Ces deux cas illustrent la convergence entre les secteurs numériques et spatial dans le New Space [Nardon, 2017].
Les enjeux commerciaux jouent un rôle central dans les avancées technologiques du secteur spatial [Paulino, 2020]. Soutenus par une volonté politique de baisse des coûts d'accès à l'espace, les acteurs du *New Space* ont pour objectif de réduire les coûts de lancement et de production des systèmes spatiaux [Gudmundsson, 2018, Robinson and Mazzucato, 2019]. Cela se traduit par le développement de lanceurs réutilisables et l'adoption de modèles de production de masse avec des composants standardisés et la miniaturisation des satellites [OECD, 2023, Quintana, 2017].

Ainsi, le New Space symbolise le développement des activités spatiales en suivant une logique de rentabilité et de croissance économiques. Les décideurs publics établissent désormais l'expansion des marchés civils commerciaux comme une priorité de leur politique spatiale [Besha and MacDonald, 2016, Robinson and Mazzucato, 2019]. Les agences spatiales modifient leur modèle de fonctionnement, en servant de client d'ancrage des nouveaux entrants, leur permettant de se développer en leur formulant une demande de services spatiaux, comme évoqué en début d'introduction avec l'exemple de l'ASE et Spire Global. De plus, le gouvernement américain a publié en 2021 son programme 'Cadre des Priorités Spatiales des États-Unis' dans lequel il détaille les principaux axes de développement futur du secteur. L'une des priorités énoncées est la mise en place d'un « environnement politique et réglementaire favorisant le développement d'un secteur spatial commercial américain compétitif »[11]. Parmi les activités commerciales mentionnées, le document met en avant les technologies

[11] Traduction de l'auteur ; `https://www.whitehouse.gov/wp-content/uploads/2021/12/united-states-space-priorities-framework-_-december-1-2021.pdf`

spatiales, les applications spatiales, et les services s'appuyant sur la technologie et l'infrastructure spatiales. De la même manière, le programme spatial européen pour la période 2021-2027 met l'accent sur la maximisation des « bénéfices socio-économiques [liés à l'espace] pour favoriser la croissance et la création d'emplois »[12] comme l'un de ses principaux objectifs. Un des leviers clairement identifiés est l'adoption et l'utilisation des données, des informations, et des services issus des programmes européens d'observation de la Terre Copernicus et de système de positionnement par satellite Galileo[13].

Dès lors, la place croissante des considérations commerciales conduit à s'interroger sur la contribution du secteur spatial dans l'économie. Pour l'évaluer, il est nécessaire de collecter des données économiques et de mettre en place des méthodes adéquates afin d'estimer l'importance du secteur spatial dans les activités économiques [Hertzfeld, 2002]. Dans le contexte particulier du New Space, où l'intensité des activités spatiales commerciales s'accélère et leur nature évolue, disposer de tels outils est crucial. L'évaluation du secteur spatial commercial profite à la fois aux décideurs publics et aux acteurs commerciaux du spatial. Les premiers peuvent ainsi mieux cibler et appliquer des politiques spatiales efficaces, tandis que les seconds peuvent identifier les marchés spatiaux dynamiques et orienter leurs décisions d'investissement.

En ce sens, le Space Forum de l'OCDE[14] joue un rôle important. Composé d'experts du spatial et d'un comité de pilotage incluant les représentants d'agences spatiales de différents pays[15], il travaille à harmoniser les définitions des activités spatiales et à formuler des recommandations en termes d'évaluation et de récolte des statistiques. Au niveau national, différentes initiatives se démarquent. L'agence américaine de statistiques Bureau of Economic Analysis (BEA) développe depuis 2020 le « Space Economy Satellite Account », un outil statistique qui permet d'estimer la contribution du secteur spatial à l'économie américaine. Les agences spatiales du Canada, du Royaume-Uni et de l'Australie publient des rapports annuels sur les activités spatiales commerciales de leur pays [Australian Space Agency, 2021, CSA, 2020, UK Space Agency, 2022]. Pour la France, l'agence nationale de la statistique

[12] Traduction de l'auteur.

[13] `https://eur-lex.europa.eu/EN/legal-content/summary/eu-space-programme-2021-2027-european-union-agency-for-the-space-programme.html#`

[14] `https://www.oecd.org/innovation/inno/space-forum/`

[15] Le comité inclut : ASI (Italie), CNES (France), CSA (Canada), DLR (Allemagne), KARI (Corée), NASA (États-Unis), NOSA (Norvège), NSO (Pays-Bas), SSO (Suisse), UKSA (Royaume-Uni), et ESA (Europe).

(INSEE) a récemment publié une étude en partenariat avec le CNES, se concentrant principalement sur le segment manufacturier du spatial[16].

C'est dans le contexte de ces réflexions sur l'évaluation du secteur spatial, en pleine transformation depuis le New Space, que s'inscrit cette thèse. Elle vise à améliorer la compréhension des évolutions au sein de ce secteur, et à identifier les principaux enjeux liés à la définition des activités qui le composent. Plus précisément, la thèse aspire à contribuer aux travaux en cours sur les développements méthodologiques nécessaires à la mesure des activités spatiales commerciales, en proposant un outil d'évaluation adapté à un ensemble particulier d'activités du secteur spatial.

Dans cette thèse, nous nous concentrerons sur un ensemble d'activités particulièrement dynamiques à l'ère du New Space, communément désignées par l'expression *segment* ou *secteur spatial aval*. Comme l'indique le titre de la thèse, ce segment renvoie à l'ensemble des secteurs commerciaux qui exploitent les infrastructures, données ou signaux d'origine spatiale dans leur activité. Plus précisément, OECD [2022b] définit le segment spatial aval de la manière suivante : « Downstream space activities comprise the provision of products and services that rely on satellite signals or data, aimed at consumer and business markets. » L'ASE en donne la définition suivante : « Downstream means all those activities based on space technology, or using a space-derived system in a space or non-space environment, that may result in an application, product or service to the benefit of the (...) economy or society. »[17]
Le segment aval peut aussi se définir par opposition à la partie amont du secteur spatial, qui comprend les activités de R&D, la production des systèmes spatiaux et les services de lancement [OECD, 2022b].
On distingue généralement trois grands domaines d'applications pour caractériser les services spatiaux développés en aval. Le premier est le domaine des télécommunications par satellite, qui inclut notamment la diffusion de contenus télévisés par voie satellitaire, les services de téléphonie par satellite, et les services de haut débit par satellite. Le second domaine est la navigation par satellite, qui comprend les services de géopositionnement et les équipements de navigation. Enfin, le dernier domaine est l'observation de la Terre, qui renvoie aux activités de mesure et de surveillance de la Terre grâce aux satellites. Cela inclut des activités telles que la vente d'images

[16] `https://www.insee.fr/fr/statistiques/6525061#consulter`
[17] `https://www.esa.int/About_Us/Corporate_news/Downstream_Gateway_bringing_space_down_to_Earth`

et le développement de services pour la gestion des ressources et des catastrophes naturelles, la surveillance du climat, etc. [CSA, 2020].

L'objectif général de ce travail de thèse est la construction d'une méthodologie d'évaluation du poids économique du segment spatial aval et la proposition d'indicateurs de mesure associés. Pour cela, nous élaborons un outil d'évaluation qui prend en compte les spécificités du segment spatial aval ainsi que les évolutions récentes du secteur spatial dans son ensemble, que nous définissons au préalable. Nous mettons également en œuvre une première application de cet outil sur le segment spatial aval en France en 2021, en présentant les principaux indicateurs économiques élaborés.

Notre recherche se structure en trois grandes parties. Avant toute démarche empirique, il est indispensable d'analyser l'objet d'étude en le situant dans son contexte et en examinant les enjeux autour de son évaluation. La **Première Partie** de la thèse, comprenant les chapitres 1 et 2, examinera en détail les caractéristiques du New Space et soulignera les principales motivations à développer une nouvelle méthode d'évaluation pour le segment spatial aval à partir de l'analyse des études existantes. La **Deuxième Partie**, incluant les chapitres 3 et 4, sera consacrée à la présentation de notre outil d'évaluation du secteur spatial aval ainsi qu'à la première application de la méthode au cas du secteur spatial aval en France en 2021. Enfin, dans la **Troisième Partie**, composée du chapitre 5, nous adopterons une approche théorique afin d'approfondir les enjeux autour de l'évaluation du segment spatial aval. Nous y explorerons les spécificités des biens économiques produits dans le spatial aval, à savoir des biens basés sur la donnée et l'information, afin d'en analyser les implications sur la définition de la chaîne de valeur du segment spatial aval.

Plus précisément, nous allons organiser notre recherche de la manière suivante.

Principaux objectifs et contenus de la thèse

Dans cette thèse, nous allons tenter d'associer la construction d'une méthode de mesure du secteur spatial aval à une réflexion sur les activités qui la composent et sur la nature des biens qui y sont produits. En effet, toute ambition de développement d'un outil d'évaluation doit nécessairement s'accompagner d'un cadre conceptuel qui justifie le besoin d'une nouvelle méthode et qui analyse de manière approfondie l'objet de l'étude. Le premier aspect sera abordé dans la **Première Partie** de la thèse, tandis que le second aspect sera davantage exploré dans la **Troisième Partie** de la thèse. La **Deuxième Partie** présente quant à elle le volet empirique de la thèse, élément central de notre contribution.

Partie 1 : Pourquoi évaluer le segment spatial aval ? Contexte et enjeux méthodologiques

Dans cette première partie, nous définirons le contexte dans lequel s'intègre notre travail méthodologique en le décrivant comme celui d'un changement structurel du secteur spatial (**Chapitre 1**). Ensuite, nous présenterons une revue de littérature des principales études qui traitent de l'évaluation du secteur spatial pour définir les enjeux de l'évaluation de la partie aval du secteur (**Chapitre 2**).

Chapitre 1 – Le New Space, une transformation structurelle du secteur spatial

Ce premier chapitre vise à dépeindre le contexte dans lequel intervient notre travail de thèse. Nous proposons une analyse du New Space et soutenons qu'il correspond à un changement structurel du secteur spatial où les activités aval jouent un rôle fondamental.

Nous nous appuyons sur la littérature spécialisée dans le domaine spatial ainsi que sur différentes études publiées par les institutions concernant les tendances du secteur pour décrire le New Space selon trois grandes caractéristiques. La première que nous mettons en évidence est le changement d'orientation des politiques spatiales [Robinson and Mazzucato, 2019] et le nouveau rôle des agences [Heracleous et al., 2019, Zervos and Siegel, 2008]. La seconde caractéristique identifiée dans ce chapitre est l'entrée de nouveaux acteurs issus du secteur numérique [Nardon, 2017]. Enfin, la troisième caractéristique du New Space que nous analysons est l'émergence de nouveaux modèles industriels axés sur la rentabilité de l'espace [Davidian, 2020]. La description de ces évolutions nous permet de mettre en évidence que le New Space correspond à une accélération du développement commercial de l'espace, dont les sources principales sont les nouvelles opportunités de marché en aval de la chaîne de valeur. Les possibilités d'usages commerciaux des infrastructures spatiales sont étendues avec les avancées technologiques dans le domaine du numérique.

Nous concluons ce chapitre en nous interrogeant sur la nature des évolutions présentées et en émettant l'hypothèse d'un changement de paradigme techno-économique dans le secteur spatial. Nous nous appuyons sur plusieurs exemples pour illustrer cette hypothèse et terminons en exposant les implications de ce changement de paradigme en termes de dynamique industrielle du secteur spatial. Cette première réflexion nous permet de poser les bases du cadre conceptuel de la thèse et guidera ainsi le développement méthodologique entrepris dans la deuxième partie.

Chapitre 2 – Estimer le secteur spatial aval : objectif et revue de littérature

Les activités spatiales ne sont pas reconnues comme un secteur d'activités distinct et unique dans les classifications officielles des activités économiques [Hertzfeld, 2002, OECD, 2022b]. S'il est tout de même possible d'identifier les secteurs qui comprennent les activités manufacturières (e.g. construction aéronautique et spatiale) et les services de lancement dans les nomenclatures de référence, les activités d'exploitation des infrastructures et des données spatiales sont plus difficiles à détecter au sein de ces nomenclatures. Comme nous l'avons mentionné plus haut dans l'introduction, un certain nombre d'études sont publiées par les agences spatiales et des organismes publics ou privés afin de proposer des statistiques sur l'économie spatiale. Le Chapitre 2 propose donc une revue de la littérature où nous nous concentrons sur les estimations du segment spatial aval existantes. Puisque l'objectif de la thèse est de développer une méthodologie de mesure de ce segment d'activités, nous nous demandons si les études existantes ne fournissent pas déjà des outils méthodologiques adaptés permettant cette évaluation.

L'analyse porte sur 25 études publiées par 15 sources différentes. Les dates de publication vont de 2009 à 2022. Nous sélectionnons les études selon deux critères : 1) le champ de l'évaluation comprend le segment spatial aval et 2) l'étude fournit une explication, même succinte, de la méthode de mesure employée. Notre revue repose sur deux objectifs distincts. Premièrement, elle vise à identifier les principaux chiffres du secteur spatial aval avancés dans les études afin d'avoir une première représentation de ce segment. Les résultats concernant le segment aval, qui rendent compte par exemple des revenus en fonction des domaines d'application, nous permettent d'avoir une meilleure compréhension des enjeux autour de son évaluation. En effet, dans les cas où l'ensemble du secteur spatial est évalué, on observe que les recettes générées par les activités aval sont supérieures aux recettes issues de l'industrie manufacturière

[Booz & Company, 2014, UK Space Agency, 2021].

Le second aspect de notre revue porte sur les outils méthodologiques employés pour mesurer le segment spatial aval. L'analyse approfondie des études nous permet d'identifier des limites importantes dans la manière d'évaluer le segment spatial aval spécifiquement. Une premier défi méthodologique concerne la définition même du segment spatial aval. Les études utilisent l'outil de la chaine de valeur pour représenter les activités spatiales successives. Nous identifions des différences dans la définition du périmètre du segment aval. Ces différences sont surtout marquées pour la limite aval de la chaine de valeur, et impliquent des difficultés à établir des comparaisons entre pays pour un même segment. Un second défi méthodologique majeur concerne les méthodes d'évaluation employées pour mesurer le segment spatial aval. A travers l'analyse de cet ensemble d'études, nous aboutissons à des conclusions dont nous tiendrons compte pour notre propre développement méthodologique. Premièrement, une évaluation du segment aval basée uniquement sur la collecte de données des acteurs *connus* par les auteurs experts du spatial ne permet pas d'appréhender les évolutions récentes du secteur et les potentiels nouveaux entrants. Deuxièmement, une évaluation du segment spatial aval qui repose uniquement sur une approche d'identification des secteurs d'activités impliqués dans l'aval (comme c'est le cas pour celles utilisant le cadre input-output) mène à une identification partielle des activités spatiales aval.

Les enseignements tirés de cette revue des études sur l'économie spatiale servent de justification et de point de départ au développement d'une méthodologie dédiée à l'évaluation du secteur spatial aval. En effet, cette analyse nous permet d'adopter une approche exploratoire qui combine l'expertise de l'évaluateur et l'utilisation des classification industrielles afin de développer une méthode originale d'identification des entreprises spatiales aval.

Partie 2 : Méthodologie d'évaluation du secteur spatial aval et première application pour la France en 2021

La deuxième partie de la thèse sera dédiée au développement de la méthode de mesure des activités spatiales aval et la proposition d'indicateurs économiques associés. La démarche d'évaluation proposée comportera deux étapes : une procédure d'identification des acteurs impliqués dans des activités spatiales aval (Chapitre 3), et la mesure des revenus générés par ces acteurs (Chapitre 4). Chacun des chapitres inclura une partie de présentation détaillée de la démarche, et une partie d'application au segment spatial aval français en 2021.

Chapitre 3 – Identification des entreprises spatial aval françaises : une approche basée sur le texte

Dans ce chapitre, nous proposons d'identifier les entreprises ayant des activités spatiales aval à partir de leur citation dans la presse. Plus précisément, nous utilisons une technique appelée Named-Entity Recognition pour extraire les noms d'entreprises à partir de leur contexte de citation dans des articles numériques traitant des activités spatiales aval.

La Named-Entity Recognition (NER) est une tâche de text-mining qui consiste à repérer puis classer des noms propres dans un texte de façon automatisée [Liu et al., 2022, Maurya et al., 2022]. Ces noms propres peuvent être des lieux, des personnes ou des organisations. Ainsi, nous établissons un dictionnaire de noms d'entreprises françaises et formulons un ensemble d'instructions (ou *règles*) à appliquer au texte de presse pour identifier les noms d'entreprises impliquées dans des activités aval. Cette approche particulière de la NER est appelée *rule-based approach* [Chiticariu et al., 2010]. Elle est particulièrement adaptée à notre objet d'étude puisqu'elle repose sur des règles transparentes et facilement ajustables par l'expert en charge de l'évaluation en fonction des spécificités du secteur et de ses évolutions futures.

Puisqu'il s'agit de proposer un outil méthodologique, nous nous appliquons à décrire chaque étape de la méthode de façon à ce qu'elle soit aisément reproductible à intervalles réguliers et dans différentes zones géographiques. Ensuite, pour démontrer l'efficacité de la méthode, nous l'appliquons au secteur spatial aval français. À partir des premiers résultats que nous présenterons de manière détaillée dans le chapitre, nous décrivons les étapes de calibrage et l'évaluation de la méthode d'identification pour nous assurer de son efficacité. Nous terminons le chapitre par un ensemble de considérations à prendre en compte pour le renouvellement de la méthode (intervalles de reproduction, adaptation des règles au contexte d'évaluation).

Enfin, la méthode proposée ne repose pas sur une démarche *ex nihilo*. Elle permet d'enrichir des bases de données d'entreprises spatiales aval existantes en identifiant des entreprises non connues auparavant. De plus, nous avons pris soin de proposer un outil de détection dont certains éléments sont ajustables au contexte d'évaluation.

C'est la raison pour laquelle sa mise en place suppose l'implication d'experts ayant une connaissance des tendances qui marquent le secteur spatial au moment de l'évaluation.

Chapitre 4 – Mesurer le secteur spatial aval en France en 2021 : méthode d'évaluation et indicateurs économiques

Le chapitre 4 aborde la mesure des revenus générés par le segment spatial aval. L'objectif est de fournir une stratégie d'évaluation du poids économique du secteur spatial aval à partir des entreprises spatiales aval identifiées dans le chapitre précédent.

Nous adoptons une procédure en deux étapes. Premièrement, nous élaborons une enquête auprès des entreprises spatiales aval pour collecter des données sur leur chiffre d'affaires, les segments d'activités aval auxquels elles appartiennent, et la part de leur chiffre d'affaires qui correspond à ces activités. Nous utilisons ces informations pour construire une nouvelle catégorisation des entreprises spatiales aval *par cercles d'acteurs*, et élaborer une série d'indicateurs économiques sur les recettes générées par ces entreprises. La deuxième étape de la procédure d'évaluation consiste à généraliser les résultats de l'enquête à l'ensemble des entreprises spatiales aval identifiées. Nous proposons des méthodes d'imputation pour estimer les chiffres d'affaires manquants, puis pour attribuer un cercle d'acteurs aux entreprises qui ne participeraient pas à l'enquête.

Comme pour la partie d'identification, ce chapitre contient une section d'explications détaillées des étapes de l'évaluation et des indicateurs indispensables à la reproduction, suivie d'une section consacrée à l'application. Nous présentons ainsi les premiers indicateurs de mesure du segment spatial aval en France pour l'année 2021.

Partie 3 : Analyse théorique de la dynamique de création de valeur dans le secteur spatial aval

Dans cette troisième et dernière partie, nous complétons notre recherche sur l'évaluation du poids économique du segment spatial aval par une réflexion théorique sur la dynamique de création de valeur dans ce segment. Les activités spatiales aval reposent sur la valorisation des données, signaux et informations générés ou transmis par les satellites. Ainsi, une partie significative de la valeur produite dans ce secteur est liée à la donnée et à l'information. Nous proposons ainsi d'ouvrir une réflexion sur les propriétés particulières de la donnée et de l'information comme biens économiques à l'ère de la transformation numérique. Cette analyse nous permet d'ouvrir des pistes d'approfondissement de la méthode proposée qui impliquent la dynamique de la donnée et de l'information dans la compréhension des évolutions futures du secteur spatial aval.

Chapitre 5 – Donnée, information et création de valeur à l'ère du numérique

Les activités spatiales aval impliquent l'exploitation des infrastructures spatiales ou de ses extrants (e.g. données et signaux) pour produire des biens et services basés sur l'information. Dans la théorie économique, l'information est un bien à la nature particulière présentant les caractéristiques de non-rivalité et de non-appropriabilité [Arrow, 1962, Varian, 1999]. Ce chapitre propose un travail d'approfondissement des concepts issus de l'économie de l'information. Premièrement, nous détaillons les principales propriétés de l'information et analysons dans quelle mesure celles-ci s'appliquent à la donnée comme bien économique.

Deuxièmement, nous analysons la donnée et l'information dans le nouveau paradigme techno-économique dominé par les technologies numériques. Nous explorons comment les propriétés particulières de la donnée et de l'information sont amplifiées dans ce nouveau contexte de production et d'utilisation de données de plus en plus variées et en grand volume.

Ensuite, nous considérons la donnée et l'information dans une perspective dynamique pour appréhender les effets de leurs propriétés particulières sur les processus de création de valeur. Pour cela, nous empruntons les éléments théoriques développés par Nicholas Georgescu-Roegen dans son approche bio-économique des activités productives[Georgescu-Roegen, 1971]. Nous définissons les données et l'information comme des facteurs flux particuliers au sein des processus économiques, que nous appelons *facteurs flux conservatifs*. Nous étayons notre analyse par une représentation des processus économiques qui intègrent de la donnée et de l'information à partir de l'approche flux-fond [Georgescu-Roegen, 1970, 1984].

Enfin, nous discutons des implications de ce développement théorique sur notre objet d'étude, le segment spatial aval. A partir des résultats de l'enquête aux entreprises aval portant sur la manière dont elles produisent ou exploitent les données d'origine spatiale et non-spatiale dans leur activité et sur les usages de leur biens et services par leurs clients, nous formulons des pistes d'approfondissement pour les évaluations futures. Ces réflexions portent sur la prise en compte des spécificités de ce secteurs et l'évolution du périmètre des acteurs à considérer.

Part I.
Why Evaluate the Downstream Space Segment? New Space and Methodological Challenges

Chapter 1

New Space and the Structural Change of the Space Sector

1.1 Introduction

At the 14th EU Space Conference in 2022, Commissioner Thierry Breton declares:

> "The space sector is undergoing a massive transformation. [...] The booming of private operators changes the business model of space, combining both large and small industry, space and digital ecosystems. This is a major opportunity for Europe."

This excerpt synthesizes the main markers of New Space that we describe in the present chapter. The first marker is the decisive role of States in the transformation of space activities (Section 2). Thierry Breton speaks of a *major opportunity for Europe.* Indeed, we highlight in this chapter that New Space is associated with the reorientation of space policies toward the development of commercial space. Public support to the sector, which is still massive, is no longer only driven by political or scientific objectives. Space and its commercial applications are perceived as a lever for economic growth.

The second marker is the democratization of space activities with the entry of new private actors in the early 2000s (Section 3). We identify from the literature on New Space the main characteristics of these private investors and their motivation for entering the space sector. We analyse how these actors are changing *the business model of space* to expand civil space markets.

This fundamental aspect of New Space will be discussed throughout the chapter. The digital revolution, the expansion of big data and the associated technological advances (AI, cloud computing, etc.) provide the space sector new market opportunities based on the exploitation of space data [Nardon, 2017, Vidmar, 2020].

This chapter concludes with a theoretical reflection to understand and characterize this *massive transformation* of space activities (Section 4). We assume that a paradigm shift has occurred in the space sector toward a demand-pull model. We affirm that this structural change has implications both in terms of industrial and data dynamics. This reflection lays the foundations for a conceptual framework to be considered in the development of a new evaluation methodology.

1.2 Redefinition of the role of space agencies and market-oriented space policies

1.2.1 States and space agencies' support for the development of commercial space

New Space corresponds to the opening up access to space to private companies. Governments and space agencies have supported this trend with the progressive and worldwide privatization of space activities, the change of contractual practices, and the reorientation of space policies toward the development of commercial space applications.

Historically, the space sector was the domain of states. In the spacefaring nations, governments were both contractors and exclusive users of technologies, providing technical requirements to manufacturers [Pasco, 2017]. Until 1982 in the US, NASA and the Department of Defense (DoD) were the only customers for launch vehicles manufacturers. The government handled the launch of civilian and commercial capabilities, and most satellites were owned by the agencies [Canis, 2016]. The 1984 US Commercial Launch Space Act, which facilitates the development of private-sector launch services, constitutes an important milestone in the shift of national space policies. On the European side, Landoni and dt ogilvie [2019] report similar political support for liberalizing the space industry. The authors describe the process of consolidation and privatization of initially state-owned industrial leaders (Aerospatiale in France, Alenia Spazio in Italy, and British Aerospace in the UK) between 1960 and 2000. More recently, the Indian government announced the opening of the space industry, previously under the exclusive control of the national space agency, to private enterprise [Rohera, 2021].

The shift in government space policies is marked by an evolution in agencies' procurement models. Under New Space, public bodies are no longer formulating demand for technology, but rather a demand for turnkey services. The Commercial Resupply Services and Commercial Crew programs are prime examples of this [Heracleous et al., 2019]. The first program was initiated in 2006 by NASA to subcontract unmanned cargo and supply delivery missions to the International Space Station (ISS) to private companies such as SpaceX and Orbital Sciences. The Commercial Crew Program, which began in 2010, contracts out manned missions to the ISS to Boeing and SpaceX. In this new mode of contracting, which mainly takes the form of public-private partnerships, space technology no longer belongs to the States but to the industry. Cost-sharing between the agency and industry replaces the prevailing cost-plus financing model, which allowed full risk coverage for manufacturers [Canis, 2016, Heracleous et al., 2019]. The development of public-private partnerships also highlights a new positioning of institutions in the space sector. Formerly exclusive clients, their role is now to support space activities and to address market failures in terms of innovation [Zervos and Siegel, 2008, Landoni and dt ogilvie, 2019].

1.2.2 Space policies for the expansion of space applications and services

Once focused on security and sovereignty concerns, space policies are formulating new objectives based on civil market creation [Besha and MacDonald, 2016, Canis, 2016, Robinson and Mazzucato, 2019]. National support for the development of commercial applications is one of the main thrusts of these market-oriented policies.
In Australia and the United Kingdom, the space sector is perceived as a major lever for economic growth and employment. In both countries, space policies are focused on expanding the satellite industry and its application markets rather than the spacecraft industry. The Australian Space Agency has advanced as the top 3 "National Civil Space Priorities" of its civil space strategy: Position, navigation and timing; Earth observation; and Communications technologies and services [Australian Space Agency,

2019]. These three segments account for $867 million in investment, or 43% of the total investment[18] over the 2019-2028 period [Australian Space Agency, 2021]. In the UK, more than 70% of space sector revenues are generated by the downstream applications segment (direct-to-home broadcasting, user equipment, fixed satellites communications services, location-based services) [UK Space Agency, 2021]. The first recommendation made in the *Space Innovation and Growth Strategy 2014-2030* report[19] is to "develop [...] new space applications by promoting the benefits of Space to business and Government and engaging service providers." [Space IGS, 2014]. This objective was notably pursued with the establishment of the Satellite Catapult Centre in 2013 in Harwell. The organization aims to facilitate the use of space technologies by space and non-space private companies and to support them in the development of commercial space applications.
In the United States, support for the privatization of space transportation is perhaps the most visible aspect of the shift in space policy [Mazzucato and Robinson, 2018]. NASA intends to leave the further development of low-Earth orbit markets such as telecommunications and Earth observation to private players by reducing regulatory barriers [Besha and MacDonald, 2016]. In this respect, the expansion of commercial space services has been encouraged since the 1990s with the authorization to exploit high-resolution satellite images for commercial purposes [Pasco, 2019]. NASA's support for the economic development of low-Earth orbit has contributed to the emergence of downstream space start-ups such as Planet, which hold dominant positions in the satellite data market.

Eventually, the space policies conducted in Europe also reflect the desire to develop civil and commercial uses of space with increasing support to space companies and start-ups. Policy briefs and reports published by European institutions point out the economic potential of space technologies and data in terms of civil applications and services [PwC France, 2019, ESA, 2016, European Commission, 2013]. In the late 1990s, the European Union (EU) launched two major space programmes: the Galileo satellite navigation system and the Copernicus Earth Observation programme. The Copernicus programme provides free and open access to data and geo-information services. The objective of the Copernicus Open Data Policy is to promote the civil use of EO data in areas such as land and ocean monitoring, climate change, and

[18] 62% of the budget is public, with the remainder coming from the private sector and international space agencies.

[19] This work was prepared by the Space Innovation and Growth Team formed in 2009 as a joint initiative between the government, academia, and the space industry.

Table 1.1: Horizon 2020 Space budget by topic

Topic	Budget (in million euros)	Share in total budget
Technology, science, and exploration	542.4	57%
Earth observation - Copernicus	**204.6**	**22%**
Satellite navigation - Galileo	**146.9**	**15%**
Space business, entrepreneurship, and education	8.3	1%
Access to space, secure and safe space environment	50.5	5%
Total	952.7	100%

Source: French Ministry of Higher Education

emergency management [Jutz and Milagro-Pérez, 2020]. Operational since 2016, the European Global Satellite Positioning System Galileo is an open service designed for civilian use. Strong efforts are made in the EU to support the development of civil applications from the two programmes. The EU Horizon 2020 framework programme for research and innovation has funded 450 space-related projects for a budget of €953 million between 2014 and 2020. Table 1.1 shows that nearly 40% of Horizon 2020 Space funding is won by projects developing Copernicus and Galileo-based services. The ESA, which initially focused on R&D and coordination of the EU space industry, is expanding its missions to promote downstream space applications and services [Robinson and Mazzucato, 2019]. The agency has implemented programmes such as Integrated Applications Promotion [Lebeau et al., 2013], and ESA Space Solutions, which includes the ESA Business Applications Programme and Business Incubation Centres. These programmes provide technical and financial support to entrepreneurs or companies in the development of products or services that use space assets.

In addition to the European measures, we observe in France the same strategy to develop commercial space with intensive public support to space start-ups. This objective is first visible within the French space agency CNES, with the creation in 2022 of the New Space sub-department and the Space Observatory within the

Strategy Department. The observatory aims to analyze economic changes within the space sector and help companies identify dynamic commercial markets. Second, the economic recovery plan "France Relance" initiated in 2020 after the health crisis contains a space component that provides €365 million of additional investment in the space sector. The French space agency CNES (Centre National d'Etudes Spatiales) operates the plan and has implemented two schemes primarily intended for space start-ups and SMEs, both in upstream and downstream segments. The first is a call for tenders addressed to small and medium-sized companies that develop dual-use space technologies (for civil and military applications). The second is a call for projects called "Pitch Days" to finance space applications in specific domains defined by French Regional Councils such as mountain development, maritime, land management, logistics, and agroecological transition [20]. In September 2021, 70% of the space sector's recovery plan beneficiaries were start-ups and SMEs. This example and the Horizon 2020 programme mentioned above illustrate two elements of change in space policy. First, the industrial strategy extends to the downstream part of the sector with the promotion of space applications. Second, the ambition for France and Europe to maintain their competitiveness in the industry against US New Space companies by supporting small businesses and start-ups through public procurement.

The following section describes the main characteristics of new space entrants.

1.3 New business and industrial dynamics driven by commercial uses of space

1.3.1 New space entrants: characteristics and motivation

The fast spread of digital technologies occurring since the late 1990s has disrupted most industries, lowering barriers to entry, increasing business dynamics and introducing new business and industrial models [Calvino and Criscuolo, 2019]. The space sector is no exception to this trend. The move into the New Space era is associated with the entry of private players intending to develop the commercial potential of space. For these new players, the convergence between the digital industry and satellite technologies is a significant growth opportunity. New space entrants are of two types: already existing private companies - mostly coming from the ICT sector -

[20]The information is available here: `https://www.connectbycnes.fr/en/space-tour-2021`

and new space start-ups [Denis et al., 2020].

The first category of entrants are large companies operating in information technologies, big data, and the Internet industry. They are founded by entrepreneurs with substantial investment capacities and dominant market position in their core market area [Denis et al., 2020, Pasco, 2017]. Their core business is not *a priori* linked with space, the digital champions enter the space sector to diversify their activities. They have specific characteristics that may explain their incentive to invest in a high-tech, high risk and capital intensive industry such as space [European Commission, 2013]. Most of them are high-tech multinationals with a highly skilled workforce, well established in their sector and generating significant revenues. These companies introduce an innovation culture based on an entrepreneurial logic, reinvesting a large share of their income to R&D activities for both technological and industrial improvements [Nardon, 2017]. This first wave of new space players is embodied by Jeff Bezos with Blue Origin and Elon Musk with SpaceX. Both companies provide commercial spacecraft and space transportation services. Before that, their founders were leaders in their respective markets: the online Amazon bookstore service and the online payment services Paypal. When they entered the sector in the early 2000s, both players aimed to reduce space transportation costs to the International Space Station. Another example is digital giant Google acquiring Earth Observation (EO) satellite operator Skybox Imaging in 2014. The company is sold to Planet Labs in 2017 in exchange for supplying EO data to Google.

The second category of entrants is new space start-ups. From 2000 to 2018, more than two hundred space companies were created worldwide, corresponding to approximately twenty-two billion dollars invested in space start-ups [Bryce Space & Technology, 2019a]. In the early 2000s, the number of new ventures created was steady and relatively low (around four new space ventures per year) but increased by 55% in 2009-2010. We believe the substantial growth in the number of new entrants coincides with the first successful launch of SpaceX. The success story of SpaceX sent a positive signal to private U.S. investors. More generally, the first wave of entrants have democratized access to space and demonstrated the potential associated with its commercial exploitation. The second wave of entrants subsequently benefited from this weakening of barriers to entry to space. The French start-up Unseenlabs illustrates this trend. Founded in 2015, the company is developing a constellation of Earth observation nano-satellites for maritime surveillance. Its first satellite was deployed in 2019, only four years after the company was created. In 2021, the company raised 20 million euros both from public and private investors: the investment fund of the French Ministry of Defence, and private venture capitalists Omnes Capital and 360 Capital.[21]

Another feature of New Space, in line with the rise of space start-ups, is the growing interest of private investors in space [Bryce Space & Technology, 2022, ESPI, 2019, Dedieu et al., 2016]. According to Bryce Space & Technology [2022], 52 billion dollars were invested in space startups betwen 2000 and 2016, with 70% since 2016[22]. Historically, the demand for space systems came primarily from institutional organizations and public project owners and thus relied heavily on public funds [Lebeau, 2008]. While this is still the case, particularly in Europe with over 60% of industry activity funded by institutional programs [Eurospace, 2020], the space sector is attracting private investors such as venture capital funds, private equity firms, angel investors, corporations, and commercial banks [Bryce Space & Technology, 2019a, OECD, 2019]. This diversification of funding sources reflects private investors' confidence in the ability of the space sector to generate significant returns. Moreover, the risk taken by investor is mitigated by government guaranties and the importance of public procurement.

Private investment in space startups remains concentrated in terms of volume in the U.S and largely impelled by billionaire angel investors from the digital industry such as Jeff Bezos (Blue Origin), Richard Branson (Virgin Galactic), Elon Musk (SpaceX), Bill Gates (Kymeta), and Paul Allen (StratoLaunch) [OECD, 2019]. However, space is attracting more and more investors: 596 investors provided funding to startups in 2021, 63% of them are new investors [Bryce Space & Technology, 2022]. Private funding of new space ventures is also gaining ground in China and Europe, to a lesser extent and with the support of public authorities. Between 2014 and 2020, 1.3 billion euros were invested in European space startups, 70% of which by venture capital funds (from a mix of private and public funding) [ESPI, 2021].

At first glance, the segment that benefits most from private capital is the upstream, with more than two-thirds of global investment ([ESPI, 2019, Dedieu et al., 2016]). The

[21]The information is available here: `https://www.latribune.fr/entreprises-finance/industrie/aeronautique-defense/newspace-unseenlabs-dechire-avec-une-levee-de-fonds-20-millions-d-euros-883357.html`

[22]These figures may be underestimated, especially for non-U.S countries, as information on private investment is difficult to track.

launch industry and small satellites and cubesats are the segment where most private capital is concentrated [Vernile, 2018]. This suggests that most new entrants are specialized in activities such as launch operations and satellite system manufacturing. However, this observation must be qualified. First, most existing studies on space markets indicate the existence of a statistical bias regarding the measurement of downstream activities and the difficulty of precisely identifying the companies that belong to this segment ([ESPI, 2019, PwC France, 2019]). The second aspect refers to an additional and crucial feature of New Space for our research area: the shift to vertical integration of space companies and the concentration of economic value in the downstream part of the space value chain ([OECD, 2014, Delponte et al., 2016, Nardon, 2017, Robinson and Mazzucato, 2019]. This point will be further discussed in the following section.

1.3.2 Emergence of new business and industrial models

Before New Space, the prevailing industrial model was based on high technological complexity and long development cycles to ensure the reliability of space systems [Robinson and Mazzucato, 2019]. As with large network infrastructures, the development costs of space technologies were massive and could only be supported by national governments [OECD, 2012, Lebeau, 2008]. The production rate of space infrastructure for satellite communication systems was relatively low with around ten satellites and launch vehicles produced per year. The commercial launch services market operated on business-to-government and business-to-business models, where institutional and government acted as contractors to a handful of large manufacturers in a narrow oligopolistic position. New entrants intend to drastically reduce spacecraft manufacturing costs and launch costs to develop commercial space [Nardon, 2017].

In New Space, market objectives take precedence over technical considerations [Davidian, 2020]. The innovation strategy adopted in New Space encompasses three interrelated goals: increasing the profitability of space activities, introducing new business models, and providing new space-based market solutions [ESPI, 2019, Dedieu et al., 2016]. To achieve the first objective, the industry concentrates its innovation efforts to substantially diminish the development and production costs of launch infrastructures and satellite [Denis et al., 2020, Robinson and Mazzucato, 2019]. This involves miniaturizing satellite platforms and payloads, adopting mass-production models with standardized components, and increasing production rates [Vidmar, 2020, Peeters, 2021]. In addition, test and decision times are significantly reduced, generating a higher risk of failure, now more equally shared between institutional and private investors.

The disruptive innovations[23] brought to the upstream sector are part of a broader shift in space companies' business models toward vertical integration [Vidmar, 2020]. Large players that recently entered the sector are targeting downstream markets such as global connectivity or surveillance. They identify a demand for service and and move up the value chain to meet downstream demand. New entrants tend to cover all activities, from satellite manufacturing to data processing to the supply of space-based products and services [ESPI, 2019, PwC France, 2019]. The example of mega constellations in the communication industry is quite striking. The Starlink satellite constellation project led by SpaceX seeks to launch around 12,000 small satellites into Low-Earth orbit (LEO)[24] to provide global Internet access. The company is involved in all project phases, from satellite development to launch on the Falcon 9 rocket and in-orbit operations. This case of extreme integration is not common to all New Space players. In the same segment, OneWeb relies on a less integrated business model. The company collaborates with incumbent firms, such as Airbus Defense and Space Intelligence for small satellites manufacturing and Arianespace for launch operations and constellation deployment. Two distinct economic models are emerging in New Space. On the one hand, a few large companies rapidly acquire technical know-how and adopt vertically integrated industrial models. On the other hand, new space ventures rely massively on the traditional manufacturing industry. The latter model often leads to partnerships between companies or mergers and acquisitions (e.g., Thales Alenia Space and Telespazio groups partnership with SpaceFlight Industries on BlackSky constellation).

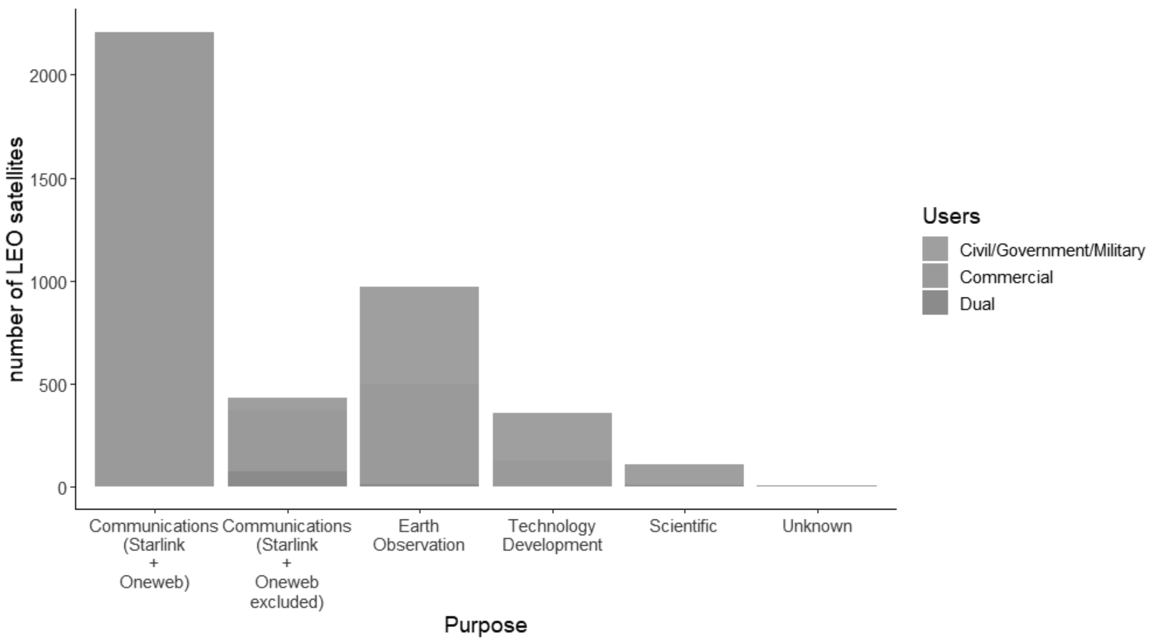
The development of commercial space is mainly concentrated in LEO. Indeed, 90% of operational satellites with purely commercial uses are low-Earth orbit satellites[25]. The commercialization of LEO is driven by lower launch costs, technical advances in

[23]Disruptive innovation refers to a type of innovation that offers a new value proposition to the market [Christensen, 2013]. It first appears to be less performant than existing technologies due to its newness and lack of refinement, and therefore only addresses a limited fringe of new consumers in the short run.

[24]LEO is an orbit at an altitude of less than 2,000 km above the Earth's surface.

[25]Purely commercial satellites account for more than 70% of the satellites currently in orbit. Dual-use satellites, i.e. those used for both commercial and military or civil purposes, are not included. Source: UCS Satellite Database

Figure 1.1: Operating Satellites in Low-Eath Orbit in 2022, by Purpose and Type of Use



Source: UCS Satellite Database

miniature satellites, lower regulatory barriers, and the emergence of new application markets enabled by the digital industry [Giannopapa et al., 2022, Lerner et al., 2016]. Figure 1.1 shows that the satellites deployed are mainly communication and Earth observation satellites. More than a half of LEO satellites belong to the Starlink and Oneweb constellations. The growing need for high speed connectivity (broadband Internet services, Internet of Things) has given a new lease of life to the satellite communications segment. Although the use of satellite broadband still pales in comparison to terrestrial technologies (fibre, DSL), LEO constellations offer several advantages for the telecommunications sector [OECD, 2019]. Latency, i.e. transmission time-lag, is a major issue for services such as video conferencing, telemedicine, stock trading and financial transactions. Due to the short distance from Earth, latency effects are substantially reduced compared to Geostationary orbit (GEO)[26]. The second advantage of large LEO constellations is their global coverage [Orlova et al., 2020]. This aspect gives satellite Internet a real competitive advantage as it can be used to complement terrestrial technologies to reach remote geographical areas.

[26]The geostationary orbit (GEO) is 36,000 km above the Earth's surface.

With digitalization, Earth observation satellite operators have seen new market opportunities emerge for satellite data. The increasing use of big data technologies to provide commercial services involves important data flows. Satellite images have become one of the sources of big data and are exploited in various sectors such as precision agriculture [Huang et al., 2018]. In this regard, LEO Earth observation constellations allow collecting images in large volume and at a much higher frequency (daily or even hourly) than traditional EO satellites [Lerner et al., 2016, OECD, 2019]. More than 900 Earth observation satellites are deployed in LEO, 55% are commercial or dual use satellites (Figure 1.1). The constellation operated by the US company Planet Labs accounts for more than one-third of the commercial EO satellites in LEO.

A last feature that logically follows from commercial objectives pursued by new space actors is the concentration of economic value at the end of the space value chain. The economic efficiency of the model mentioned above relies on the idea that increasing space systems production and launch rates necessarily lead to the disruption of existing markets and the creation of new mass markets [ESPI, 2019]. The development of space applications is not only driven by technological progress but also by potential users [Dedieu et al., 2016]. This market-driven logic has at least two consequences for downstream activities. First, the increasing heterogeneity of downstream space value chains and the diversification of space applications and services. New downstream markets related to the digital sector emerge such as global connectivity, geo-information systems, Internet of Things, and Machine to Machine networks. The second consequence is the extension of the space value chain to the end-user with the development of business-to-consumer services [Robinson and Mazzucato, 2019, Nardon, 2017]. While space infrastructure becomes standardized, space-based services are increasingly customized and therefore diversified.

In this section, we have depicted the primary changes occurred with New Space. In the next section, we will examine more specifically the nature of this change and analyze its implications in terms of the dynamics of the space sector.

1.4 Toward a demand-pull paradigm in the space sector ?

1.4.1 The hypothesis of a paradigm shift

References dealing with the space economy affirm that the space industry has been a demand-driven industry from its inception [Barbaroux, 2016, Barbaroux and dos Santos Paulino, 2013]. Space technologies and applications emerged in the post-1945 period during which governments made massive investments in military R&D [Mowery, 2010]. At that time, it is true that space assets were developed in mission-oriented R&D programmes and were primarily motivated by public policies challenges and political prestige [Robinson and Mazzucato, 2019]. States and defense organizations were prime investors and purchasers, shaping space technologies with well-defined technical requirements and supervising the industry [Heracleous et al., 2019].

However, we believe that a shift has occurred between the early stages of development of space activities and the New Space. This evolution describes the transition toward a demand-pull model in the space sector. The demand-pull theory analyses the causal relationship between economic growth and innovation activities. Space technologies first evolved independently of market constraints, with a single buyer, the State, controlling the entire value chain of a few sellers. Economic considerations, such as productivity gains due to the introduction of the new technology, entered into account only when space infrastructures opened to civil use in the form of indirect benefits [Mowery, 2010].

We presume that the transition towards a demand-pull model historically corresponds to the end of the Cold War and has been confirmed by the first achievements of the large private players. In the first stage of space liberalization in the late 1980s, public authorities intended to convert space resources formerly used for military purposes into economic assets. This resulted in the rapid growth of the satellite telecommunications sector. However, it is the convergence with the Internet economy that has challenged the prevailing industry models. Space industry no longer evolves independently of sectors of activity; digital entrepreneurs perceive and exploit space infrastructures as an economic asset, an instrument to develop their activity such as information systems. In this demand-pull vision, value creation in the space sector does no longer rely on the technology itself, but on the commercial goods and services produced from the technology.

In the market-pull approach, technological progress is not an exogenous variable but is shaped by economic factors and changes in market conditions [Schmookler, 2013, Rosenberg, 1974]. In this respect, market signals, i.e., changing demand characteristics, are determinant factors of innovation. The recent innovative activities within the space sector focus on three areas: launchers, satellite miniaturization, and electric propulsion of satellites. Faced with aggressive competition from new models of LEO constellations of small satellites, the traditional market for geostationary telecommunications satellites is in decline. To boost sales and withstand the downward pressure on prices for access to space, the Franco-Italian manufacturer Thales Alenia Space is launching in 2023 the "Space Inspire" line of digital, in-flight reconfigurable geostationary satellites[27]. This technological innovation enables commercial operators to change the service provided by the satellite directly in orbit and thus adapt to shifts in demand. For instance, a satellite dedicated to direct-to-home television (a shrinking market) could be reconfigured into a broadband connectivity satellite (a fast-growing market). The case of reconfigurable satellites illustrates the innovation model that is currently prevailing in the space sector. Technical progress is the result of adaptation to growing downstream markets.

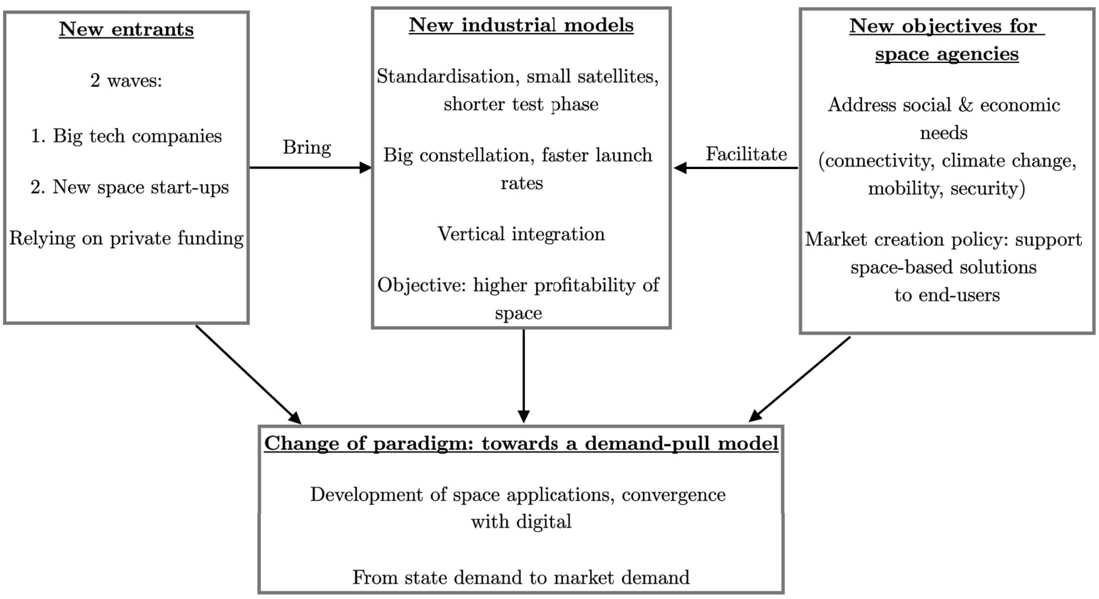


Figure 1.2: New Space characteristics - synthesis

[27] Thales Alenia Space is also active in the LEO constellations market, notably with BlackSky, Iridium NEXT, Kinéis, and Omnispace projects.

Figure 1.2 proposes an analysis of New Space based on the assumption of a paradigm shift in the space sector. It lays the theoretical foundation for a future conceptual framework to rely on for evaluating commercial space. We assume that New Space is marked by three trends: the entry of new companies and private investors that develop or finance space-based goods and services. These new entrants aim to facilitate access to space and apply industrial models not formerly applied in the space sector. Finally, states support the success of these players by orienting space policies towards market-oriented objectives. These characteristics reflect a structural change in which space activities are pursued according to a demand-pull logic. In this new dynamic, the growth of the space sector is no longer based solely on state demand but on market demand for space-based services.

The French space company Unseenlabs mentioned in Section 3.2 typically belongs to the second wave of new entrants. The start-up intends to create a constellation of about twenty *cubesats*, small and light satellites that are therefore cheaper to produce and launch. The company is adopting a vertical strategy, i.e., it is involved in the entire satellite value chain from the development and production of satellites to their in-orbit operation and the provision of final services based on satellite data to users. Public authorities support the firm's development through regional and governmental investment funds. Unseenlabs is a good example of market-pull logic. It is no longer the former pattern of a company developing a technology according to the instructions of a public order. The company identifies a market opportunity and develops a turnkey service for boat monitoring for public and private users.

1.4.2 Implications on the space value chain: the data dynamic approach

The demand-pull paradigm hypothesized in the previous section has implications for the dynamic of value creation in the space sector. First, it affects the industrial dynamics and value chain of the sector. The commercialization strategy of space activities has put emphasis on the downstream space applications segment. New Space companies are investing in the sector either with the aim of developing new markets or to enrich their existing service offer. We have described in this chapter the process of simplification of upstream space activities: standardization of space components, reusable launchers, miniaturization of satellites, easier access to space.

Conversely, the downstream space segment is becoming more complex. Its development results both in a multiplication of application domains (agriculture, insurance, transport, maritime, etc.) and an extension of the downstream value chain. Downstream companies are space companies, often involved in several segments of the value chain, but also non-space companies that exploit the infrastructure to enrich their product and service offer. Space applications and services most often combine space data or signals with other types of data (e.g. terrestrial and user data) and techniques. Besides, space-based services can be used by companies further down the chain to improve their products or services. This raises the challenge of identifying the downstream actors in the space sector, but also of measuring the revenues that come strictly from the use of space infrastructures or data.

Second, the incursion of digital technologies into the space sector has changed the dynamics of satellite data production and use. The rise of mega constellations in LEO reflects a growing demand for satellite data. Commercial satellite launches are ramping up, pointing out an evolution in the demand for large volume, fast frequency, easier and real-time access to data. As data flows increase, downstream players adopt digital technologies such as AI and cloud computing to facilitate mass data collection and processing.

Another evolution regarding space data in New Space is the trend, admittedly marginal and limited but observable, towards open data. In navigation, satellite positioning systems developed by States are based on the principle of open access to signals and free exploitation. In Earth observation, several platforms of institutional and, to a lesser extent, private initiatives have been developed in Europe. They are part of the wider EU open data policies, such as the Directive on "Open Data and the Re-use of Public Sector Information" adopted in 2019 to facilitate the access and use of public funded data [Council of European Union, 2019, Harris and Baumann, 2015, OECD, 2019]. The European Union provide satellite images and information from its Copernicus Earth observation programme. All Copernicus data are available and free of charge on a set of cloud-based platforms, such as the Copernicus Open Access Hub or more recently the Copernicus Data and Information Access Services (DIAS). They give access to Sentinel satellite images, but also to more sophisticated services and indicators for land, sea, and climate monitoring. European open data policies have several objectives. First, they provide large amounts of data for scientific research. Second, they allow states and institutions to develop adapted services for societal challenges such as climate monitoring. Finally, they aim to encourage small companies and start-ups to develop commercial services based on these data.

This suggests that the value does not lie in the production of raw space data, but in what it brings when exploited and combined with other technologies, data or information in the broadest sense. This has clear implications for the industrial dynamics of the sector, particularly for Earth observation operators for whom the sale of images is the core business. They must extend their activities down the chain and integrate their data into increasingly customized services. The vertical integration movement of space companies described in section 2.2 reflects this evolution. Earth observation start-ups tend to produce their own satellites, operate them and provide value-added services based on the data collected.

These elements are actually linked to the nature of data as an economic good. Data are non-rival goods: they can be consumed at the same time and as many times as desired without being diminished [Jones and Tonetti, 2020, Romer, 1990]. In the case of the space industry, once the Earth observation satellite is in orbit and the first set of images is created, it can be used simultaneously by several downstream companies and duplicated with zero or almost zero additional cost. These fundamental attributes of data, which make them akin to public goods if considered imperfectly appropriable, differentiate them from conventional physical goods. They imply that data are a source of considerable economic value, but make the assessment of this value complex. In this respect, the data value chain cannot be linear as the traditional value chain of physical goods. For instance, a given Earth observation image may be perceived as having little utility at a certain time but may have a considerable value afterwards, for example, when studying a natural disaster. Similarly, space data can gain or lose value depending on its combination with other data sources, such as end-user data. We propose here to link the evolution of the nature of satellite data and the emergence of new industrial models in the space sector. We also consider the implications of this change on the downstream space sector to identify its actors and assess its economic importance.

1.5 Conclusion

The purpose of this first chapter was to set the contextual background for our thesis by describing the principal trends in the space sector. In light of the literature dealing with space activities, which includes both academic articles and reports from space agencies and institutions, we argue that New Space corresponds to a structural mutation of the space sector. The characteristics of this transformation allow us to introduce both methodological and conceptual issues that are at the heart of our research question.

First, we show that governments and space agencies support the growth of private investment in space and have a particular interest in the development of the downstream services sector. Upstream, they formulate a demand for turnkey services and leave the design of launchers and their operation to the industry. Downstream, we have seen through the example of the European Union that the use of space data and signals for civil applications is perceived as a significant economic stake by states. Financial public support is increasingly targeted at downstream companies and access to space data from European programmes is facilitated, free, and open. Second, we describe a new industrial dynamic with the entry of digital players in the space sector. They reduce the costs of access to space, enabling start-ups and small space companies to emerge, and opening up new opportunities in terms of civil space applications.

In light of these elements, we conclude the chapter by hypothesizing a paradigm shift in the space sector to a demand-pull model. Specifically, we suggest that technological advances under the New Space and the new industrial dynamic described result from downstream market signals. The digital revolution offers new perspectives for the use of space data. The development of large constellations in low Earth orbit and the trend towards vertical integration of firms indicate that the commercial development of space is necessarily linked to the expansion of civil applications. This assumption gives a crucial dimension to the issue of measuring downstream space activities. The convergence with the digital sector allows us to question the role of satellite data in this new dynamic of the space sector. We propose to link the question of measuring the size of downstream sectors to the analysis of the evolution of data attributes in the context of New Space.

The assumptions made in this chapter constitute the premises of a conceptual framework on which our evaluation methodology relies. They can be summarized as follows:

(i) The space sector is moving toward a demand-pull paradigm in which technology development is increasingly driven by the demand for downstream space applications and services.

(ii) New Space introduces a new industrial dynamic that disrupts the traditional space value chain, especially downstream segments.

(iii) New Space and the convergence with the digital industry give particular importance to space data. The conceptual framework must address the attributes of data as an economic good and their implication when assessing the size of downstream space.

The following chapter focuses on the methodological aspect of our research question. We raise the issues associated with the measurement of the space sector, with a particular emphasis on the downstream segment. We review existing studies and highlight the limitations that motivate the construction of a new methodology for evaluating the economic importance of downstream space activities.

Chapter 2

Estimating the Downstream Space Sector: Purpose and Review

2.1 Introduction: Motivation for evaluating of the size of the downstream space sector

The space sector is commonly defined and represented using the value chain framework. This approach describes the linear sequence of value-adding activities, ranging from space R&D and the manufacture of launchers and satellites to providing space-based products and services to end users. In most cases, the space sector value chain is composed of two sub-sectors:

1. The *upstream sector* relates to the delivery of space technologies, i.e., satellite, launcher, and ground system design and manufacturing [Booz & Company, 2014, Oxford Economics, 2009]. It includes sub-systems, equipment, components, and related software supply [ESPI, 2019]. Launch operation services are most often defined as part of the upstream sector [OECD, 2022b].

2. The *downstream sector* covers all activities related to the commercial exploitation of space facilities and data to deliver value-added products and services to end-users [Moranta, 2022, OECD, 2022b]. Space services are commonly categorized into three application areas: communications (e.g., Direct-to-home broadcasting, satellite radio, broadband Internet), Earth observation or remote sensing (e.g., mapping services, ocean levels monitoring, responses to natural disasters), and navigation (e.g., location-based services) [Booz & Company, 2014, CSA, 2020].

In the previous chapter, we reported on a structural change in the space sector with New Space. The digital transformation has enhanced the commercial potential of space activities, offering new market opportunities from space technology exploitation. This implies a new industrial dynamic that affects the entire value chain with new players investing both upstream and downstream and new business models in which space companies' profit comes from the expanded use of space products and services by end users. The consulting firm BryceTech estimates that global commercial revenues generated by satellite services and the ground segment (i.e., the downstream sector as defined above) reached $260 billion in 2021, compared to $78 billion in 2005 [BryceTech, 2022, OECD, 2007]. On the other hand, upstream revenues (launch industry and satellite manufacturing) are estimated at $19 billion in 2021, up from $11 billion in 2005/ The consulting firm Euroconsult, whose reports serve as a primary source of information for many space institutions and agencies, indicates that global revenues from upstream space activities (manufacturing industry and launch services) amounted $33 billion in 2021, and downstream satellite services revenues reached $285 billion [Euroconsult, 2022]. Unlike BryceTech, these estimates include government spending in space. More generally, the figures given for the space sector in this chapter should be taken with caution. We will see that studies do not necessarily cover the same scope of activity. Based on these estimates, the downstream space sector has more than tripled between 2005 and 2021 (with 230% growth), while revenues from the upstream manufacturing sector have increased by 70% over the same period. These figures first show that the downstream sector accounts for the largest share of commercial space activities. Second, they illustrate the hypothesis we made in the first chapter that the commercialization of space activities during New Space has favored the development of the downstream sector, whose growth is much higher than that of upstream.

The commercial development of space and its main characteristics lead us to the question of how to evaluate the space sector. The ability to assess the economic value generated by the exploitation of space infrastructure is critical for both institutions and the private sector investing in space. For the former, it would facilitate the definition and implementation of effective and well-targeted space policies. For the latter, it would help identify dynamic space markets and motivate investment decisions. Naturally, the question of measuring economic returns to space activities has been widely addressed long before New Space and growing commercial space activities. Among the methods commonly applied to capture the economic benefits of space expenditures are impact analyses. They evaluate the quantitative effects of space programs or policies on an outcome of interest such as GDP. Evans [1976], Midwest Research Institute (MRI) [1988] use macroeconomic tools (macroeconomic forecast, production function approach) and econometric modelling to estimate the average return on each dollar invested in NASA R&D.

Cost-benefit analysis (CBA) is another approach commonly applied to assess the socio-economic impacts of space activities. For instance, Mathematica Inc. [1975] conducts a cost-benefit analysis to measure the "economic benefits to the U.S. economy from secondary applications of NASA technology". The study focuses on the knowledge generated by four NASA technologies used for purposes other than their original objectives. More recently, Murthi et al. [2007] evaluates the effects of public investment in space infrastructure in India, focusing on telecommunications and Earth observation. A study commissioned by the Italian space agency aims to quantify the benefits of national and European space policies on the Italian space sector with CBA [Florio et al., 2022]. Cost-benefit analysis is also applied for specific space programmes. Eumetsat [2014] quantifies the socio-economic impact of its *Metop* weather forecasting satellites. Based on the results, the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) decides whether to invest in the second generation of polar-orbiting meteorological satellites. CBA is a decision support tool to determine whether a resource allocation is efficient [Prest and Turvey, 1966]. It accounts for the costs and benefits of a policy, programme, or planned investment over time to quantify a net present value. The assessment is most often conducted *ex-ante* but can be used to analyze the outcome of a project. The first step is to identify all the costs and benefits associated with the project and assign them a monetary value. A *discount rate* is applied based on expected inflation to give a present value to the future costs and benefits. If the ratio between the sum of benefits and costs is positive and greater than 1, the project is expected to have a positive socio-economic impacts. A sensitivity analysis often complements the CBA to account for possible changes in the assumptions of the initial model (e.g., additional costs).

From a microeconomic perspective, the *Bureau d'Economie Théorique et Appliquée* (BETA) methodology evaluates specific economic effects called *spinoff effects* of European space projects [Bach et al., 1992]. Spinoffs are indirect long-term effects of space technologies transfer to non-space activities. The BETA methodology essentially relies on interviews with companies that received contracts from ESA and aims to capture the variety of learning processes acquired through their participation in a space programme. The BETA methodology provides minimum qualitative and quantitative estimates of these impacts and classifies them into several categories: technological, commercial, organizational, and work-factor-related effects. Indirect effects are expressed in terms of value added for companies under contract with ESA. The respondents estimate two types of coefficients [Cohendet, 1997]. *Q1* coefficients evaluate the distribution of the influence of each factor (technological Q1T, commercial Q1C, organizational and method Q1OM) on the increase in sales of the respondent firm. *Q2* coefficients estimate the share played by the work for ESA in each of the three factors (Q2T, Q1C, and Q2OM). The final effect is calculated by multiplying Q1 and Q2 coefficients by the increase in sales of the interviewed firm in the period considered. The authors also propose to estimate the value of *critical mass*, i.e., the minimum cost incurred by the company to qualify for space contracts if it had not been awarded an ESA contract. The BETA evaluation method is limited to the beneficiaries of agency contracts. It does not intend to capture the overall socio-economic benefits of public investment in space. PwC France [2019] uses an adapted version of the BETA methodology to estimate the revenues of intermediate users (downstream space companies) and end users enabled by the EU Copernicus programme's EO data. This approach was initially built to evaluate the indirect economic effects of European space projects, but can be applied to any innovation policy.

More generally, impact models aim to quantify space technology-induced economic gains. They evaluate downstream space activities as economic spin-offs of space technology transfer. The question addressed by impact analyses is: how access to space infrastructure and data enables intermediate users to achieve productivity gains, improve the quality of their services, provide new services, and increase their sales? Our thesis explores a different level of analysis. It investigates the methodological question of measuring the economic importance of downstream space activities. In other words, we are interested in analyzing downstream space as a sector and seek to know how it is defined, how the companies that are part of it are detected, and how the economic size of the sector is evaluated. The need for statistics on the downstream sector is intimately linked to our paradigm shift hypothesis. If technological development in New Space is driven by increased demand for space data and signals, then space applications can no longer be considered only through socio-economic impact analysis. The remainder of this chapter is therefore dedicated to reviewing studies dealing with measuring the size of the space economy. We pay particular attention to downstream space evaluation. We present the selected studies, their methodological approach, and their main results. We then comprehensively discuss the limitations of the studies in measuring downstream space, highlighting two aspects: the definition of the downstream sector and the method used to measure its economic importance.

2.2 Literature review: presentation of the studies and downstream space key figures

Our review draws on papers, surveys, and reports spearheaded by space agencies, institutions, or trade associations. We have deliberately restricted the analysis to publicly available studies that specify the evaluation perimeter and the methodology used. For this reason, countries involved in space activities but whose market reports are not public or do not present the measurement method are not included in the review. Table 2.1 summarizes the references reviewed according to the geographic area and the scope of the evaluation they cover. The extensive analysis with information on the method and main results is in Appendix 2.5 of the chapter.

The first set of references includes studies estimating the size of specific downstream application segments. EARSC [2019, 2017, 2013] and Lafaye [2017] focus on commercial activities based on Earth observation. European GNSS Agency [2017, 2015][28] deal with satellite navigation downstream markets. EUSPA [2022] covers both EO and Global Navigation Satellite System (GNSS) application markets. When defining the scope of activities considered, the reports essentially adopt the value chain approach. EUSPA [2022] identifies three types of activities in the GNSS market: components and receivers manufacturers (chipsets, antennas), system integrators delivering GNSS devices, and value-added service providers. As observed in EUSPA [2022] and EARSC [2019, 2017, 2013], the EO downstream value chain covers a wider range of activities. It includes mission and user ground stations (satellite operation, lease of satellite capacities, collecting, storing, and selling data) and value-added services and products (e.g., Geographic Information Systems companies) supply. The

[28]The European GNSS Agency became the European Union Agency for the Space Programme (EUSPA) in 2021. The agency is attached to the EU and aims to manage the Copernicus, Galileo, and EGNOS programmes.

results of the studies are based on data collected from public and private reports and company surveys. Both quantitative and qualitative data are collected (revenues, employment, application domain, number of receivers used for GNSS services).

Table 2.1: Publications by scope of evaluation and geographical area

Scope	Reference	Geographical area
Downstream applications	Lafaye [2017]	France
	EARSC [2019]	Europe
	EARSC [2017, 2013]	Canada & Europe
	European GNSS Agency [2017, 2015] EUSPA [2022]	World
Space economy	CSA [2020]	Canada
	Liu et al. [2019]	China
	ESPI [2020, 2019] Eurospace [2020]	Europe
	BryceTech [2021] OECD [2022b, 2012]	World
Impact of space activities	Booz & Company [2014] PwC France [2019] Technopolis group [2012]	Europe
	UK Space Agency [2022, 2021] UK Space Agency [2019, 2016] Oxford Economics [2009]	United Kingdom
	Highfill et al. [2020] Highfill and MacDonald [2022] Whealan George [2019]	United States

The second set of publications deals with the evaluation of the overall commercial space sector (”space economy” in table 2.1). BryceTech [2021] is part of a series of reports published annually on behalf of the U.S.-based Satellite Industry Association. It provides figures on the global satellite industry (operational satellites by mission type, satellite industry direct revenues). CSA [2020] is the latest report published by the CSA assessing the size of the Canadian space sector. The report presents the results of the annual survey of space companies. It gives detailed figures such as registered patents, commercial revenues, and employment by value-chain segment and region. In addition, CSA proposes an evaluation of the economic impact of space activities in Canada. Liu et al. [2019] performs an initial identification of Chinese commercial space companies based on their business segment (satellite manufacturing, operation, application services). Finally, the European Space Policy Institute (ESPI) conducts annual surveys of space companies and organizations [ESPI, 2020, 2019]. We also included in our review the *OECD Handbooks on Measuring the Space Economy* [OECD, 2022b, 2012]. This publication is purely methodological and intends to provide guidance to institutions in defining and assessing the space sector.

The last set of references evaluates the socio-economic impacts of space activities (”impact of space activities” in Table 2.1). Most publications use the input-output approach to assess the wider economic benefits of space activities. Booz & Company [2014], and PwC France [2019][29] respectively assess the socio-economic impacts from space in the EU and the benefits of the Copernicus programme for the European Commission. UK Space Agency [2022, 2021, 2019, 2016], Oxford Economics [2009] estimate the size of the UK space industry. CSA [2020] also assesses the impact of space activities on the Canadian GDP using the input-output method. This methodology measures the final GDP impact resulting from the injection of a certain amount of spending (e.g., public funding, R&D investment) into the economy. Three types of economic effects are considered: direct effects (spending associated with space system’s manufacturing), indirect effects (suppliers’ expenditure for materials), and induced effects (salaries of space sector’s employees spent on consumer goods and services). It provides metrics such as gross value added, employment, and labor income. In the United States, the Bureau of Economic Analysis (BEA) proposes an enhanced version of the input-output method with the construction of a *Space Economy Satellite Account* that identifies industries involved in the production of space commodities [Highfill and MacDonald, 2022, Highfill et al., 2020, OECD, 2022b]. We will analyze this approach in more detail in the section on methodological limitations.

Differences appear in maturity degree and market structure between space application domains when comparing estimation results. EO downstream services, although in expansion, is the less mature application segment. EUSPA [2022] indicates that the global EO downstream revenues amounted to 2.8 billion euros in 2021 against 199 billion euros for the navigation market and 133 billion euros for the

[29] In addition to evaluating the contribution of Copernicus to the EU GDP, PwC France [2019] estimates the ”revenues of intermediate and final users enabled by Copernicus data and products” with the BETA methodology.

telecommunication services industry (telecommunications services and user equipment) [BryceTech, 2022]. A relatively low number of companies hold the satellite imagery market. It is fragmented with, on the one hand, few large space players (the U.S. company Maxar held 30% of the global EO markets in 2017, Airbus Defence and Space 12%) and non-space players (value-added resellers and big data analytics companies such as Atos represented 36% of the market in 2018), and on the other hand many small players [PwC France, 2019]. More than 65% of EO downstream players are "micro-companies," i.e., companies with less than ten employees [PwC France, 2019, EARSC, 2017]. Regarding the demand for EO data and services, over 60% of the EO downstream revenues emanate from the public sector (national and European public bodies, international institutions, public research institutes) [EARSC, 2017].

The navigation market is the downstream segment that has enjoyed the most significant growth since the early 2000s, reaching almost 200 billion euros of global revenues in 2021 compared to 120 billion euros in 2008 [EUSPA, 2022, European GNSS Space Agency, 2010]. Like EO services, the GNSS downstream segment is also composed of a few large companies and a significant number of SMEs. The sector is characterized by a growing trend of mergers and acquisitions, especially in the navigation device industry, where five companies accounted for 60% of income in 2015 [European GNSS Agency, 2017]. Both studies on GNSS downstream activities suggest that road and location-based services dominate navigation markets with respectively 50% and 43% of total revenues over the period 2015-2025.

The global satellite communications market is the more mature downstream segment [ESPI, 2018]. It comprises two sub-segments: broadcast services (Direct-to-Home satellite television, radio) and broadband services (Internet access). Satellite TV still dominates satellite telecommunications despite a decline in recent years. Satellite telecommunication operators' revenues amounted $13 billion in 2015 [Dedieu et al., 2016]. They concentrate among a few large historic players such as Intelsat, Eutelsat, SES, and Inmarsat [OECD, 2014]. The commercial telecommunications market has been experiencing a significant decline since 2015, with five orders of commercial satellites in 2017 compared to twenty-five in 2014 [OECD, 2019]. This trend is explained by the stagnation of the traditional telecommunications market and the rise of new market segments [Bondiou-Clergerie, 2019]. Broadband connectivity and the Internet of Things are a new growth opportunity for the satellite telecommunications segment [OECD, 2019]. Satellite broadband is still in the early development phase but is spurred by mega-constellation projects such as OneWeb, Starlink, and Amazon's Kuiper.

2.3 Limitations in estimating the size of downstream space markets

Our review raises limitations in how studies address the evaluation of the downstream space segment. These issues relate to 1) the representation of the downstream segment and 2) the methodology for measuring the direct revenues associated with these activities.

2.3.1 Different definitions of downstream space

When assessing the economic value of downstream space markets, a first methodological challenge relates to the proper definition of downstream space. From one publication to another, the representation of the downstream space segment – and a fortiori the scope of actors considered to assess it – differs, which makes comparability of the data difficult. These differences are observable at two stages of the space value chain: the upstream-downstream distinction and the downstream end of the sector. Figure 2.1 presents the space value chain and shows the differences we identified in the perimeter of each segment between the studies of our corpus.

2.3.1.1 Position of the ground segment in the space value chain

The ground segment covers activities relating to in-orbit satellite management ("mission ground segment" in figure 2.1), satellite operation services, and supply of user equipment ("user ground segment" in figure 2.1). Comparing the studies of our corpus, we find differences in how these activities are categorized in the value chain. First, some studies do not adopt the upstream-downstream distinction. Liu et al. [2019] defines three types of space companies: manufacturing, operational, and application. Similarly, BryceTech [2021] distinguishes four segments in the commercial satellite industry: satellite manufacturing, launch industry, ground equipment, and satellite services. Booz & Company [2014] and Oxford Economics [2009] represent the ground segment as intermediate or midstream activities in the value chain, making a distinction with downstream value-added products and services. This segmentation generates confusion since, in most cases, it is not used in the evaluation phase nor in the presentation of the results in the same reports. Finally, Highfill and MacDonald [2022], Highfill et al. [2020] do not provide a representation of the space sector per se. This can be explained by the fact that the objective of the study is not to assess the size of the space sector and identify space companies but to evaluate the impact of space technologies on the U.S. economy.

ESPI [2021] adopts another upstream-downstream distinction, using the mapping of space actors proposed by the investment company Seraphim Capital. The study defines upstream space in 3 sub-segments: build, launch and data. The data category corresponds to the mission ground segment of our value chain. It includes manufacturers and operators of EO satellites such as Planet and the Finnish start-up Iceye. Thus, mission operation services belong to the upstream segment in this example (arrow 1 in figure 2.1). The remaining publications define both mission and user ground segments as downstream space activities (arrow 2 in figure 2.1) [CSA, 2020, EUSPA, 2022, PwC France, 2019, UK Space Agency, 2022]. This representation is suggested in OECD [2022b, 2012] for a harmonized definition of the space economy. It is consistent with the idea that space activities progressively result from a demand for service. The upstream part of the sector concerns only space infrastructure development and manufacturing, whereas commercial activities once satellites are in orbit are part of the downstream segment.

We have noted a first difference in the segmentation of space markets with satellite operations activities and ground equipment. This aspect shows that the space value chain is not stabilized and leads to differences in the estimation of the downstream space segment depending on whether satellite operations are part of it. In addition to generating data comparability issues, the differences in the upstream-downstream distinction between studies highlight the difficulty of the value chain representation in capturing the vertical integration trend observed in chapter 1. In the case of LEO constellations, companies develop and build the satellites and provide operations and downstream services. The value chain approach helps define space activities, but it is questionable whether it is enough to identify the actors that comprise each activity segment.

2.3.2 What downstream end?

Another issue in reviewing the studies is identifying the value chain downstream end. To what level of exploitation of space data or signal is an organization considered

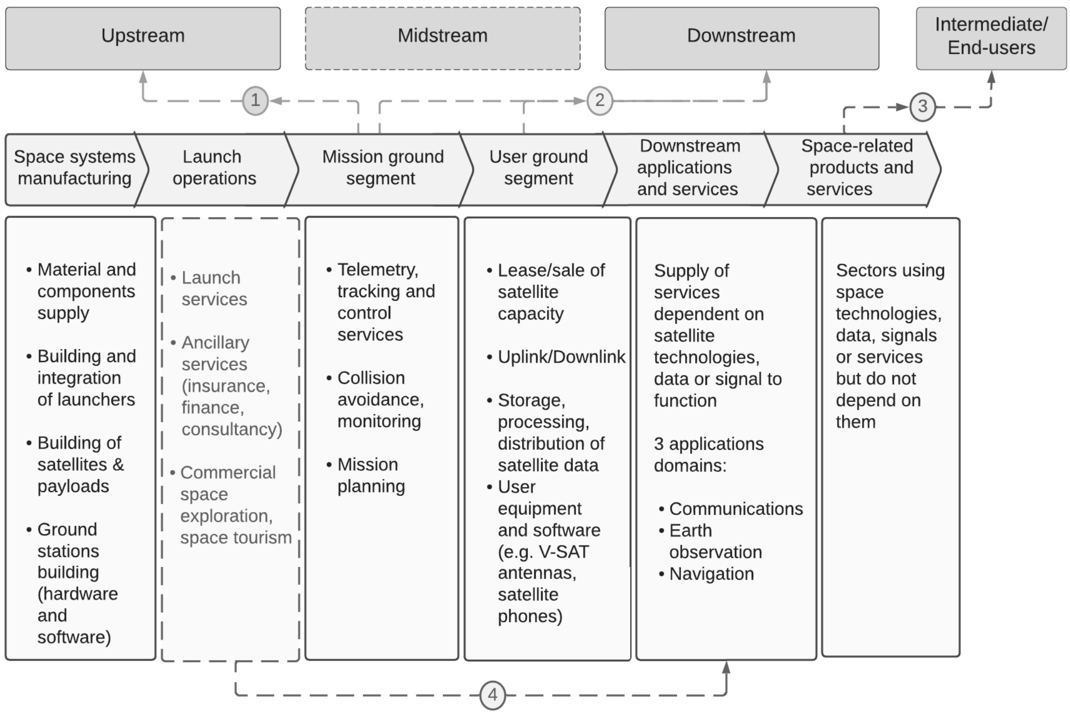


Figure 2.1: Space value chain representations

to belong to the downstream space segment? How can this level be measured? The evolution of the perimeter of downstream activities is linked to the question of commercial uses of space and their fast development in New Space. Space-based products and services are expanding both in volume and variety. A direct consequence of this trend is the extension of the space sector's value chain: creating added value from space data involves a longer process of data and signal transformation and the potential recombination of space assets with other data sources. Therefore, some studies of our corpus underline that identifying downstream players is challenging. It encompasses increasingly heterogeneous business segments and may even include companies and organizations whose activities' portfolio is not entirely integrated into the traditional space value chain [ESPI, 2019, OECD, 2019].

The OECD is working on harmonizing the terminology and the methodologies used by the space community to represent and measure the commercial space sector. OECD [2022b, 2019] makes a distinction between the downstream space sector and *space-derived activities*. The latter comprises "all economic activities hosted in non-space economic sectors, but derived from the application of space technologies." This implies that companies or organizations that do not belong to the space sector, i.e., whose core business and main source of revenue do not directly depend on the provision of satellite signal and data, should not be included in assessment studies of the downstream space segment. In other words, companies who use space-based products and services for their activity but for whom space assets are substitutes should not belong to downstream space but to a broader category called the space economy.

Most studies adopt the downstream-end user distinction in representing the space value chain (arrow 3 in figure 2.1) [BryceTech, 2021, CSA, 2020, EARSC, 2019, 2017, 2013, ESPI, 2020]. For instance, BryceTech adopts a rather conservative approach in defining a space company [BryceTech, 2021, Bryce Space & Technology, 2019b]. The satellite services sector includes companies providing services that rely on space systems (consumer services, e.g., satellite TV, radio, and broadband), fixed and mobile satellite services providers, and analytic services providers based on "data collected extensively from space-based systems." In this specific representation, the EO services segment refers to satellite operators that provide "optical and radar images to the open market" and value-added services either with only EO data or in combination with terrestrial systems. This definition of the EO downstream market excludes a priori players that are not directly involved in processing raw data and that use, for instance, geospatial information in which EO data have been combined with other sources. For instance, meteorological data are increasingly utilized with other information and knowledge to provide services whose functionality goes beyond simple weather forecasting (e.g., solar maps). Conversely, EARSC [2021] defines geographic information service companies whose "focus is in other sectors but where EO data is used" as part of the EO services industry, distinguishing them from end users. Moreover, it seems that BryceTech reports do not consider the existence of a GNSS-based services downstream market for the navigation domain. Only revenues from GNSS devices and chipsets, included in the ground equipment category, are estimated for this application market. Location-based services developed partly from GNSS signals are thus not included in the assessment. The Canadian report includes navigation services, indicating that it is the second largest source of revenue for navigation-related space activities after products and applications [CSA, 2020].

The reports published by the UK space agency admit a different understanding of downstream space activities from that advocated by the OECD. First, they categorize

launch and satellite insurance services ("ancillary services") in the downstream segment, unlike most other studies of our corpus [UK Space Agency, 2022, Oxford Economics, 2009]. Second, the 2022 report defines all operation services, including launch services, as downstream activities. In the same report, commercial space exploration and space tourism are also part of the downstream segment [UK Space Agency, 2022]. These differences are indicated by arrow 4 in figure 2.1.
UK reports adopt a definition of downstream space applications similar to the other studies. However, they provide an assessment of the share of "UK GDP that is supported by satellite services" (GNSS, EO, meteorological, and communications services) [UK Space Agency, 2021]. For instance, in 2016-2017, satellite navigation services supported nearly 15% of UK GDP. This result, while surprising in comparison to the other impact evaluations reviewed in this thesis, highlights the difficulty in defining the downstream limit of the space value chain. Identifying downstream activities raises the question of determining the level of service dependency on the infrastructure, data, or space signal for it to be considered a downstream application or a space-derived service.

2.3.3 Methodological limitations

The differences in value chains observed between the evaluations highlight the problem of identifying space activities in statistical classifications of economic activities. Space activities do not form a distinct economic sector in countries' official industry nomenclatures (e.g., North American Industry Classification in North America, Nomenclature d'Activités Françaises in France). Whether in space systems manufacturing or the supply of space-based value-added services, *space companies* do not exist in statistical terms. Upstream space companies are scattered in several industries such as aerospace, defense, and engineering. The problem is even more acute for the downstream segment, which is composed mainly of service companies. Satellite telecommunications is a separate category in the statistical classifications, but companies providing Earth observation services or satellite location-based services correspond to various activity codes.

Studies that estimate direct revenues generated by the space sector generally rely on survey data. Respondents are companies "whose activities include the development and use of space assets and/or space data" [CSA, 2020] or, more restrictively, whose "main business [...] (in revenue share) is part of the space value chain" [ESPI, 2021]. UK Space Agency [2022] collected 152 responses from "UK-based space organizations". However, the reports provide very little information about how space businesses and organizations were identified prior to the survey. We understand that, for most studies, the space companies surveyed are known to agencies and institutions. They are either part of the space community network, have responded to a call for funding, have contracted with space agencies for a mission, are part of a directory of space companies, or have participated in a space-related event. Based on the references of our literature review, we highlight that there is no systematic method for identifying space actors and a fortiori downstream space actors. The first arguments that motivate such a method are the lack of recognition of space activities as an economic sector in official statistics and the harmonization of space economy assessment between countries. Moreover, the current way of identifying space companies or organizations is inconsistent with the demand-pull hypothesis we made in this thesis's first chapter. If the space value chain is defined from the activities of companies already known by the space community, how can we capture the structural transformation of commercial space? We hypothesized that New Space companies more and more respond to downstream market demand and that the offer of space services is no longer only the result of a technological opportunity. In our opinion, the current approach to representing and identifying does not allow us to follow the evolution of the downstream space value chain (new entrants, new application domains, etc.).

To overcome the difficulty of defining the space sector using activities classifications, we have seen that some studies conduct evaluations on the economic impact of space activities [Booz & Company, 2014, CSA, 2020, Highfill and MacDonald, 2022, PwC France, 2019, UK Space Agency, 2022, Whealan George, 2019]. Most of them apply the input-output model or an approach derived from this model. The objective of such evaluation is to answer the question: what is the contribution of space activities to the national economy in terms of GDP and employment? Input-output analytical framework describes interdependences in an economy by analyzing the flows of intermediate and final goods and services among industries. The sectors are presented in a double-entry matrix that shows the output distribution of each industry among the other sectors of the economy and, similarly, from which sectors a given sector's inputs originate [Miernyk, 2020].

The BEA study for estimating the U.S. space economy follows three steps [Highfill and MacDonald, 2022, Highfill et al., 2020]. The first step is the definition of the space economy with the construction of a Space Economy Satellite Account[30]. The authors identified 200 commodity codes, i.e., goods and services included in the space sector, based on survey data and consultation with experts from the space community. The commodities are classified into three categories: low, medium or strong support to the space economy. The study's second step consists in distinguishing space commodities from non-space commodities within a product code when necessary. For this purpose, the authors used a classification of manufacturing products that gives a more refined level of product code. The third and final step in the method is using supply-use tables (i.e., input-output matrices from the U.S. System of National Accounts) combined with the space commodity information from steps 1 and 2 to estimate economic activity by sector. The macroeconomic aggregates obtained are gross output (value of intermediate and final goods and services), added value, jobs supported, and wages by sector and in total.

This tool allows comparison of the space sector across countries and with other industries within a country [OECD, 2022b]. BEA produces many national statistics based on input-output matrices. Additionally, countries' activity classification systems that use the International Standard Industrial Classification (ISIC) are compatible. OECD [2022b] indicates that the BEA study prompted several countries, including France, to build satellite accounts for official and harmonized statistics on the size of the space sector.
However, we note several limitations in the input-output approach to measuring the direct revenues of downstream space activities. First, national accounting does not fit to evaluate a particular activity segment. It gives access to general information at the industry level. It does not allow for the representation of the value chain of the space sector and the distinction between upstream and downstream revenues. Similarly, the product approach does not capture the structure of downstream markets. We do not have information on individual companies' contributions to downstream space. This information helps detect the industrial transformations such as vertical integration, entry of new private space players, and the expansion of downstream firms. For example, Highfill et al. [2020] inform us that in 2018, the "agriculture, forestry, fishing, hunting, mining, and utilities" industry contributed 10 million dollars to

[30]Satellite accounts were integrated into the U.S. NAICS in 2008 for activities that are not defined as economic sectors (e.g., ocean economy, digital economy, and now space economy) but whose size is critical enough to be estimated with the national accounting tool.

the U.S. space economy. Which companies are involved in the production of these commodities? Are they space companies or non-space companies? What kind of space-based data or signal do they use?

Second, when looking at the NAICS-based codes used in the U.S. space satellite account, we see that many are manufacturing products from the upstream space sector (spacecraft, components, instruments) or R&D services associated with upstream [OECD, 2022b]. The satellite communications segment is well delineated in classification systems and is, therefore, the only visible downstream segment. By contrast, satellite navigation includes navigation equipment, device, and systems manufacturing codes but does not cover location-based services. The task is even more perilous for Earth observation services. One of the codes that include applications based on remote sensing images is: "541370 - Surveying and mapping (except geophysical) services". In this case, we cannot access a more detailed code to distinguish between space and non-space activities since these are services and not manufacturing products. In our view, this approach effectively measures traditional mature space activities such as the aerospace industry or satellite communications. However, it necessarily underestimates downstream application segments. We believe that an assessment method exclusively based on standardized industry classifications provides a partial and conservative representation of space-based application services.

Finally, the analysis based on the input-output framework seems inadequate for assessing the economic importance of downstream space activities, as it does not consider the specific characteristics of this segment outlined in the first chapter. Part of the downstream activities consists of exploiting satellite data[31] combined with other inputs of varied nature to provide value-added space services. In chapter 1, we emphasized that data is different in nature from standard economic goods. For instance, a piece of metal that forms the structure of rockets does not have the same characteristics as a satellite image used for weather forecasting. On the one hand, the piece of metal can only appear in one cell of the input-output matrix as an input to the spacecraft assembly industry. On the other hand, the same satellite image can appear simultaneously and without diminishing as input of the meteorology and precision agriculture sectors. Ten years later, this same image could be reused without losing its utility as an input for other sectors of the economy. The non-rivalry attribute of satellite data does not allow us to estimate the production value that integrates data

[31] We use the term "data" in a broad sense, e.g., satellite images, navigation signals, etc.

(e.g., space-based value-added services) with the input-output analysis.

2.4 Conclusion

This chapter aimed to introduce the methodological challenge of the thesis. We investigated how commercial space activities, and more specifically downstream activities, are measured in the context of New Space. Our literature review drew on a selection of studies intending to evaluate the space sector's size or economic importance. The corpus of references we analyzed is composed of studies conducted by consulting firms on behalf of national space agencies and space institutions or performed by the agencies and institutions themselves. Although some of them include an evaluation of the socio-economic impacts of space activities, our review primarily focused on publications estimating the size or *economic footprint* of the space sector and providing minimum information on the methodological approach used. Two aspects guided our review: 1) how studies define the downstream sector in the space value chain, and 2) what methodological tools have been developed and applied to estimate the economic footprint of commercial space activities and the downstream sector in particular.

The first difficulty we observed when estimating the size of downstream space relates to the scope of activities the downstream space sector covers. We noted significant differences in segmentation between upstream and downstream space sectors. The intermediate sector, which links in-orbit activities and terrestrial space applications, is either integrated into the upstream sector, is a separate segment, or belongs to the downstream sector, depending on the study. In addition, we had difficulty identifying a single, delimited downstream boundary of the space value chain in the publications considered. Much of the corpus aligns with OECD [2022b, 2012]'s definition based on the notion of *reliance* on satellite signals and data [OECD, 2022b]. From this perspective, the downstream sector includes the activities aiming to provide products and services derived from satellite exploitation. The OECD distinguishes the downstream sector from space-derived activities, which result from space technology transfers to non-space companies. Despite the examples of activities provided in the organization's handbook (e.g., operations of space and ground systems, supply of space-based services, and products supporting consumer markets), we saw that this distinction remains unclear when identifying space service suppliers in practice. In the case of Earth observation, for example, the difference between downstream companies and end users of Earth observation is tenuous, leading to notable differences in activities included in the Earth observation segment between studies.

Secondly, we explored the methodological approaches developed to estimate the size of commercial space activities and, more particularly, the downstream space sector. The main difficulty we raised, which explains the sector's definition issues mentioned, is that commercial space activities do not form a distinct industry in official industrial classifications. Therefore, statistics on the space sector, such as revenues, employment, number of space companies per segment, and company characteristics, cannot be developed only from existing statistical tools (extraction of sector statistics from activity codes). It is especially challenging for downstream applications, which include mostly service activities spread over various sectors. The evaluation studies we considered in this chapter collect data through surveys of downstream space companies. However, they do not indicate how space companies are identified. Most of the surveys are conducted with companies from internal databases of space agencies and institutions. We highlight that there is no systematic, open, and reproducible method to identify the companies forming the downstream space sector.
One approach we examine in our review is using impact assessment tools to estimate the size of the space sector within an economy. The studies rely on input-output matrices to estimate the economic value (expressed in GDP contribution, gross output, and wages) produced by firms involved in space goods and services production. Once again, we observed that this method captures very partially the value generated in the downstream space sector. Indeed, the first step of this method consists in identifying space products in selected industries. This approach is effective in the case of physical goods (launchers, satellites, user equipment), but its ability to identify Earth observation and navigation services is more questionable.

The literature review allows us to draw several lessons for developing a new assessment methodology of downstream space activities. Since the space sector cannot be delineated via industry classification systems, one of the major methodological stakes is to define it by accounting for the changes to which it is subject. We have observed significant changes in business models with customers who no longer formulate a demand for technology but a demand for space services. Take the example of a provider of air quality monitoring services. Instead of buying or leasing a satellite, which it will have to operate by investing in ground stations and developing expertise, the company can share a satellite with other customers and use ground station services to retrieve the data it needs. This new business model implies significant changes in the value chain, such as vertical integration of space companies, and partly explains the challenges of representing the operations segment (launch services and ground segment). This new dynamic is comparable to cloud computing in the digital industry: customers subscribe to a shared computer software service that provides the same utility as owning the hardware. Satellites and ground stations for operations are becoming space services. Given this evolution, the upstream segment would include only purely manufacturing activities of launchers and satellites. The remaining space activities would belong to the downstream sector. The notion of downstream space companies also remains to be defined in the specific context of New Space. Companies close to satellite technology (e.g., telecommunications satellite operators) are easily identified as downstream companies. However, it is more difficult for companies that exploit satellite data they have not purchased but obtained via data platforms to detect whether they belong or not to downstream space. Is the factor of proximity to space technology relevant to defining the reliance of service companies on space data and infrastructure? Therefore, the method for measuring the downstream space sector must rely on a theoretical analysis of the space sector to capture the evolution of upstream and downstream space activities. The issue of defining the downstream sector is central to the design of the method. It requires to specify the nature of the dependence of a service company on space infrastructure that allows it to be identified as a downstream space company.

2.5 Appendix

Table 2.2: Summary Table of Studies Reviewed (1/2)

Reference	Evaluation perimeter	Data and method	Main results
Booz & Company [2014]	European space economy	Input-output analysis	Upstream revenues: €6.5B (2011) Downstream revenues: €43.7B (2012)
BryceTech [2022, 2021]	Global satellite industry	n/a	Global revenues: $386B Telecommunication services: $115 Remote sensing: $3 (2021)
CSA [2020]	Canadian space sector	Survey Input-output analysis	Direct revenues: $5.5B Value Added: $2.5B (2019)
EARSC [2021, 2019, 2017, 2013]	European EO services industry	Public information, survey Methodology n/a	515 EO companies Total revenues: €1.25B (2019)
ESPI [2020, 2019]	Private investment in the EU space sector	ESPI database, public information, survey Methodology n/a	Private investment: 56 deals, €188M (2019)
European GNSS Agency [2017, 2015]	Global GNSS markets	Private database, public information Forecasting model	GNSS downstream markets: €45B Value-added services: €50B (2015)
EUSPA [2022]	Global EO and GNSS markets	Forecasting model	GNSS downstream revenues: €199B to €492B EO downstream revenues: €2.8B to €5.5B (2021-2031)
Highfill and MacDonald [2022], Highfill et al. [2020]	US space economy	Data: BEA's Supply-Use Tables Input-output analysis	Direct revenues: $177.5B Value added: $108.9B

Table 2.3: Summary Table of Studies Reviewed (1/2)

Reference	Evaluation perimeter	Data and method	Main results
Lafaye [2017]	French downstream EO companies	Private database, public information Methodology n/a	92 companies (30 non-commercial organizations)
Liu et al. [2019]	Chinese commercial space companies	Company identification: online searches (website, media), interviews No assessment	78 companies (44 downstream)
UK Space Agency [2022, 2021, 2019, 2016]	UK space industry	Private and public information, survey, impact evaluation	Total direct revenues £16.5B; Space applications £12,2B (2019,20)
Oxford Economics [2009]	UK space sector	Input-output analysis Econometrics	Total revenues: £5.9B Downstream revenues £5.08B Value added: £2.8B (2007)
PwC France [2019]	Copernicus programme downstream markets	Input-output analysis	Public investment: €8.2B Economic benefits from downstream markets: €16.2 to €21.3B
Technopolis group [2012]	Benefits of public investment in space (methodological note)	Enhanced input-output model Economic, social and environmental impacts	n/a
Whealan George [2019]	Commercial space industry in Florida and the US	Input-output analysis	Value added: $38B Employment: 36.000 jobs from a 7% industry growth (US, 2016)

Part II.
Methodology for Assessing the Downstream Space Sector and First Application for France in 2021

Chapter 3

Identification of French Downstream Space Companies: A Text-Based Approach

This chapter was co-authored with

Pierre PELLETIER

3.1 Introduction

As implied by the term 'downstream space segment,' most studies that analyze and evaluate commercial space activities use a value chain representation [CSA, 2020, EARSC, 2021, UK Space Agency, 2022, PwC France, 2019]. This mode of representation requires understanding the various activities and stages that contribute value to a product or service until it reaches the end user. It is the essence of assessment studies that employ the input-output framework to estimate the space sector's contribution to the economy [CSA, 2020, Highfill and MacDonald, 2022]. However, we highlighted in Chapter 2 that the use of input-output matrices necessitates selecting sectors that incorporate space activities at the beginning of the analysis. This approach is effective for measuring manufacturing activities but less when applied to value-added services further down the value chain, where activity codes are highly diverse. Consequently, studies measuring the size of the space economy in terms of its contribution to industry sectors result in an incomplete downstream segment assessment.

Therefore, we propose adopting an approach to identify the players involved in downstream space activities. Understanding the companies participating in the sector would allow us to overcome the problem of explicitly identifying downstream activities within activity codes, a challenge that arises with sector-level analysis. Rather than attempting to estimate the contribution of space within various industry sectors, the idea focuses on detecting the sectors most represented among downstream companies and using this information to identify unknown downstream space businesses. Moreover, we believe knowledge of companies involved in downstream space activities is fundamental in the context of New Space. We defined New Space as a paradigm shift in the space sector, marked by the entry of actors intent on developing new downstream space markets. Given this evolving landscape, understanding and monitoring the sector's industrial dynamics is essential when developing an evaluation method for the downstream space segment. It entails detecting the economic actors in the segment and new entrants, analyzing their characteristics, measuring the economic value their activities generate, and following this evolution over time.

This chapter introduces an identification method based on a text-mining approach. In concrete terms, this method exploits text, a form of unstructured data, to extract new, structured information [Tan et al., 1999], using specific tools that we will detail later. Text mining is increasingly recognized in economic and financial literature as a rich source of information [Gentzkow et al., 2019]. It offers the possibility of formulating predictions in various fields, such as stock prices determination [Antweiler and Frank, 2004] or measuring uncertainty in macroeconomic policies [Tobback et al., 2018]. Recent studies have even exploited the text of corporate websites as an alternative for assessing their economic activity and innovation potential [Arora et al., 2013, Gök et al., 2015, Libaers et al., 2016]. Text data mining is not limited to these areas. It can also be used to analyze companies' activities, to propose an alternative industrial classification to the reference classifications [Al-Hassan et al., 2013, Hoberg and Phillips, 2016, Kile and Phillips, 2009], or even to determine a company's industrial code from its business plan [Tong and Fortino]. The textual sources exploited in this approach are varied and encompass scientific articles, speeches, web pages, publications on social networks, or even emails [Hassani et al., 2020].

In the context of our research, the information we seek to extract from the text is the names of companies registered in France and involved in downstream space activities. For this purpose, our primary textual source is the French-language newspaper press. There are several reasons for this choice. Firstly, newspapers are a rich textual source in terms of data volume. The digital version of the press offers abundant content, which is particularly useful in text mining and machine learning methods that require large amounts of training data. The second argument is that the press is an effective source of information and economic analysis. For example, it is used to measure the state of an economy [Bybee et al., 2021] or to gather information on fraud in finance literature [Miller, 2006]. Thirdly, the digital format of the press is a rapid information medium, so certain players such as investors consult it in near-real time [Hong and Han, 2002]. In this way, the use of the press fits in particularly well with our exploratory approach, the aim of which is to discover downstream companies that are new, in the sense of being unknown to industry experts, in an evolving sector.

We propose to apply a method for identifying named entities to press data called Named Entity Recognition (NER). This approach refers to '[the extraction of] specific "proper nouns" from unstructured texts, such as the names of people, places, and institutions, in addition to dates' [Liu et al., 2022]. In our case, we are dealing with organizations with downstream space activity. Named-entity recognition is a Natural Language Processing (NLP) task that enables machines to read, understand and interpret human language [Gorinski et al., 2019]. Thus, the NER follows a two-step process: it detects the entities cited in a text, then classifies, or *labelise* them according to categories (place, person, organization). Maurya et al. [2022] conducts research very similar to ours, both from a methodological point of view and in terms of the field analyzed. Indeed, the study is based on NER's tool for detecting satellite names, rocket names, and space agencies from Wikipedia and Google news texts. They use a machine learning model, i.e., a system pre-trained to annotate entities in a text correctly.

There are two possibilities for identifying entities in text in NER: a statistics-based approach and a rule-based approach. Unlike Maurya et al. [2022], which employs the statistical method, we use the rule-based approach. In other words, we establish a word dictionary and formulate a series of instructions (or rules) to apply to the press text to match the names of companies involved in downstream activities. Because of their diversity, company names are particularly challenging to recognize compared to personal names. Indeed, company names can be common nouns used as proper nouns, surnames, or first names. It makes it particularly difficult to capture this type of entity using a purely statistical model, and context is essential to limit the over-identification of words like company names in a text. The rule-based approach thus enables experts with knowledge of the companies to be detected to formulate customized detection rules.

In this chapter, we provide a detailed description of our NER rule-based approach for detecting downstream company names and present the results of the first application of the method to the case of the French downstream sector. The remainder of the chapter is organized as follows. In the next section, we provide a general presentation of the method, detailing the main steps and the textual data collected. As our methodological tool is intended to be replicated, we detail in Section 3 the query that enabled us to collect articles on the subject of downstream space activities and formulated rules. In Section 4, we present the main results of the first application of the method and detail the number of downstream space companies it enabled us to discover. Section 5 evaluates specific aspects of this first version of the method. Finally, with the methodological aim in mind, we address the question of the time required to implement the method and its replication. We conclude the chapter with a discussion, reviewing the main contributions of our approach and suggesting avenues for improvement.

3.2 General overview of the rule-based approach for downstream space company names identification

This section comprehensively describes our methodology for identifying French downstream space companies. The proposed approach is designed to supplement internal databases with companies previously unknown to space actors and to overcome the traditional statistical tools' limitations discussed in the introduction. We adopt a **rule-based method** based on textual information from press databases. The general principle is to extract from news about the French downstream space sector a relatively compact list of companies engaged in downstream space activities. We implement a series of hand-crafted rules to recognize company names associated with the sector within a raw text of information on the downstream space sector. Figure

3.1 summarizes the critical stages of our approach.

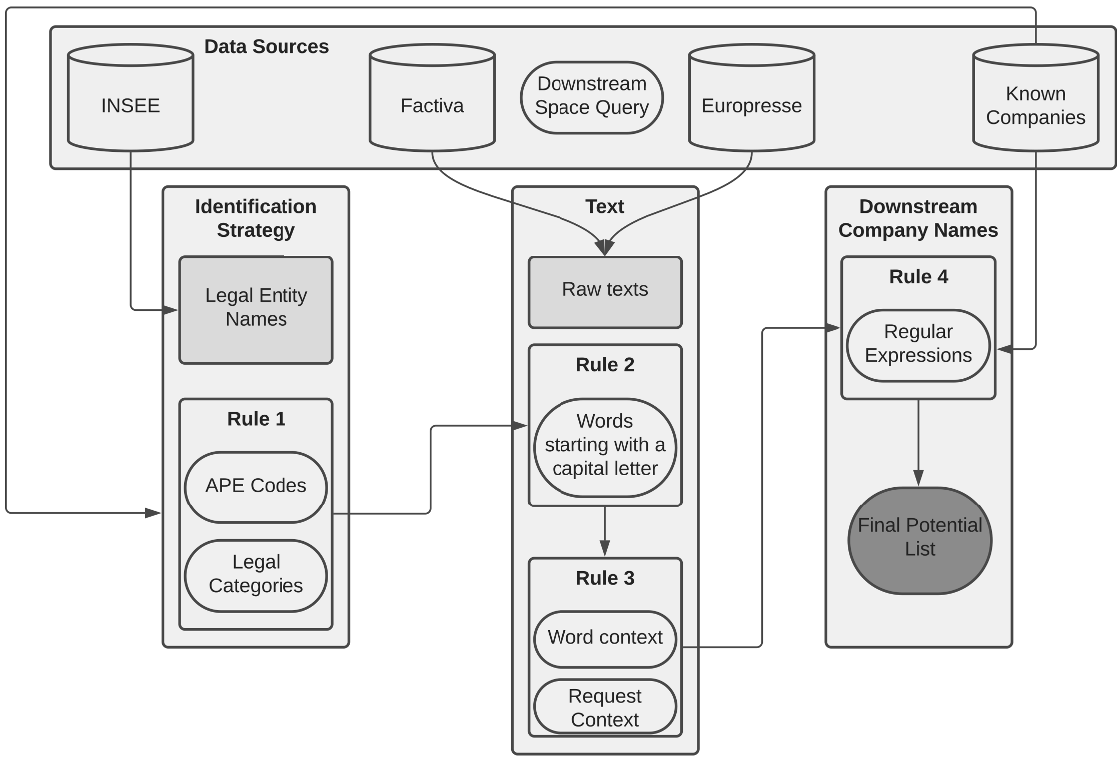


Figure 3.1: Rule-based strategy

3.2.1 Data collection

Our identification methodology depends on four data sources. The first two sources are the electronic press databases Factiva and Europresse, which we used for article collection. Factiva was our primary resource. It is a business-oriented platform offering access to more than 30,000 international sources, including newspapers, journals, business magazines, and newswires. We utilized Europresse as a supplementary source. The coverage of Europresse is European publications. We used this database to enrich our data with French sources not included in Factiva, such as *Le Monde*, *Libération*, and 89 other sources, including regional publications. The most prominent journals are listed in Appendix 3.8.1, Figure 3.7.

The third database, referred to as the *Sirene database* throughout this chapter, encompasses all French-registered companies. This registry, provided as open data by the French national statistics office (INSEE), includes variables like company name, identification code, creation date, and main activity. We selected the legal units file, which pertains to the firm as a legal entity, rather than the establishment file, with our primary interest being company names. This legal unit file contains a variable called 'Dénomination usuelle' for names. The Sirene file is updated monthly, meaning a company registered in a particular month, *m*, will be included in the file for the following month, *m+1*. We used the June 2022 version of the file named 'Fichier StockUniteLegale'[32]. This database housed 23,414,852 records, with 13,766,039 corresponding to active entities. Among these active companies, we identified a total of 6,961,344 unique names.

The final database we utilized, referred to as the 'known companies' database in Figure 3.1, contains the names of 220 French-registered downstream space companies. We assembled this database from several different information sources. Firstly, we leveraged a database on the French downstream ecosystem from a 2016 study conducted by a consulting firm on behalf of CNES. Secondly, we gathered information on downstream space companies that were recipients of the first space section call of the France Relance program in 2021[33]. Thirdly, we sought out information on space companies involved in incubation programs led by CNES, specifically the *CONNECT by CNES* initiative, which the French space agency launched to support space-related businesses[34]. Lastly, we gathered company names from the history of the *ESA BIC* incubation program run by the European Space Agency (ESA)[35].
The known companies database served as a solid foundation for our method because it includes companies definitely involved in downstream space activities. Consequently, we utilized the characteristics of these known companies (names, industry codes, legal status) to construct two rules for our identification strategy.

[32]The file is available at: `https://www.data.gouv.fr/fr/datasets/base-sirene-des-entreprises-et-de-leurs-etablissements-siren-siret/`

[33]France Relance is the economic recovery plan launched by the French government in the wake of the Covid-19 health crisis. The results of the call for the space sector are available at: `https://www.entreprises.gouv.fr/fr/actualites/crise-sanitaire/france-relance/france-relance-premiers-laureats-du-volet-spatial`

[34]The information is available here: `https://www.connectbycnes.fr/en/space-for-good`

[35]The list of company names is available at: `https://commercialisation.esa.int/startups/` Country: France, Space Connection: Downstream

3.2.2 Method and rules description

3.2.2.1 Extraction of newspaper articles dealing with downstream space activities

The first feature of our methodology involved extracting a corpus of text specifically related to downstream space activities from the press databases. To achieve this, we developed a Boolean query. This query formulated a topic that enabled us to retrieve news articles on France's downstream space sector that could potentially mention companies within this sector. In essence, this first step aimed to narrow down the textual data sphere, firstly to avoid detecting company names unrelated to our area of interest as much as possible, and secondly, to reduce the volume of text to analyze. We present the *downstream space query*, along with details on its construction and additional restrictions applied (geographical scope, publication period, source removal) in section 3.3.1.

The query returned 48,900 articles, 28,400 from Factiva, and 20,500 from Europresse. We manually downloaded the articles in HTML format. We restructured the information they contained (source, author, headline, lead paragraph, body text) into variables using web scraping libraries[36]. The number of articles captured per year ranged between 2000 and 2500 throughout the period (Figure 3.2).

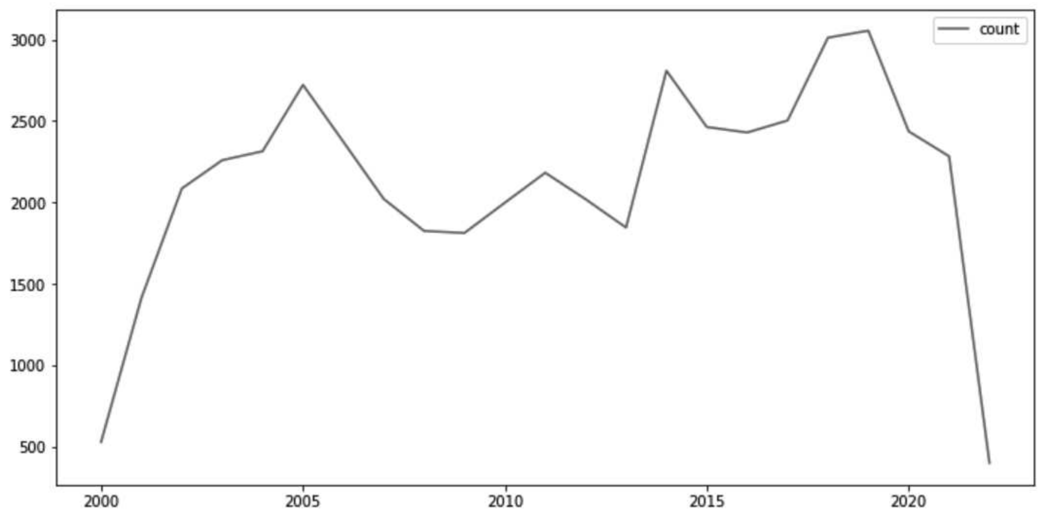


Figure 3.2: Downstream space query results: Number of articles published per year

[36]We carried out all the implementation steps with Python software. However, the methodology is replicable with any other programming software.

Once the text was restructured, we eliminated all words that did not start with a capital letter. This critical step forms the **Rule 2** of the identification method.[37] The remaining text contained only punctuation (if a company's name is an acronym or hyphenated), numbers (if a company's name starts with a number), and words starting with a capital letter (e.g., proper nouns and the first word of sentences). From a computational standpoint, this significantly reduced the number of words to compare with the company names from the Sirene database, consequently shortening the processing time. Methodologically, we assumed that since company names are proper nouns, they would always be quoted in the press with an initial capital letter. With this text modification, we minimized the risk of losing company names unless they were incorrectly quoted in the original article. Furthermore, we observed from the Sirene database that many French company names were very common words. Additionally, news articles often employ non-technical language and simple vocabulary. By eliminating lowercase words from our corpus, we aimed to limit the capture of *false positives*, i.e., instances where a company name matches a word in the text but does not refer to a downstream company. Finally, we removed all accents to ensure harmonization.

3.2.2.2 Matching the newspaper text with the dictionary of French company names

The second aspect of the method involved creating a dictionary of company names from the Sirene database. This task required refining the database of French-registered companies to a subset of companies to further limit the number of companies compared with the press text. In doing so, we defined an additional rule (**'Rule 1'** in Figure 3.1): selecting companies with pre-defined APE ('Activité Principale Exercée') codes and legal categories. Limiting the dictionary of company names to business entities belonging to specific sectors of activity allowed us, in the same way as the query with the press articles on downstream activities, to form a list of companies that are part of a field of activity. We selected APE codes and legal categories based on those of downstream space companies in our internal database. We provide the details of the selected codes in section 3.3.2. It is important to note that a company with a unique national identification number may have multiple names. It is often the case for companies with acronyms or extended full names. The Sirene database provides four

[37]The numbering of the rules does not indicate any temporality in their implementation. Rules 1 and 2 were defined at the same time.

name variables ('Dénomination usuelle,' 'Dénomination usuelle 1', 'Dénomination usuelle 2', and 'Dénomination usuelle 3') which we all considered throughout our procedure. Conversely, two active companies with different identification numbers may share the same name. Despite identical names, we maintained two distinct observation lines in our dictionary.

We converted all company names in the dictionary to lowercase and removed accents for consistency. Furthermore, some names in the dictionary included acronyms indicating the company's legal form. We removed these acronyms[38], as newspaper articles do not necessarily include the legal form when citing companies.

Next, after removing all lowercase words, we used the Sirene database subset– our company dictionary– to detect company names in the articles that came from our downstream space query. We automated this task using a simple algorithm: it compared each company name in the dictionary with each word from the downstream space query text that started with a capital letter. If a word from the article corpus was completely identical to a company name in the Sirene dictionary, it added that company to the list of potential downstream space companies. Given that this was an exact match operation, both databases compared had to be harmonized (all words in lowercase and without accents). We applied the word comparison to the entire text of each downloaded article (headline, lead paragraph, article body). The result was a list of company names registered in the Sirene database that were cited at least once in one or more articles from our downstream space query corpus.

Though limiting our analysis to capitalized words reduces noise, we still faced instances where we detected words in the Sirene list that corresponded to company names but referred to different entities in the articles. For example, 'Paris' appeared in our list because it was a company registered in the Sirene database and fell within our selected industry codes and legal status. However, many articles mentioned 'Paris' about the city. At this stage in our procedure, we identified Paris as a potential downstream space sector company. We faced this issue with place names, geographical areas, individual names, and acronyms. Another challenge is that our article corpus did not exclusively cover downstream space activities. An article might mention a downstream space company in one paragraph but discuss topics and companies unrelated to this sector (and the space industry in general) in the rest

[38] We removed acronyms for Société Anonyme (SA), Société À Responsabilité Limitée (SARL), and Société par Actions Simplifiée (SAS) with and without periods between letters.

of the text. Consider an article from a regional publication about a funding plan for local businesses, which includes a company relevant to our research. If we found all mentioned company names in the Sirene dictionary, we would add them to the list of potential downstream space companies. Consequently, using the query as the only filter applied to press databases did not guarantee that we only retrieve text dedicated to downstream space sector information.

3.2.2.3 Minimizing False Positives by Considering Citation Context

To mitigate the incidence of false positives stemming from the above issues, we implemented a third rule (**'Rule 3'** in Figure 3.1) we call the 'Word Context' and 'Request Context' as shown in Figure 3.1. This rule relies on two lists of words that establish a context for citing downstream space companies. The first list comprises nouns, verbs, and past participles frequently used in sentences that mention a company name, such as 'start-up,' 'founded,' and 'provide.' By assuming these words are near a cited company in the text, we can lower the likelihood of incorrectly matching names from our Sirene dictionary with unrelated terms in the articles. The second list consists of the main keywords from the downstream space query. We ensure that the matched companies appear within a context that aligns semantically with space activities. This step allowed us to narrow down the company identification to the sections of articles dedicated to downstream space activities. After several trials, we determined a 30-word window for both lists (30 words before and after potential company names). We provide both word context lists in section 3.3.4. For this task, we had to revert to the full text from the query (including lowercase words). The computational time was short since, for each article, we used the list of potential downstream space companies resulting from the initial matching.

By this point in our methodology, we had a list of French companies mentioned in press articles stemming from a query on downstream space activities. We implemented a final, stringent fourth rule directly on this list, called 'Regular expressions' (**Rule 4**) in Figure 3.1 and further detailed in section 3.3.5. It includes a set of character sequences reflecting as closely as possible the patterns of downstream company names. We chose the patterns that frequently appeared in known companies' database. While this rule may lead to the exclusion of 'true positives' - downstream companies identified by the matching procedure - it aids in eliminating a large number of false positives from the list. We applied the list of regular expressions to the list of potential companies, yielding a final reduced list for manual sorting by an expert. The sorting process involved three steps: First, we verified the company's active status since it might be in the Sirene file but no longer operating. Second, we check the context of its mention in the raw text. This step sometimes enabled us to validate the company name as the article explicitly described the company's downstream activity or discussed an entirely different topic. Otherwise, we checked the company's website to confirm whether it belongs to the downstream space sector based on the company's description.

The subsequent section delves into each facet of the identification methodology we have devised for this research. First, we offer an extensive explanation of the query constructed to compile a corpus of articles discussing downstream space activities. Following that, we outline each of the carefully formulated rules implemented to distill the names of downstream space enterprises from this corpus.

3.3 Construction of the query and choice of identification rules

The development of our method and rules takes into account two primary issues. The first pertains to the method's processing time. Our approach involves an automated part (matching procedures) and a manual part (sorting the final list of potential downstream companies). The objective is to limit the volume of text the computer processes, thus speeding up the procedure and reducing the number of companies to sort. The second aspect relates to the performance of the method: the text reduced by the query must be relevant enough to identify as many downstream companies as possible. This optimization of text should confine the *semantic universe* processed to the downstream space sector.

We start by presenting the query and explaining how we developed it. Following that, we describe each rule implemented. Two of these rules apply to the Sirene dictionary (Rule 1 and Rule 4), while the other two apply to press article texts (Rule 2 and Rule 3).

3.3.1 The downstream space query

The motivation behind building a query is to formulate a search topic within press databases and delimit a perimeter for identifying downstream companies. Its correct formulation is crucial to ensure the performance of the method. Indeed, the matching with the dictionary of company names is performed with the articles from the downstream space query. It must be developed or verified by experts informed on the latest space sector developments, including downstream activities.

Consequently, we dedicated considerable attention to the formulation and optimization of the query. The more relevant it was, the more likely we were to find companies mentioned in the resulting articles. This step required, on the one hand, the incorporation of words and expressions that covered the activities we were interested in as comprehensively as possible. On the other hand, the query needed to be sufficiently restrictive to exclude articles dealing with activities related to, but not directly part of, the space sector (e.g., non-satellite geographic information) and those dealing with space systems manufacturing. The final query is in Box 3.3.

(1) [satellit* NEAR4 (application* OR service* OR solution* OR operat* OR donnee* OR data OR imag* OR communication* OR telecommunication* OR broadband OR broadcast* OR connectivity OR diffusion OR telediffusion OR cartographi* OR geoinformation OR geo-information OR (information geographique) OR geoloca* OR geoposition* OR geo position* OR position* OR navigation OR surveillance OR monitoring OR tracking)

OR earth observation OR observation de la terre OR teledetection OR remote sensing

(2) OR (downstream space NEAR4 (industr* OR compan* OR provider* OR sector* OR market* OR application* OR service*))
OR (((secteur* OR industrie* OR economie* OR segment* OR ecosysteme*) W/1 spatial*) NEAR30 (aval OR applications)) OR (service* a valeur ajoutee NEAR30 (spatia* OR satellit*))]

(3) AND (francais* OR french OR france)

Figure 3.3: The downstream space query (with Factiva language)

Our study focuses on companies registered in France. Therefore, we formulate a query with French and English keywords to capture articles from the foreign press in case the method is replicated by including English-language publications dealing with the French downstream market. We developed a single query with words used in English and French, and others specific to each language. However, for this first application of the method, we restrict our search to articles written in French. The addition of English-specific keywords did not change the number of results.

The query is organized into three distinctive parts. Part (1) in Box 3.3 allows collecting articles referring to activities specific to each application domain (communication, satellite imagery, and navigation). Most of these words taken in isolation do not refer only to space. Therefore, we specified they must not be four words away from *satellit**. We preferred this term to *space* (or *spatial* in French) to avoid collecting papers on non-satellite geographic data. Two expressions (*remote sensing/teledetection* and *Earth observation/Observation de la Terre*) are used almost exclusively in the space domain and are therefore not subjected to this restriction.

With part (2) of the query, the goal was to obtain publications that inform about downstream space activities in a more general way. Contrary to the first part, the vocabulary used is different in French and foreign publications. After several attempts and analyses of the database, we confirmed that 'downstream space' is the most widely used expression in articles written in English to qualify activities related to the exploitation of satellite data. In addition, we introduced as a constraint that *downstream space* must not be three words away from *industr* OR compan* OR provider* OR sector* OR market* OR application** to collect articles tackling the commercial nature of these activities.

We adopted a similar approach for the French part. The nuance is that the expression *downstream space* (*spatial aval* in French) is not as common as in the English-speaking press. Therefore, we used the same words as in English to obtain articles on commercial space activities (*secteur* OR industrie* OR economie* OR segment* OR ecosysteme**) followed directly by *spatial** to specify the domain of activity. Then, we indicated that this expression must not be thirty words away from *aval* or *application** (i.e., downstream or application) to target the type of activities we were looking for in the space sector. The distance of thirty words ensures that the words are cited in the same paragraph. Finally, we added the expression *service* a valeur ajoutee* (i.e., value-added service) in the same paragraph as *spatia** (i.e., space) or *satellit**. It is commonly used in French to describe this type of activity, especially in telecommunications.

The third part of the query consists in adding as a constraint to the two previous blocks the words *French* (or *Français**) or *France* in the articles. We tested several options to target articles that cite French companies, such as using the region criterion proposed by the interface. However, introducing the geographic limitation directly into the free text allowed more results. This also left the possibility of using the query on other article databases.

In addition to the boolean query, we added additional criteria for the Factiva search. We limited the article language to French, thus including foreign press written in French. We removed the sources *EUR-Lex* and *Le Mensuel d'Agefi Luxembourg*. The former is the official journal of the EU, and the latter is a Luxembourg newspaper dealing with European economic and financial news. Although they are potentially rich sources of information, the articles in these publications were very long. In the first matching tests with the Sirene dictionary, we detected thousands of company names in each paper. In addition to adding much noise, the size of the articles significantly slowed computation time.
We started the search with articles published from January 1, 2000, to limit the number of papers to download because this period corresponds to the premises of the New Space. Finally, we removed duplicates, republished news, recurring pricing, market data, obituaries, sports, and calendars.

We performed the last query in Factiva on February 12, 2022. It resulted in 28,400 articles published between 01/01/2000 and 12/02/2022. We uploaded the HTML pages of the articles.

In addition, we exploited the Europresse platform as a secondary newspaper data source. Our version of Factiva did not give access to the publications of the French daily newspapers *Le Monde* and *Libération*. Moreover, Europresse includes many additional regional sources, particularly those dealing with South-West of France news in which the space sector is established. We used the same query as in Factiva, removing the France restriction (block (3) in Box 3.3) since we were only looking for articles published in French sources. French articles may deal with another country's downstream space sector news, but we assumed this noise was limited considering the selected sources (mainly regional publications). We also removed the English part in Block (2) and adapted the query formulation to the Europresse language.

As with the Factiva database, we uploaded the HTML pages of the 20,500 articles from the query published between 01/01/2000 and 12/02/2022.

3.3.2 Rule 1: Industry Codes and Legal Categories

The selection of industry codes and legal categories is the **first rule** of the identification method. We applied it to the Sirene database to create a "Sirene dictionary," which contains only companies with activities and legal categories corresponding to those of known companies. This operation aimed to reduce the number of company names compared with the text from the downstream space query and the computation time.

We have seen in Chapter 2 that industrial classifications, except for the satellite telecommunications segment, do not alone allow the identification of firms with a downstream space activity. Most share an activity code with firms unrelated to the sector under consideration. However, we have noticed that specific industry codes were recurrent in our internal database of downstream companies. These are related to information and communication activities, specialized, scientific and technical activities, and business support activities. We therefore applied the list of industry codes from our internal database to the Sirene company database to create a dictionary of companies whose industry codes corresponded to those of known downstream companies. Table 3.1 lists the industry codes (Activité Principale Exercée (APE) codes) selected to build the Sirene dictionary.

The list includes thirty-three activity codes out of the 732 sub-classes of the French industry classification (Nomenclature des Activités Françaises). We deliberately filtered the Sirene database with the most detailed code level. Indeed, the NAF is organized into five levels: the section with one letter, the division with two digits, the group with three digits, the class with four digits, and the sub-class with four digits and one letter. Each level provides a more detailed description of the activity of companies in the sector. The more precise the code, the fewer the number of firms that match that code. We selected sub-classes so the rule was restrictive enough, and the dictionary comprised a few companies.

Our code selection covers a wide range of activities. It includes activities related to equipment: the manufacture of navigation and communication equipment (divisions 26 and 30 in Table 3.1) and the sale of equipment (divisions 46 and 47). We naturally kept activity codes related to telecommunications (61 division). Division 52 includes known downstream space companies delivering telecommunication and navigation services in the transportation sector. Division 58 corresponds in our known companies database to downstream companies providing mapping services and software integrating satellite data (e.g., for agriculture). Known companies with an APE code of divisions 62 and 63 provide meteorological services, data processing, and Earth observation services. Divisions 71, 72, and 74 refer to engineering, scientific, and technical activities, but the known downstream space companies in these sectors are very similar to those of divisions 58, 62, and 63. Division 66, referring to insurance activities, stands out from the other selected sectors. One known downstream company with this APE code develops parametric insurance using Earth observation data.

APE code	Description
26.30Z	Manufacture of communication equipment
26.40Z	Manufacture of consumer electronics
26.51A; 26.51B	Manufacture of instruments and appliances for measuring, testing and navigation
30.30Z	Manufacture of air and spacecraft and related machinery
46.51Z	Wholesale of computers, computer peripheral equipment and software
46.52Z	Wholesale of electronic and telecommunications equipment and parts
46.90Z	Non-specialised wholesale trade
47.78C	Other sundry specialized retail sale
52.21Z; 52.23Z	Service activities incidental to land transportation; to air transportation
58.29B; 58.29C	Development tools and programming languages software publishing; Application software publishing
61.10Z; 61.20Z	Wired telecommunications activities; Wireless telecommunications activities
61.30Z	Satellite telecommunications activities
61.90Z	Other telecommunications activities
62.01Z	Computer programming activities
62.02A; 62.02B	Computer consultancy
62.03Z	Computer facilities management activities
63.11Z	Data processing, hosting and related activities
63.99Z	Other information service activities
64.20Z	Activities of holding companies
66.22Z	Activities of insurance agents and brokers
70.10Z	Activities of head offices
70.22Z	Business and other management consultancy activities
71.12A; 71.12B	Engineering activities and related technical consultancy
72.19Z	Other research and experimental development on natural sciences and engineering
72.3Z	Data processing
74.90B	Sundry professional, scientific and technical activities
82.99Z	Other business support service activities

Table 3.1: Industry codes selected for the Sirene dictionary

Legal Status	Description
3120	Foreign commercial company registered with the RCS
5499; 5460	Société à Responsabilité Limitée (SARL); Other Cooperative SARL
5599; 5699	Société Anonyme (SA) with a board of directors
5710	Société par Actions Simplifiée (SAS)
5800	European company
6599	Civil company

Table 3.2: Legal status selected for the Sirene dictionary

We performed the same operation with legal categories. In other words, we listed the different legal statuses of known downstream companies and filtered the Sirene database by restricting the firms belonging to these legal categories. The list of legal status codes is in Table 3.2. This part of the rule was not very discriminating since most French firms belong to the following three legal categories: Société à Responsabilité Limitée, Société Anonyme, and Société par Actions Simplifiée. Nevertheless, it allows for removing all non-profit organizations such as associations, trade unions, and state administrations.

After implementing the rule to the French company database, we obtained a Sirene dictionary that included companies belonging to the selected sectors of activity and whose legal category corresponded to one of the legal codes in Table 3.2. Initially, the Sirene database contained 6.9 million legal units still active. **The Sirene dictionary after the APE and legal category rule included 650 thousand observations**. The next step of the method was comparing each company name variable of the Sirene dictionary with each capitalized word from the downstream space query. When an exact match occurred, we saved the company name in the potential list of downstream companies.

3.3.3 Rule 2: Words starting with a capital letter

The third rule also applies to the text resulting from the downstream space query. As described in section 2, it involves keeping only the words from the text that start with a capital letter. We have kept punctuation, symbols, and numbers to avoid losing company names that contain them (for example, acronyms with dots between each letter). This step of the method was also an important source of noise reduction. It prevented a company from the Sirene dictionary with a name also used in everyday language from being matched. For example, the company named 'Sun' could only be matched if the word 'sun' was cited with a capital letter in the text.

Besides, this rule allowed us to reduce the text volume to be compared with each company name from the Sirene dictionary. Therefore, the computation time was lower than if we kept the entire press articles.

3.3.4 Rule 3: Word Context and Query Context

The Word Context and Query Context rule (**(Rule 3)** relies on a list of names from business and downstream space activities lexicons. We applied it to the text from the query to reduce the scope of identification of downstream companies to a particular semantic context. During the development phases of the method, we realized that removing only lowercase words in the query text and building up the Sirene dictionary was not efficient enough to substantially reduce the number of false positives. Many companies registered in the Sirene dictionary matched a word in the query text when they were not downstream space companies. We decided to reduce the detection scope of company names in the text. Therefore, we implemented a second matching procedure: We compared the potential list of downstream company names resulting from the first matching with the raw text (with lowercase words). The constraint was the following: names of the potential list must appear in the text within thirty words of at least one of the words of Table 3.3 **or** one of the words of the query (Box 3.3). The window of thirty was the most efficient in terms of filtering and known companies kept. Generally, a thirty-word window requires that words are in the same paragraph.

Word class	Words
Nouns and groups of nouns	entreprise*, start-up*, startup*, société*, PME, TPE, spin-off*, spinoff*, filiale*, groupe*, pépite*, jeune* pousse*, licorne*, acteur*, spécialiste*, le bureau d'études, opérateur*, fournisseur*, client*, fondateur* de, fondatrice* de, co-fondateur, cofondateur*, cofondatrice* de, co-fondatrice* de, directeur général de, directrice générale de, dirigeant* de, gérant* de, consortium*, PDG, DG, SA, SAS, SARL, créateur, entrepreneur, incubateur, SATT, pôle de compétitivité, booster, cluster, ecosysteme, business model, modele d'affaires, business plan, capital risque, levée de fonds
Verbs	conçu* par, créé*, fondé*, a développé, développe*, fournit, fournissent, opère*, spécialisé* dans, spécialisé* en

Table 3.3: List of context words

We built the word list (Table 3.3) using a sample of five hundred articles from the downstream space query. In each article citing a downstream company name, we analyzed the context of its citation, i.e., the words preceding and following it in the paragraph. We identified nouns, verbs, and past participles frequently mentioned in proximity to downstream company names. Some nouns refer to the company in general (e.g., *entreprise*, *start-up*, *filiale)* (subsidiary), others refer to company activities (e.g., *opérateur*, *le bureau d'études* (the consulting firm)). The list also contains words that refer to the company management (e.g., *fondateur* (founder), *directeur général* and *PDG* (CEO), *entrepreneur*). In the case where an article cites several downstream companies, we selected words or expressions that refer to a group of companies (e.g., *pôle de compétitivité* (competitive cluster), *ecosysteme*). The verbs often describe the company (*créé* (created), *fournit* (provides, delivers), *spécialisé dans*, *opère*). We included all word forms (masculine/ feminine, singular/plural) in the analysis.

3.3.5 Rule 4: Downstream Regular Expressions

Once the second word comparison is performed on the text restricted to the semantic context (Word Context rule, section 3.3.4), we obtain a final reduced list of potential downstream company names. However, the list is still too extensive for expert hand-sorting. We introduce a final rule we call Downstream Regular Expressions. It involves selecting only companies detected in the text whose names contain one of the recurring character strings of downstream space company names. This operation is drastic because it considerably reduces the number of companies to sort out by eliminating false positives. However, it also removes true positives, i.e., downstream space companies whose names contain no regular expression. In this respect, we use it with great care: the companies resulting from this rule are only a sample of the total company list where downstream companies are potentially over-represented. It allows us to process a decent number of companies. We therefore do not entirely exclude from the analysis companies that do not belong to this sample.

Regular expressions
agr, data, e-, farm, geo, ima, lab, map, nav, ocea, sat, sea, service, solution, space, system, tech, tele, terr

Table 3.4: Downstream regular expressions

Table 3.4 lists the regular expressions defined. We exploited the database of known downstream companies to build the selection of character sequences in which we observed recurring identical expressions among company names. Some patterns refer directly to the space sector (*space, sat*), and others give an indication of the application sector of the company (*agr, geo, ima, map, nav, ocea sea, terr*). Finally, some selected expressions are related to company's activities (*data, lab, service, solution, tele*).

After applying the rules outlined in this section, we obtained a reduced list of **companies that potentially belong to the downstream space sector**. This list results from the matching between the Sirene company dictionary and press articles on the topic of downstream space.
We now move on to the results of the first application of our rule-based identification method.

3.4 Results of the first implementation and new downstream space companies identified

In essence, the first application of our identification method aimed to discover new French downstream space companies, that is, companies that were not known prior to its implementation. In other words, we seek to determine whether the method has successfully enriched the 'known' downstream space companies' database.

Before presenting the final results, let us review the intermediate results, i.e., the number of companies obtained after applying each rule. The application results and the various figures given here are summarized in Figure 3.6.

Initially, we had a file from the Sirene database containing 6.9 million active legal units registered in France. The application of **Rule 1** (APE Codes and Legal Categories) reduced the number of companies to match with the query text by 90%, resulting in the *Sirene dictionary* containing 650,000 observations.
Next, the list of companies obtained after matching the Sirene dictionary with the query text (only words starting with a capital letter, **Rule 2**) included 30,084 companies. Thus, this step in the method allowed for a 95% reduction in the number of companies to sort compared to the Sirene dictionary.
To apply the following rule, we remind that we returned to the raw text of the downstream query to determine the citation context of the 30,084 companies obtained

in the previous step. By restricting the citation window with the Word Context and Query Context lists to 30 words, we reduced our list of potential downstream space companies to 22,862 observations.
Finally, the fourth rule was to keep only company names that contained a 'regular expression.' This rule was drastic, as it allowed us to obtain a final list to sort composed of 1,475 companies.
Therefore, our rule-based identification method resulted in a reduced list of companies (from 6.9 million to 1,475) potentially involved in downstream space activities.

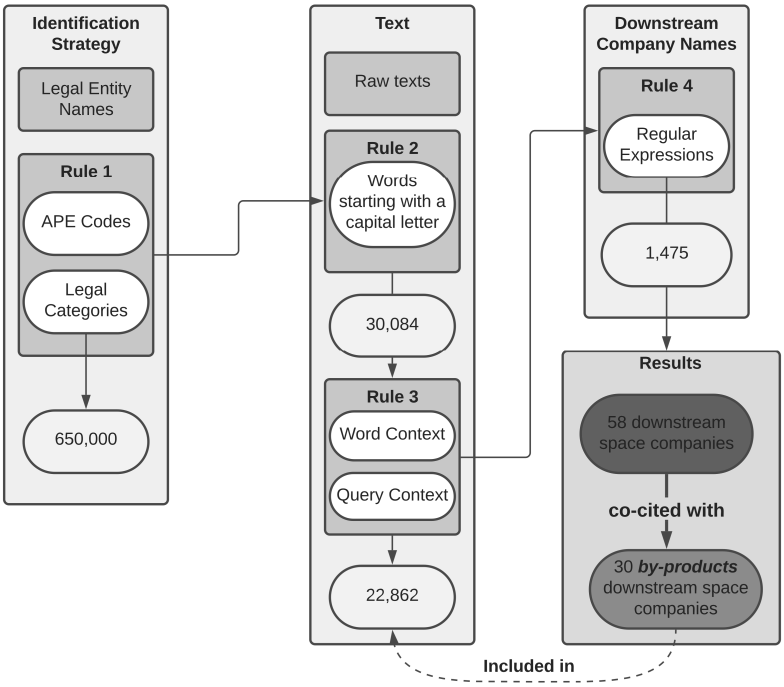


Figure 3.4: Results of the first implementation

The final step in the identification method was to manually sort the 1,475 companies obtained by applying the rules. The procedure for determining whether a company was effectively part of the downstream sector was in two stages:

1. We returned to the press articles in which these companies were mentioned to see if we could directly classify them as downstream space companies. The sorting procedure stopped here if the article included a description corresponding to a downstream activity.

2. If the citation context in the press was not sufficient, we conducted a search on the company's website. If specific keywords appeared in the description of its offer (e.g., 'satellite,' 'earth observation image,' 'GPS,' 'GNSS'), we classified it as a downstream space company.
   In some cases, companies were mentioned in application projects or the results of calls for projects led by institutions (CNES, ESA, etc.). This information was also taken into account when sorting the companies.

The strict implementation of the method allowed us to identify 58 new downstream space companies. In other words, 4% of the companies in the list of company names containing regular expressions were indeed downstream space firms. Given the restrictive rules applied (particularly the last 'regular expression' rule), we find this result satisfactory.

In addition, we detected 30 new downstream space companies we call *by-products* of the method. These companies did not contain a regular expression in their name but appeared in the list of potential companies from the first three rules. During the manual sorting procedure, we noticed papers citing some companies from the list of 1,475 with other companies included in the list of 22,862. Co-citation was of several kinds: companies conducting similar activities or in the same geographical area, participating in the same program or call for projects, or having supplier-client relationships.

Source	Number of companies
Known database	
Reference database	220
Other*	26
Rule-based method	
Method outputs	58
By-products	30
Total	334

* "Other" refers to companies identified through online research after the method was developed. We will include them in the evaluation stage.

Table 3.5: Summary of the final downstream space company database by source of identification

Eventually, we obtained 88 new downstream companies using our

rule-based named entity recognition approach. Table 3.5 summarizes the final number of downstream space companies that form our database by source of identification. The 'known database' includes the companies we did not identify through our identification method. This database is split into two parts: the 'Reference database,' which served as our basis for formulating the rules, and the 'Other', comprising 26 companies that we discovered via alternate ways (like Internet research and participation in downstream space-focused events) **post** method application. We have included these companies in our evaluation in the next chapter. The 'Rule-based method' column corresponds to the companies that we detected through our approach. Overall, our database of downstream space companies, compiled in 2022, gathers 344 companies, more than 26% of which were detected using our method. We will assess the economic weight of the French downstream space sector in 2021 from this database in Chapter 4 of the thesis.

3.5 Calibration and Evaluation of the rule-based identification method

The identification method presented in this chapter results from a series of adjustments and tests to ensure its effectiveness. We started by verifying critical steps of the identification procedure to evaluate their robustness. In the absence of existing studies offering a comparable method for detecting downstream space sector companies, we determined the performance of our process by comparing it to a statistical approach to Named-Entity Recognition.

In this section, we first focus on a crucial step of the method, the query formulation. Then, we evaluate the performance of the rules specifically applied to the text and compare it with the statistical approach.

3.5.1 Query verification

To verify the query's ability to obtain news articles that contain downstream space company names, we used our 'Known companies' database. We analyzed how many known companies the query could detect to understand how this step contributes to the overall performance of our approach and where the potential weak points might be for future applications of the method. As a reminder, our method relies not on a machine learning model but on the simple application of rules in newspaper text. Therefore, we have not split our 'known companies' database into a learning and test subset, as our approach has no machine learning process.

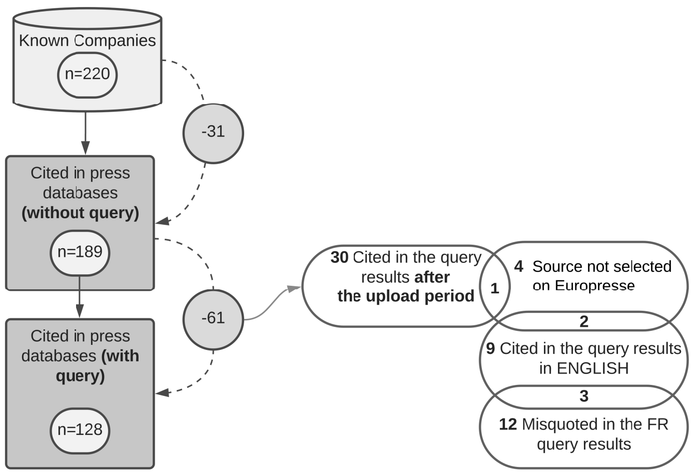


Figure 3.5: Loss of known companies upon query application.

Figure 3.5 illustrates the number of known companies before and after the application of the method and indicates the sources of loss after the method's application. The initial known company database included 220 known companies (Figure 3.5). The first step was checking whether known companies appeared at least once in the two press databases without applying any restriction. We found that 31 of the 220 companies were neither mentioned in Factiva nor Europresse. In other words, our method could not capture 14% of known downstream companies since they did not appear in any newspaper text from the databases. Several explanations can be assumed for this. Firstly, these companies were not well-known and did not make the news to appear in the press. Secondly, these companies are still too young and their activity too recent to be the subject of an article.

In the second step, we examined how many known companies among the 189 remaining were cited in the downstream space query results. We expected this part of the method to be an important source of loss since we reduced the text from all available published electronic articles to a corpus restricted by keywords, publication period, and language. Eventually, we found 128 - nearly 70% - of the remaining known companies mentioned in the papers from the query. Looking at further details on the losses, we observed that half of the companies were not cited in the query keyword context but appeared in the query results **after** the upload period. It supports our previous hypothesis, namely that there is a delay between the company's creation date and its mention in the press. It is the case for companies that are not part of space programs or not involved in calls for proposals, for example.

Misquotes are the second most significant source of losses: 15 known companies were mentioned in the query results but not under the exact name of their legal unit registered in Sirene[39]. To illustrate this point, we had the case of a company whose official name, i.e., defined in the Sirene file, was '*X* France.' This company was mentioned several times in the query but without the word 'France.' Thus, it did not appear in the raw text from the query. Finally, 14 known companies lost were mentioned in the Factiva query results but only in papers written in English. In addition, we lost 6 known companies by forgetting to select sources on Europresse.

In the following, we detail the part of the evaluation focusing on the formulated rules.

3.5.2 Rules performance measurement and comparison with the statistical approach

The second part of the assessment relied on the 128 remaining known companies mentioned in the French query results. First, we describe how we assessed the performance of the applied rules in filtering company names relative to how many known firms they identified. Next, we examine the results of applying the spaCy named entity recognition model to our text and compare them with our method. Figure 3.6 summarizes the two evaluation procedures.

[39] In figure 3.5, 12 known companies were misquoted in the French query results. Three were mentioned in the English query results **and** misquoted in the French query results.

3.5.2.1 Rule performance evaluation

We applied two rules to the downstream space query text: keep only the words that start with a capital letter (Rule 2) and keep only the words within a maximum distance of 30 words from the Word Context or the Query Context lists (Rule 3). Figure 3.6 provides the numbers of companies cited in the text that matched those from the Sirene dictionary, as well as the number of known companies detected after the application of each of the two rules (path 'Rule-based NER,' box 'Evaluation').

We thus measure the performance of each rule by two ratios. The first one evaluates the filtering capacity of the rule. We obtain it by dividing the number of companies after the application of the rule by the number of companies before the application of the rule. The second one evaluates the rule's capacity to detect known companies. We obtain it by dividing the number of known companies detected after applying the rule by the number of known companies before applying it. **The lower the first ratio (hereafter *Filtering ratio*) and the higher the second ratio (hereafter *Conservation ratio*), the more effective the rule.**

The matching between the Sirene dictionary (Rule 1) and the query text by applying only the rule on capital letters to the raw text (Rule 2) resulted in a list of 30,084 companies potentially involved in the downstream space sector. Initially, the Sirene dictionary contained 650,000 companies. In parallel, applying the capital letter rule and matching the remaining text with the Sirene dictionary almost did not impact the number of known companies identified by the method, with 126 companies detected out of the 128. The two lost companies had names composed of two words, the second written in lowercase. **The Filtering ratio of Rules 1 and 2 is 0.04, compared to 0.98 for the Conservation ratio.**

Then comes the application of Rule 3, which consists of keeping only the companies mentioned within a distance of less than 30 words from one of the words in the list of Word Context or Query Context (see Section 3.3.4). After applying this rule, we obtained 22,863 potential downstream space companies and detected 120 known downstream space companies. **Therefore, the Filtering ratio of Rule 3 is 0.76, and the Conservation Ratio is 0.95.** Compared to Rules 1 and 2, the Filtering ratio of Rule 3 is very high, suggesting a low power of filtering. We tried reducing the window size of words ($w = 15$ and $w = 10$), but we lost too many 'known' companies. It could suggest that the small windows caused us to lose too many companies potentially involved in downstream activities. Moreover, we decided that the company name must be close to one of the words in the Word context list **OR** Query Context list, which may explain the slight decrease in the number of companies in the potential list. We tested the application of the intersection of the two lists, resulting in a significant loss of known companies.

Overall, Rules 1 and 2 are very effective as they have a significant filtering capacity while preserving the number of known companies. Rule 3, even though its filtering capacity is lower, allowed us to reduce our list of companies to manually sort by nearly 25% compared to the list obtained after applying Rules 1 and 2.

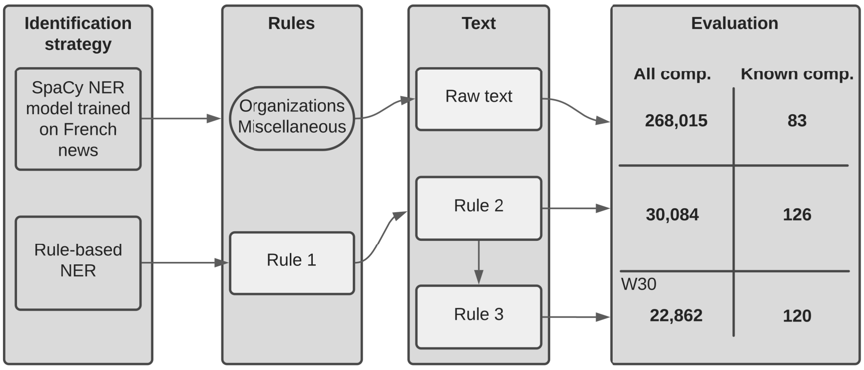


Figure 3.6: Method evaluation: Rules performance and comparison with the statistical approach

3.5.2.2 Comparison with the statistical approach

We compared our results with a statistical approach to have a more general assessment of our rule-based approach. As establishing a set of handcrafted rules to identify downstream space-related companies is time-consuming, we measure the benefits of this method compared to a more straightforward approach that does not rely on explicit rules. The statistical method has the advantage of overcoming spelling inconsistencies as it does not rely on a company names dictionary. Therefore, this approach can recognize company citations in newspaper articles even if their names have been misquoted.

We used *SpaCy*'s pre-trained named-entity recognition model to directly identify companies mentioned in the articles that resulted from the downstream space query (see path 'Spacy NER model trained on French news' in Figure 3.6). *SpaCy* is an open-source Python library designed to simplify complex NLP tasks. In simple terms, NLP is an interdisciplinary field between computer science and linguistics, which allows machines to read, understand, and extract meaningful information from human language.
SpaCy's functionality is vast. Its applications range from extracting information from text and simplifying text input to more advanced tasks such as interpreting the semantics of a given text. It assists in understanding human language and detecting significant information from raw text. SpaCy provides a highly efficient statistical system for NLP in Python, which can assign labels to groups of adjacent words. It provides a default model that can identify a wide range of named entities, including people, organizations, places, and miscellaneous items. In addition to these default entities, SpaCy allows adding arbitrary classes to the model by training it with newly provided examples.

In our research, we used SpaCy's pre-trained model for the specific task of named entity recognition, with the entity in question being company names. There are many different and efficient NER tools [Jiang et al., 2016], but we chose SpaCy as it offers a pre-trained model on French news [Jabbari et al., 2020]. Additionally, SpaCy offers a deep learning implementation to obtain dynamic word embeddings, meaning it provides words with a dense vector representation depending on the context[40].

Eventually, in applying SpaCy to our text corpus dealing with downstream activities, **the NER model labeled 268,015 entities as either organizations or 'miscellaneous entities'** (Figure 3.6, box 'Evaluation'). We added the 'miscellaneous' category in addition to organizations to maximize the number of already known companies captured. However, **only 83 known companies were labeled by the algorithm.**

[40]Documentation is available here: `https://spacy.io/universe/project/video-spacys-ner-model`

In comparison, our rule-based approach reduced the number of company names to sort to 22,862 and detected 120 known companies out of 128. SpaCy model labeled a significantly higher number of entities but identified fewer known companies than the rule-based approach. This result suggests that our rule-based approach demonstrated a better ability to filter company names and detect known companies. This comparison sheds light on the importance of expert work in constructing an identification method using named-entity recognition tools. In the case of identifying downstream space companies, the rule-based approach and the formulation of customized rules proved more effective than the purely statistical approach.

3.6 Time considerations in the application of the method

One crucial aspect of methodological development that we still need to discuss is the implementation time of the method. As this is a first application, we distinguish here between the time taken to construct the method and the time taken to apply the method. The construction of the method was quite time-consuming as it included the painstaking step of developing customized rules. Before we arrived at the version of the method outlined in this chapter, we tested several different rules and went through a series of rule adjustments. The aim was to minimize computation time while maximizing tool performance regarding identification. Thus, the method development phase extended over a year to arrive at the version proposed in this chapter.

In contrast, the strict application of the method once it was operational was much shorter. The two longest steps were naturally those done manually, i.e., downloading press data at the start of the identification procedure and manually sorting the list of companies resulting from applying the rules at the end of the procedure. For reference, we provide information on the application time per step of the procedure:

1. Downloading press articles: 28,400 from Factiva and 20,500 from Europresse.[41]
   Estimated time: 4 days

2. Matching procedure and rule application (completely automated part of the method).
   Computation time: 3 days

[41] We had access to Factiva, which allowed us to download articles in batches of 100 only. On Europresse, we downloaded articles in batches of 1,000.

3. Manual sorting and verification of the 1,475 companies obtained after applying the rules.
   Estimated time: 5 days

Therefore, we estimate **a total duration of 12 full days for the application of the method**. This duration is variable depending on the context of the application. Suppose it is a first application in another country. In that case, this duration may be longer depending on the number of articles to download and the adaptation of the query if the expert uses other sources for downloading press data. In addition, the list of regular expressions (Rule 4) could change depending on the known database where the method is applied. If it is simply a matter of replicating the method for the French downstream space segment, then this application time could be significantly reduced as the number of articles to download will be fewer (published after 2022 only). If the text to compare with the Sirene dictionary is less voluminous, one can also hope that the list of companies to check at the end of the rule application will be smaller.

We provide recommendations at two levels regarding the frequency and modalities of reproducing the method. Indeed, our methodological developments occur in the particular context of New Space, which we have described as a paradigm shift in the space industry. Thus, we have proposed a tool suited to identifying downstream space companies, given the current specifics of this activity segment. The context is considered in the choice of vocabulary for the downstream space query (also used in Rule 3), in the selection of activity codes for Rule 1 to form the Sirene dictionary, and in the choice of regular expressions with Rule 4. However, new changes may occur in the downstream space sector, with the emergence of new applications in diverse markets and the entry of new-profile companies into the downstream market. In this case, the method will need to adapt to this change.

Hence, the first level of recommendations refers to cases where the expert perceives the downstream space segment as stable compared to the last identification procedure. Without any visible structural change, it involves simply reproducing the identification method to enrich the downstream space database at regular intervals. In this case, **we recommend reproducing the method annually.** Therefore, the application time is very short since the procedure only includes a year's publications on downstream space activities. In addition, the expert keeps the same formulations for the rules and the query. Therefore, the download, computation, and sorting times are significantly reduced.

However, the expert may notice a critical change in the structure of the downstream segment following a particular shock. For instance, we can imagine the announcement of a large-scale public policy on commercial space applications or a large company investing in developing a new use for space. In this case, some aspects of the method will need reconsideration. The first point is the query formulation, which should be adapted to the new context and the newly emerged downstream activities. Then, the industry codes may need adapting (Rule 1). The expert should analyze the new context and add activity codes to include in the Sirene dictionary. Finally, the rule on regular expressions (Rule 4) might also need reviewing based on the new downstream players. The only stable rule is the one on words starting with a capital letter.

3.7 Discussion and areas for improvement

This chapter was dedicated to the first part of our methodology for evaluating the downstream space sector in France, focusing specifically on identifying the actors within this activity segment. The objective was twofold: to propose a method for detecting companies with downstream space activities and to conduct an initial test of this method. We used Named Entity Recognition tools to extract from a corpus of press text dealing with downstream activities the named entities that we associated, or *labeled*, with the pre-defined category 'downstream space companies'. Our approach was rule-based, wherein we developed a series of customized rules to extract a list of organizations mentioned in the press articles as small as possible and containing the most downstream space companies.

The first rule was to match each word in the text with a Sirene dictionary containing companies with a main activity (industry code) corresponding to that of already known downstream companies. The second rule was to keep only the words starting with a capital letter in the press text, which is generally a characteristic of proper nouns. The third rule involved keeping only the words matched after the previous steps at a maximum distance of 30 words from those appearing in a list of context words or query context. Finally, the last rule aimed to reduce the list of potential downstream space companies by applying a constraint of frequent character sequences, or regular expressions, in downstream company names.

Following the first application of the method, we obtained 88 newly detected downstream space companies. Two-thirds of them were identified through the strict application of the method. The remaining third are 'by-products' of the method and are companies that we identified during the sorting phase of the companies from the list obtained after applying Rule 4. These by-products are co-cited with the companies detected by the method and are in the intermediate list obtained after applying Rule 3. Thus, our database of downstream space companies is composed of 334 companies: 220 companies are known companies used to build the method, 88 companies are companies detected via the rule-based approach, and 26 'other' companies were identified via other sources after applying the method. Eventually, the rule-based approach has allowed us to enrich our database of downstream companies by more than a third.

The proposed approach did not arise *ex nihilo*; we started with a database of known companies and used information about these companies to develop rules and identify previously unknown downstream space companies. Similarly, the method aims to enrich the downstream space database. The tool provides a procedure for regularly updating the downstream space database with companies not yet identified as involved in downstream activities and recently mentioned in the press for this purpose or with newly created, previously unknown companies just mentioned in the press.

A second distinctive aspect of the method is that it heavily relies on expert intervention. This aspect is common when adopting a rule-based approach to NER, as this approach suggests the implementation of directives. The expert is involved in several stages: formulating the downstream space query, elaborating the rules, and sorting the potential downstream space companies' list after matching. Therefore, the method contains both an automated part and a part where the expert's intervention is required. The expert is involved both in applying the method as is and in potential future adjustments we will tackle later.

Lastly, we present a few areas for improvement. One aspect of the method is its heavy reliance on human judgment. Its application is limited to the expert's level of knowledge about the downstream space sector. Additionally, the formulated rules might seem strict and restrictive. The method's ability to capture the downstream space segment's evolution is questioned in this context. The method, as outlined in this chapter, could be a starting point toward a completely automated methodology for labeling downstream space companies. Indeed, machine learning algorithms could accelerate the process and regularly propose a list of potential space-related companies to be verified. This way, we could avoid implementing strict rules that are both restrictive and lengthy to establish.

However, as we have seen when comparing the rule-based method to the statistical approach, a method relying solely on the latter would not guarantee results. A potential extension to our work is to merge the two approaches, using first the rule-based approach to build a dictionary of space-specific company names to train a statistical model. For instance, [Jafari et al., 2020] trained a model themselves and improved the statistical approach's ability to detect entities related to the satellite domain. They used the spaCy model and predefined three categories related to the satellite domain: organizations, rockets, and satellites. They outperformed state-of-the-art NER techniques to capture items in the Satellite domain.

One aspect to consider is that the diversity of companies related to downstream space activities makes it less direct to construct a dictionary to label a corpus and train a model to capture companies specifically related to downstream space. With only 128 companies detected in our data, it is necessary to feed this list first to increase the number of potential examples of companies related to space. Building another dataset that covers all the texts mentioning these companies is also crucial. Using a reasonably long list of space-related company names and the press articles in which they appear, it is possible to train a model that can recognize space-related organizations and dispense with the rule-based approach in the future.

A last point in our results caught our attention and could be considered for future improvements. We identified a portion of downstream space companies through co-citation. We have referred to these companies as 'by-products' of the method. It would be interesting to delve deeper into this issue of co-citation to establish links between the identified downstream space companies. We only have a raw list of companies for now, but we lack information on potential supplier-customer or competitor links between the companies. Analyzing the co-citation between the known downstream companies and those identified by the method, or the co-citations between the companies identified only, is a potential path to consider in future methodological developments.

3.8 Appendices

3.8.1 Information on newspaper texts collected

Two sources are over-represented in our text database: 'AFP Infos Economiques' with 5,500 articles and 'AFP Infos Françaises' with around 5,000 articles. These two sources are from Agence France-Presse, a national press agency producing dispatches that are then used by other media to relay the information.

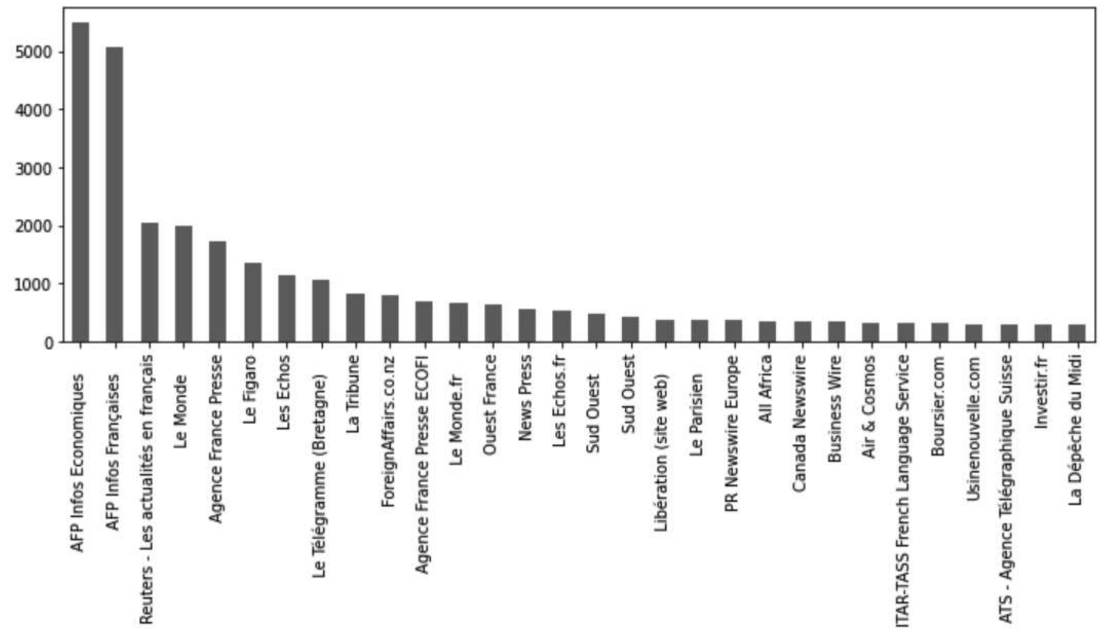


Figure 3.7: Most visible journals, number of articles per source

3.8.2 Descriptive statistics on the number of companies captured by article

The density curve below shows a peak at $n = 10$, corresponding to the number of companies most frequently detected by our method.

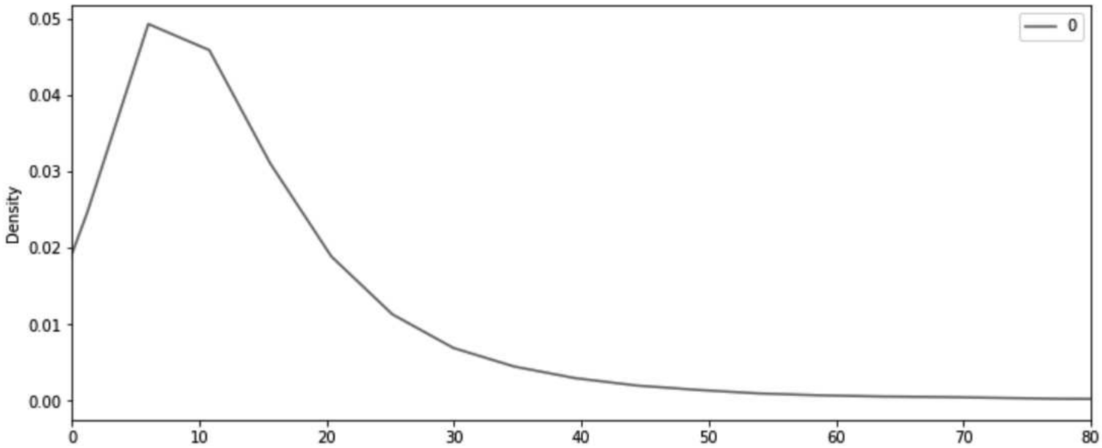


Figure 3.8: Number of companies captured per article

Chapter 4

Measuring the Downstream Space Sector in France in 2021: Evaluation Method and Economic Indicators

4.1 Introduction

In the previous chapter, we introduced our original methodology for identifying companies in the downstream space sector. We focused the study on companies registered in France in 2021. At the end of the first implementation, our instrument enabled us to identify a set of companies engaged in activities that exploit space infrastructures or their outputs to generate added value.

The present chapter is devoted to the economic evaluation of this set of players. Continuing with our empirical approach, we aim to assess the economic footprint of the French downstream space segment in 2021 using a selection of economic indicators. These metrics will provide insights into the industrial structure of the downstream space segment, the diversity of the activities it comprises, and the economic and commercial performance of the downstream space companies. Furthermore, the choice of our evaluation method and associated indicators relies on their capability to be adapted across various geographical areas and replicated at regular time intervals. This adaptability is critical for comparing downstream space markets and monitoring the sector's evolution over time. Importantly, our evaluation aims to contribute to socio-economic impact studies within the space field. A comprehensive understanding of the dynamics of the downstream space sector within the French space industry will be instrumental in formulating appropriate policy recommendations, industrial strategies, and informed investment decisions in the space sector [Dedieu et al., 2016, Ministère de l'Économie et des Finances, 2021].

Our assessment approach involves two stages. First, we conduct a survey among the downstream companies detected through our identification methodology to obtain an in-depth analysis of the sector. This survey provides us with qualitative and quantitative insights into these entities, offering a firsthand perspective on the nature of their activities in the downstream space sector and their commercial performance. This first step provides valuable data for our research, giving us an in-depth and detailed view of the sector.

The second stage encompasses an overall evaluation of the downstream space sector. This assessment integrates some imputation, applying specific survey results to non-respondents. It enables us to provide economic indicators for the downstream space sector as a whole. This evaluation approach not only gives us a more comprehensive perspective on our object of analysis but also fulfills the initial objective of our research, which is to provide economic indicators for the entire downstream space segment.

The remainder of the chapter is structured as follows. We first describe the survey and provide the main findings and economic measures applied to the respondents (Section 2). Subsequently, Section 3 is devoted to generalizing our results and evaluating the downstream space sector as a whole. Here, we detail our approach to obtaining a general measure of the downstream sector and the primary associated indicators.

4.2 Survey of French downstream space companies

The use of a survey as the first step in our assessment appears particularly suitable given the exploratory nature of our research. At this stage of implementing the method, we have identified companies that belong to the downstream space sector in France in 2021. Our identification process suggests that these firms can be heterogeneous regarding objective characteristics (size, economic activity, age) and, importantly, the degree of involvement in the downstream space sector. Conducting the survey allows us to gather comprehensive information on a subset of our downstream companies, to which we do not have access via available company databases, to produce a thorough analysis of the sector. Furthermore, the data provided by the survey is a critical resource in developing our indicators measuring the downstream space sector. This initial survey has enabled us to formulate and refine pertinent indicators based on the collected information.

In this section, we will first delineate the design of the survey and the data collected (2.1). Secondly, we will present the main results of the survey and the principal indicators for the downstream space sector we propose (2.2).

4.2.1 Survey design and data collection

4.2.1.1 Survey Implementation and Respondent Identification

The survey was performed between June 30, 2022, and October 31, 2022, via an online questionnaire.[42] The target population is the 334 downstream companies we identified with our detection tool. The first step was to obtain a contact address for each company in our database. We wanted to address the questionnaire to individuals holding information we needed directly. Therefore, we prioritized obtaining contacts of the CEOs or commercial directors of the surveyed companies. We relied on three primary sources: CNES for well-established downstream companies or those already engaged in a project or dialogue with CNES, the companies' websites where contact information is available, and various online resources (e.g., companies' social media).

We prepared a generic invitation email introducing the survey author and their affiliation, the thesis's overall goal, the survey's specific objective, and the questionnaire's link. We sent the questionnaire once to 285 companies and a second time a week later as a reminder to those who did not respond. In 12 cases where we could not find a specific contact address, we sent the survey invitation via the contact form on the companies websites. Approximately 10% of the downstream space companies in our database, 37 companies, were not contacted due to the lack of contact information.

[42] We used Limesurvey online software to design and administer the questionnaire

4.2.1.2 Structure of the questionnaire

The questionnaire is relatively brief, comprising 17 questions in its most extensive form. The appearance of some questions depends on the answer to the preceding ones. It is structured in two main parts: The first part includes questions common to all respondents. The second part is exclusively for respondents who deliver value-added services dependent on space infrastructures, data, or signals. In this chapter, we only consider the first part of the questionnaire provided in the appendix 4.5.1.[43] In the introductory page of the questionnaire, we present the survey's author and their affiliation, along with the overall objective of the survey. We also assure respondents of the anonymity and non-dissemination of individual responses. We provide the respondents with a precise definition of the survey's scope:

> *'In this survey, the **downstream space segment** encompasses all activities based on **the exploitation of satellite systems or on the production and use of data or signals of space origin to provide added-value products or services** (including user equipment).*
>
> *These activities notably include:*
>
> - ***Sale or lease of satellite capacities**,*
> - ***Production and sale of space data**,*
> - ***Provision of software and/or equipment** for the storage, processing, and use of data, signals, and information provided by satellites,*
> - ***Value-added services** dependent on space infrastructures, data, or signals,*
> - *...*
>
> ***Mission ground segment services**, on the other hand, cover the **activities of satellite operations in orbit** (e.g. tracking, telemetry, and control of satellites, uplink and downlink for signal processing, etc.)'*

[43]Note that the structure outlined here does not appear in the questionnaire sent to the companies. Its sole purpose is to distinguish the survey features that will be covered in this chapter and those that will be discussed in the next.

The questionnaire begins with three *identification questions*. Respondents were required to indicate the name of their company, whether they have activities within the downstream space segment as defined in the introduction, and whether they provide mission ground segment services. The survey ended for those who answered 'No' to both questions.

The following questions relate to the 2021 revenues of the respondents. They correspond to companies' total sales revenues before deducting costs and expenses. We chose this metrics to measure downstream space companies' economic and commercial performance for several reasons. First, revenue is information known to respondents or easily accessible. Second, it is an indicator commonly used to measure the economic value generated in the space sector [OECD, 2022b] and in other sectors in general [EPO and EUIPO, 2022, Majumdar, 1997]. Third, our evaluation aims to be reproduced at regular intervals to analyze the evolution of the downstream space sector. Revenue growth is an indicator that allows an appreciation of a firm's growth [Coad et al., 2016], and ultimately of the sector if we consider all the revenues generated.

Therefore, we asked respondents to report their company's total sales revenues in 2021. Suppose they answered 'Yes' to their involvement in the downstream space segment. In that case, the following question concerned the **share of their total revenue** corresponding to activities in the downstream space segment. Respondents had five possible options: 1. Less than 10%; 2. Between 10% and 50%; 3. Between 50% and 80%; 4. More than 80% or 5. 100%.
We established this categorization to compare downstream actors with one another. We know that companies with downstream space activities belong to diverse sectors, and the extent of their involvement in the downstream segment can significantly differ from one company to another. Hence, we suggest an alternative approach to the primary activity from the standard classifications. This categorization also provides a deeper insight into the structure of the downstream segment and its evolution: whether it consists mainly of companies primarily engaged in downstream space activities (e.g., historical actors of downstream space) and a few companies for whom downstream space activities are peripheral, or if the downstream space segment is predominantly made up of companies for whom space activities constitute a minor portion of their revenue. From a dynamic perspective, monitoring the evolution of the distribution of downstream actors within this categorization is relevant to identifying potential structural changes in the downstream segment.

Similarly, if they answered 'Yes' to the question about their involvement in the ground mission segment, the question about the part of the revenue corresponding to ground mission segment services was posed in the same way. Note here that companies can have activities in both the ground mission and downstream space segments. We have decided not to include mission ground segment services in the downstream segment, mainly because we did not consider this activity when applying our identification method in the first place (Chapter 3). In addition, the ground segment is considered in some cases as belonging to the *intermediate* or *mid-stream segment* between upstream and downstream activities [Australian Space Agency, 2021]. In future applications of the method, it could easily be integrated into the downstream space segment.

Finally, the last set of questions in the first part of the questionnaire focuses on the types of activities the companies are conducting if they answered 'Yes' to the question about their participation in the downstream space segment. For this purpose, we have used the categorization of downstream activities proposed by the OECD in its *Handbook on Measuring the Space Economy* [OECD, 2022b, p.33]. Respondents could select one or more of these options:

- Sale or lease of satellite capacities,
- Production and sale of space data,
- Provision of software and/or equipment for the storage, processing, and use of data, signals, and information provided by satellites,
- Value-added services dependent on space infrastructures, data, or signals,
- Other activity(ies).[44]

If respondents indicated more than one downstream activity, an additional question appeared. We asked them to specify the distribution of their downstream space revenue across these activities. We gave them a scale from 0 to 100 for each downstream activity. Using a slider, they indicated the share of the company's downstream revenue corresponding to this activity. For respondents who responded to only one downstream activity, this naturally corresponded to 100% of their downstream sales.

The questionnaire was tested by experts before being sent to companies. Specialists in the space sector (members of CNES and a former downstream segment actor)

[44]We asked them to specify which ones if they checked 'Other activity(ies)'. See Appendix 4.5.1.

focused on the content and relevance of the survey's scope definition and questions. In addition, we submitted the questionnaire to a researcher experienced in survey techniques to check the structure and coherence of the questionnaire (presentation of the survey, definition of the survey's scope, order and formulation of questions).

4.2.1.3 Additional company data: Amadeus database and FARE file

Moreover, we have coupled the data collected from the survey with additional data about the companies to enrich our analysis. We relied on two main sources. The first is Amadeus, a database proposed by Bureau van Dijk that compiles detailed information and financial data on European companies. The second source of information used is the FARE database (*Fichier Approché des Résultats d'Esane*) for 2019 and 2020. It brings together individual accounting data on French companies. The database combines the administrative data filed by the companies with the Ministry of Finance and the results of the Esane survey (*Elaboration des Statistiques Annuelles d'Entreprises*) produced by the national institute of statistics INSEE. We obtained the authorization to access this protected individual data following a request to the Committee of Statistical Secrecy.
Here are the main company information collected from the FARE and Amadeus databases to develop our survey results:

- **Creation date**.

- **Main activity**, referred to as the APE code (*Primary Activity Performed*) from the Nomenclature of French Activities (NAF).

- **Size category**. INSEE defines four size categories of companies. The first is 'Micro-enterprises,' which employ less than 10 people and have a turnover or total balance sheet that does not exceed 2 million euros. The second category, 'Small and Medium Enterprises' (SMEs), includes entities employing less than 250 people and having a turnover that does not exceed 50 million euros. Then, 'Mid-sized companies' are not SMEs, employ less than 5,000 people, and their turnover does not exceed 1.5 billion euros. Finally, 'Large companies' are those not classified in the previous categories.

We will use these data and other variables from the Amadeus database and FARE file to generalize our results in Section 3.

4.2.2 In-depth analysis and indicators of the downstream space sector: Survey results

4.2.2.1 Presentation of the data collected

We collected 57 responses, representing 20% of the companies that received the questionnaire. Of these, 36 respondents indicated they were active only in the downstream space segment, 15 reported they were involved in both the downstream space segment and the mission ground segment, and 1 respondent said his company was involved only in the mission ground segment. Finally, 5 respondents reported they were involved in neither the downstream space segment nor the mission ground segment.

After verification, three of the five respondents who claimed to be outside the scope of the survey were indeed companies belonging to the upstream part of the space sector. They are either involved in the production and assembly of space equipment and systems or launch services. However, according to our assessment, the remaining two companies are involved in downstream space activities. One provides a weather forecasting service. The other produces and operates autonomous vessels to collect data on the oceans and uses the satellite telecommunications network to transmit marine data to its customers. Given their marginal involvement in downstream activities, we assume that these two companies do not consider themselves to belong to the downstream space segment.

Since this research focuses on assessing the downstream space segment specifically, we will use the term 'respondents' in the remainder of the analysis to refer to companies that have affirmed they have at least one downstream space activity. Therefore, we are excluding the company that stated it only has ground mission segment services, as we did not include the ground mission segment in our scope of assessment for this first application of the method.

Table 4.1 presents descriptive statistics of responding companies compared to non-respondents. The 'Total' column gives the characteristics of all the downstream space companies we have identified (respondents and non-respondents). The first variable considered is the age of the companies. We have introduced two creation periods: before 2010 and after 2010. This distinction allows us to separate downstream companies that came before New Space from the younger companies created in the context of New Space in France. The survey respondents are relatively young

	Respondents	Non-respondents	Total
Creation period			
Before 2010	19 (37%)	132 (47%)	158 (48%)
After 2010	32 (63%)	146 (53%)	171 (52%)
Total	51 (100%)	278 (100%)	329 (100%)
Category			
Micro	30 (59%)	145 (52%)	175 (54%)
SME	14 (27%)	99 (36%)	113 (34%)
Mid-size	6 (12%)	19 (7%)	25 (7%)
Large	1 (2%)	15 (5%)	16 (5%)
Total	51 (100%)	278 (100%)	329 (100%)
No. of different NAF groups (APE)	15	31	33

Table 4.1: Descriptive results - Respondents vs Non-respondents characteristics

companies: 63% were created after 2010 versus 37% before 2010. More precisely, 20 responding companies, or more than a third of respondents, were created after 2016. The second characteristic considered is the company size category. The downstream companies that completed the questionnaire are predominantly small businesses, with nearly 60% of micro-businesses (fewer than 10 employees and less than 2 million euros in turnover) and 27% of SMEs (fewer than 250 employees and less than 50 million euros in turnover). We also have 7 respondents who belong to medium-sized and large companies, representing nearly 20% of the companies of these categories among all our companies in the downstream space segment.

Finally, the last variable used to characterize the companies is the activity sector they belong to, called *NAF groups* in Table 4.1. The Nomenclature of French Activities (NAF) offers different levels of activity classification, with detail up to four figures and a letter. We chose to use the NAF group classification rather than the APE code[45] to simplify the comparison between companies. Using the highest detail level, we would have had to deal with 43 different APE codes. By limiting ourselves to the

[45]The APE code corresponds to the NAF sub-class in the nomenclature.

NAF group, we reduce the number of activity sectors in our dataset of downstream sector companies to 33. For the companies responding to our survey, 15 different NAF groups are represented. It is a convenient result since almost half of the activity sectors of all our downstream companies are present among the respondents, despite a relatively low response rate to the survey.

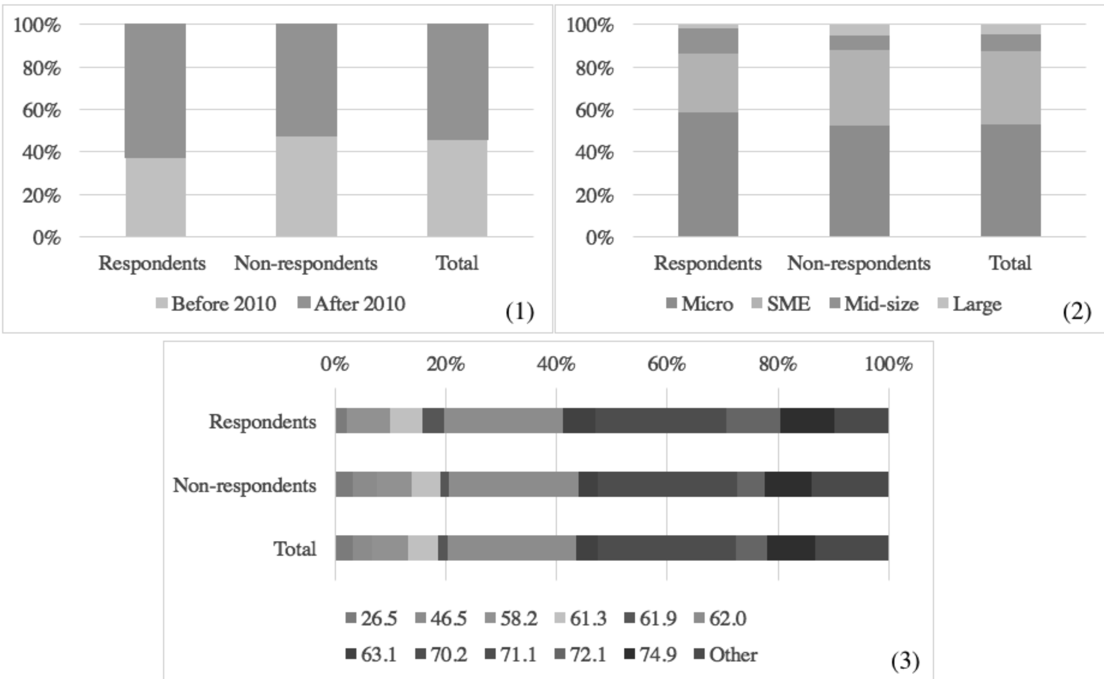


Figure 4.1: Date creation (1), size category (2), and main activity (3) distributions across respondents, non-respondents, and total downstream space companies

Figure 4.1 allows us to compare the distribution of creation periods, business categories and NAF groups among respondents, non-respondents, and all downstream space companies. For creation periods and business categories, we observe a reasonably uniform distribution from one group to another. Among respondents and non-respondents, companies founded after 2010 are more numerous. However, young companies are slightly more present in the respondent group than non-respondents and total companies groups. Micro-enterprises and SMEs are the dominant business sizes in both groups. We notice a small overrepresentation of mid-sized businesses among survey respondents.

Regarding company main activities, two NAF groups are predominant, both among respondents and non-respondents. The first is 62.0, corresponding to 'Computer programming, consultancy, and related services.' 11 out of 51 respondents, or more than 20%, belong to this NAF group. This same NAF group is present at 23% among non-respondents. The second most NAF group represented is 71.1 'Architectural and engineering activities and related technical consultancy.' 20% of respondents and 22% of non-respondents belong to this group of activities. Groups 58.2 'Software publishing services', 61.3 'Satellite telecommunications services', and 74.9 'Other professional, scientific and technical activities' are present in almost identical proportions among respondents and non-respondents. Finally, group 46.5, 'Wholesale trade services of information and communication equipment,' is not represented among the non-respondents, while 12 downstream space companies belong to this group. The 'Other' group contains all NAF groups with fewer than five companies.

We have outlined the main characteristics of survey respondents in terms of age, size and main activity. We will now proceed to present the main economic indicators derived from the responses collected.

4.2.2.2 Key indicators for measuring the downstream space sector

From total revenues to downstream space revenue indicator

We chose sales revenue as the primary performance indicator for estimating the economic footprint of the French downstream space sector. However, as indicated in Chapter 2, one of the methodological challenges in this thesis is to evaluate a heterogeneous set of activities conducted by companies more or less involved in the downstream space sector. In existing evaluations, we have not encountered any such reflection. Either studies are limited to companies whose main activity is downstream, or they consider the evaluation of this segment in terms of economic impact. In our study, downstream activities are not the core business of all companies in the dataset. For instance, some companies are mainly active upstream but have downstream activities as secondary activities. Further down the chain, others deliver services integrating only small quantities of space data. Consequently, the companies' total revenue is not a sufficiently accurate measure of the economic weight of the downstream sector.

To address this issue, we propose an evaluation in terms of **circles of players**, as depicted in Figure 4.2. We distinguish five circles, each representing a group of companies with an identical revenue share attributable to downstream space activities. The closer the circle is to the black 'core', the greater the share of revenues corresponding to downstream space activities. Therefore, the innermost circle in black includes companies whose entire revenue is attributable to downstream activities, as defined in the survey. In dark orange, the second circle comprises downstream space companies for which between 80% and 100% of revenue corresponds to these activities. The third circle includes companies whose downstream space activities are still an important source of value (between 50% and 80% of their revenues) but have diversified activities. The fourth circle consists of players whose share of downstream space in total revenue ranges between 10% and 50%. Finally, the outermost circle of players corresponds to companies with a small revenue share attributable to downstream space (less than 10%).

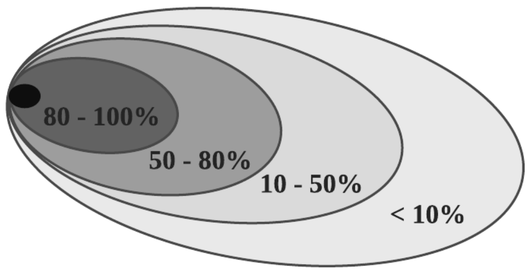


Figure 4.2: Circles of players by share of downstream space revenue.

With this representation, we avoid the problems inherent in the core business approach. Belonging to a remote circle of players, for instance, between 10% and 50%, does not mean that downstream activities are not the company's primary source of added value. Indeed, it could be that 20% of a company's total sales are associated with the provision of satellite services, and this constitutes its main source of value.

Based on the circle categorization, we introduce a **downstream space revenue indicator**, representing the revenues from companies' downstream space activities. Specifically, we calculate the downstream revenue of each surveyed company by multiplying its total sales by the share of sales attributed to downstream space activities. As we deal with ranges, we use the lower limit of the range to calculate a **minimum downstream revenue** and the upper limit to calculate a **maximum downstream revenue**. Thus, we propose three revenue measures for each company:

its total revenue in 2021, its minimum downstream space revenue in 2021, and its maximum downstream space revenue in 2021. To obtain metrics for the entire downstream space sector, we sum up the companies' total, minimum downstream, and maximum downstream revenues.

Share of Revenue from Downstream Space Activities*	N	Total Revenue**	Min. Downstream Revenue	Max. Downstream Revenue
Less than 10%	8	147,592,983	1,475,930	13,283,368
10% - 50%	16	627,428,500	62,742,850	307,439,965
50% - 80%	6	2,650,000	1,325,000	2,093,500
More than 80%	7	155,677,453	124,541,962	154,120,678
100%	14	1,262,324,294	1,262,324,294	1,262,324,294
Total	51	2,195,673,230	1,452,410,036	1,739,261,805

* Revenue shares were converted into the following intervals for downstream revenue computation: [0, 10), [10, 50), [50, 80), [80, 100), and 100%.

** Two respondents did not disclose their 2021 revenue.

Table 4.2: Respondents Downstream Space Revenue in 2021 (in euros)

Table 4.2 details the 2021 revenues from the respondent companies we obtain. The total revenue of the 51 companies surveyed amounts to nearly 2.2 billion euros in 2021. We estimate the minimum revenue imputable to downstream space activities at 1.45 billion euros, representing 65% of the total revenue. Lastly, we estimate the maximum downstream revenue to be 1.74 billion euros, or 80% of the total revenue of the respondents. The two measures of minimum and maximum revenues, with a difference of 300 million euros between the two, allow us to have both a conservative and a more expansive estimate of the revenue generated by the downstream activities of the respondents. The most notable difference between the two measures is found in the '10%-50%' circle, within which the maximum downstream revenue is five times greater than the minimum downstream revenue, increasing from 62.7 million to 307.4 million euros.

Besides, the results suggest no significant difference between the total revenue of the companies and the downstream space revenue. One reason may be that our sample of respondents is relatively small compared to the entire downstream space segment. An analysis of downstream revenues within each circle of players is required to understand this result.

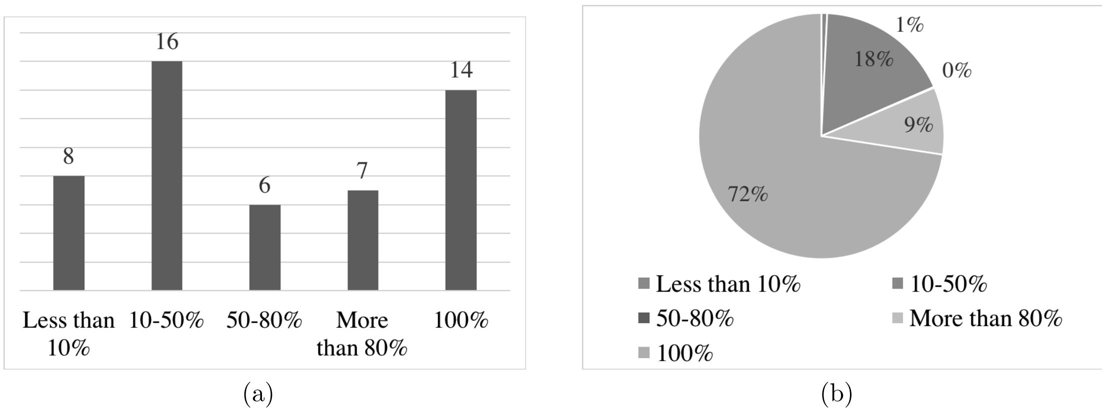


Figure 4.3: Number of respondents by circle of actors (a); Distribution of maximum downstream revenues by actor circle (b)

We observe that two particular revenue ranges contain most companies: '10%-50%' and '100%' (Figure 4.3 (a)). The former includes 16 companies, while the latter includes 14 out of the 51 downstream companies surveyed. When we consider the maximum downstream revenue, companies attributing all their revenue to downstream activities emerge as the key contributors. This group accounts for 72% of the total maximum downstream revenue. The '10%-50%' range follows next, contributing 18% to the total maximum downstream revenue. The companies within the 'More than 80%' interval emerge as the third-largest contributors to the total maximum revenue, even though they are only 7. Finally, the 'Less than 10%' and '50-80%' circles, comprising only 8 and 6 respondents respectively, contribute marginally to the total downstream revenue.

We provide the table of results for the revenues generated by the respondents' mission ground segment activities in the appendix 4.5.2.

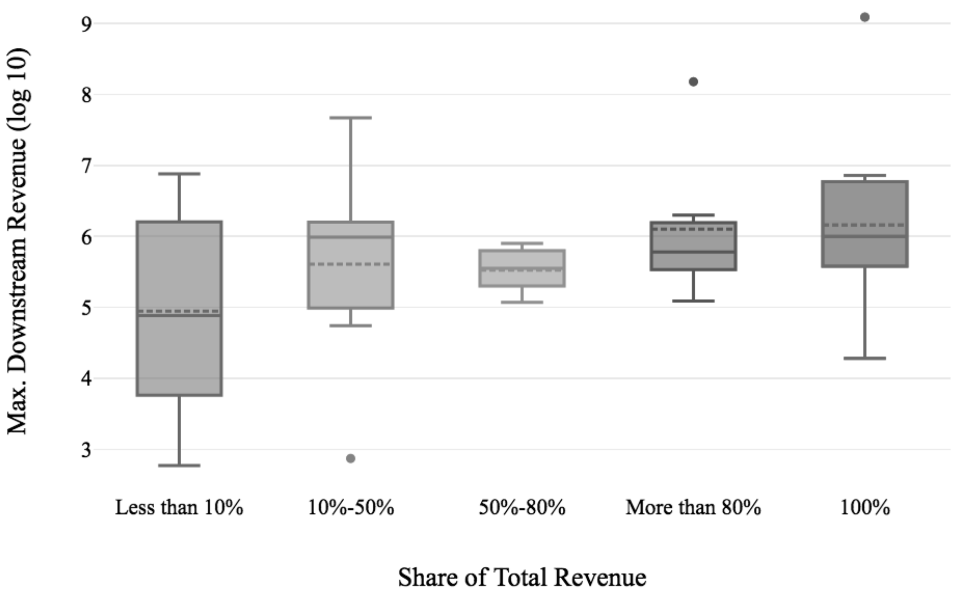


Figure 4.4: Distribution of Downstream Revenues Across Circles of Players

The boxplot in Figure 4.4 compares the distribution and variability of maximum downstream revenue among different actor circles (or revenue range intervals). Each box corresponds to a circle of downstream players and depicts the interquartile revenue range. The whiskers extending from each box illustrate the range of the data. We observe that the downstream revenues are most dispersed in the 'less than 10%' circle. Regarding the '10%-50%' circle, the gap between the median (solid line) and the mean (dotted line) indicates an asymmetry in the downstream revenue distribution. The blue dot beneath the box reveals the presence of an outlier, namely a downstream revenue value considerably lower than the others, pulling the average downwards. The '50%-80%' circle (represented by the orange box in Figure 4.4) exhibits the least variation in downstream revenues, with the mean and median values coinciding. However, as this interval has the fewest respondents, the reliability of this result may be limited. Finally, we note the presence of two outliers within the 'More than 80%' and '100%' circles. We have shown above that the latter circle is the primary contributor to the total downstream revenue from respondents. The fact that it contains an outlier, i.e., a downstream revenue substantially higher than those within its circle, demonstrates that a single company contributes significantly to the total maximum downstream revenue of the respondent companies. In other words, a single company belonging to the '100%' circle accounts for a substantial portion of the total estimated downstream revenue based on the survey data.

Analysis of downstream space activities declared

We now focus on the survey results related to the downstream activities conducted by the companies. The respondents defined their activities from five options, corresponding to the different stages of the downstream space value chain. Three main aspects guide our analysis of the collected information on downstream activities.

First, we aim to build on our downstream revenue indicator by supplementing it with a measure of downstream revenue per activity. This indicator helps determine the contribution of each activity segment in the downstream chain to the overall downstream space segment.The second aspect involves a deeper exploration of our analysis by actor circles. We seek to know what activities are conducted within the same circle and whether companies of the same circle have similar downstream space activities. Finally, the last aspect of our activity-based analysis relates to the diversification of downstream companies. To evaluate the state of a sector and track its evolution, it seems essential to know whether or not the companies have more than one activity within that sector. As a reminder, the analysis relies on data collected through a survey from a limited sample of downstream companies (51 out of 329). Therefore, the results presented should be interpreted with caution.

Downstream 2021 revenues by activity

From the information provided by respondents on the portion of their downstream revenue corresponding to each activity, we have devised a measure of **downstream revenue by activity**. Table 4.3 details the revenue by segment of the downstream value chain based on survey responses. The 'N' column in the table refers to the number of companies that reported being involved in the respective activity. A company may have multiple downstream activities.

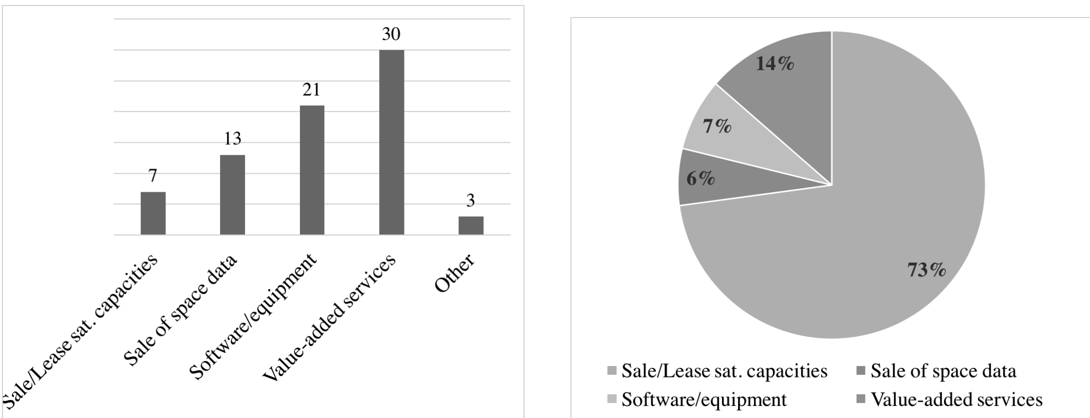
Downstream activity*	N	Min. Downstream Revenue	Max. Downstream Revenue
Sale or lease of satellite capacities	7	1,203,980,000	1,267,230,000
Production and sale of space data	13	45,195,289	103,651,599
Provision of software and/or equipment	21	48,216,782	131,948,471
Value-added services	30	154,923,606	236,336,858
Other	3	94,359	94,879
Total		1,452,410,036	1,739,261,805

* Four respondents with multiple downstream activities did not specify their revenue distribution per activity. We applied an equal revenue split among their activities by default.

Table 4.3: Respondents 2021 Downstream Space Revenue by Activity (in euros)

For the calculation of downstream revenue by activity, we followed the same procedure as for total space downstream revenue. We multiplied the minimum and maximum downstream revenue for each company by the revenue share it attributes to each activity. Take a company with total revenue of x euros. The company states that between 10 and 50% of its total revenue corresponds to downstream activities and that 80% comes from the sale of space data, the rest from the provision of value-added services. To calculate the minimum (respectively maximum) downstream revenue related to the sale of space data, we multiply x by 10% (respectively 50%) and then by 80% (share of revenues linked to the sale of data). We proceed similarly for the revenue associated with providing value-added services, replacing 80% with 20%. To obtain the total revenue by activity, we calculate the sum of the declared downstream activity revenues.

Thus, the *minimum* downstream revenue corresponding to selling and leasing satellite capacities among respondents amounts to 1.2 billion euros. The *maximum* downstream revenue related to this segment is slightly higher (1.27 billion euros). As for space data sales and software and equipment provision segments, there is a notable difference between *minimum* and *maximum* revenues. Considering the *minimum* revenues attributed to these activities, data sales generate 45 million euros in revenue and software and equipment provision 48 million. However, if we take into account the *maximum* downstream revenue, these activities generate respectively 104 million euros and 132 million euros in revenue. The value-added services segment, which 30 respondents declared being part of, represents 155 million euros in *minimum* revenues and 236 million in *maximum* revenues. Finally, three companies reported having 'other' activities generating 94,000 euros in revenue. The 'other' activities indicated are as follows: research and development, environmental evaluation, and the use of satellite images for ground data verification.



(a) (b)

Figure 4.5: Number of respondents by downstream activity (a); Distribution of maximum downstream revenues by downstream activity (b)

For a better visualization of the contribution of each activity segment to the downstream revenue, Figure 4.5 presents the number of respondents per activity (a) and the distribution of the respondents' maximum downstream revenues by activity (b). Even though the sale and lease of capacities segment includes the fewest respondents (besides 'other'), it is the largest contributor to the surveyed companies' downstream revenues. Indeed, this segment accounts for 73% of total downstream revenues. Meanwhile, the value-added services segment, comprising the highest number of respondents (30), adds 14% to the total downstream revenue. Finally, the segments of image sales and software and equipment provision, with 13 and 21 respondents respectively, contribute 6% and 7% to the total downstream revenue. Given the low number of companies that have declared having 'other' activities than the four proposed, this segment contributes marginally to the total downstream revenue. Therefore, it does not appear on Graph 4.5(b).

Analysis of downstream activities by circle of actors

We continue our analysis of the results by activity, taking a downstream actor circle perspective. Figure 4.6 provides us with information on the downstream activities declared by each actor circle. We note that the data sales, software and equipment supply, and value-added services segments are present in all actor circles. The first top segment of the value chain (sale and lease of capacities) is represented in three circles: two respondents from the '10-50%' circle, two from the 'More than 80%' circle, and three from the '100%' circle. No respondent from the most distant circle, 'Less than 10%', participates in satellite capacity sale and lease activity. However, over 20% of companies from the 'More than 80%' and '100%' circles indicated their participation in this activity.

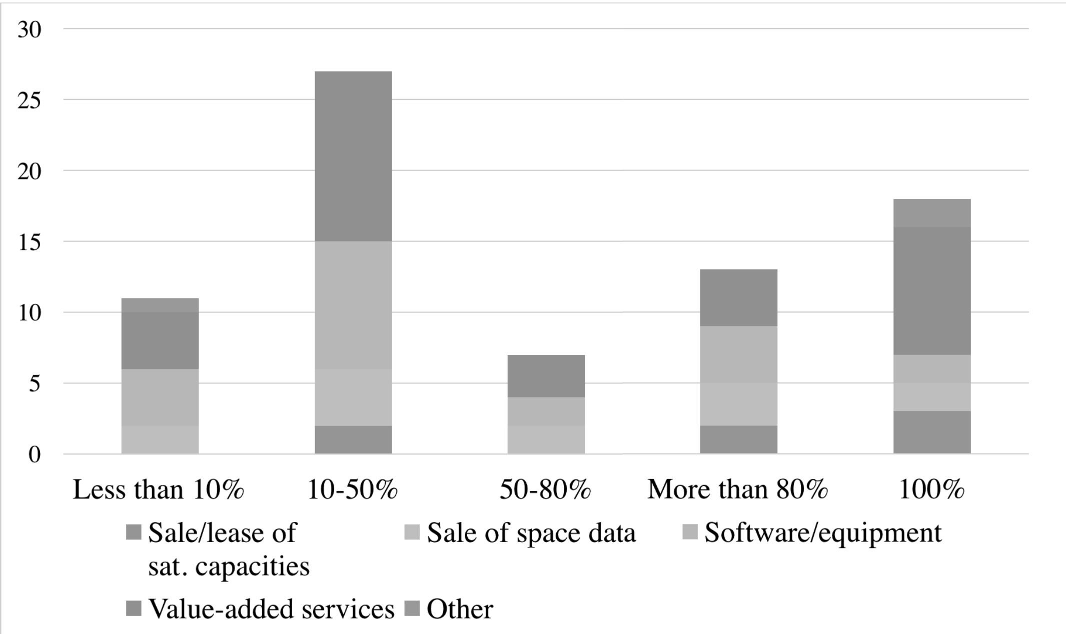


Figure 4.6: Type of downstream activities conducted by circle of actors

Furthermore, the two companies from the '10-50%' circle involved in capacity sale and lease activity declared that the other half of their total revenue corresponds to mission ground segment services. Drawing conclusions is challenging due to the low number of respondents per circle. However, we could presume that satellite capacity sales and rentals are mainly carried out by companies from circles with the most significant share of downstream revenue.

We now focus on the number of downstream activities in each actor circle (Figure 4.7). The first observation is that respondents primarily engage in a single downstream space activity in all circles. Furthermore, each actor circle includes at least one company involved in two downstream space activities. For example, in the 'Less than 10%' circle, three companies for which downstream activities generate less than 10% of their total revenue report have two downstream activities. One company has declared involvement in sales and rentals of capacities and satellite imagery sales—the second deals with satellite image sales and software and equipment supply. Finally, the last company having only one downstream activity (provision of software and equipment), also declares activities in the mission ground segment, which we did not define as a downstream activity.

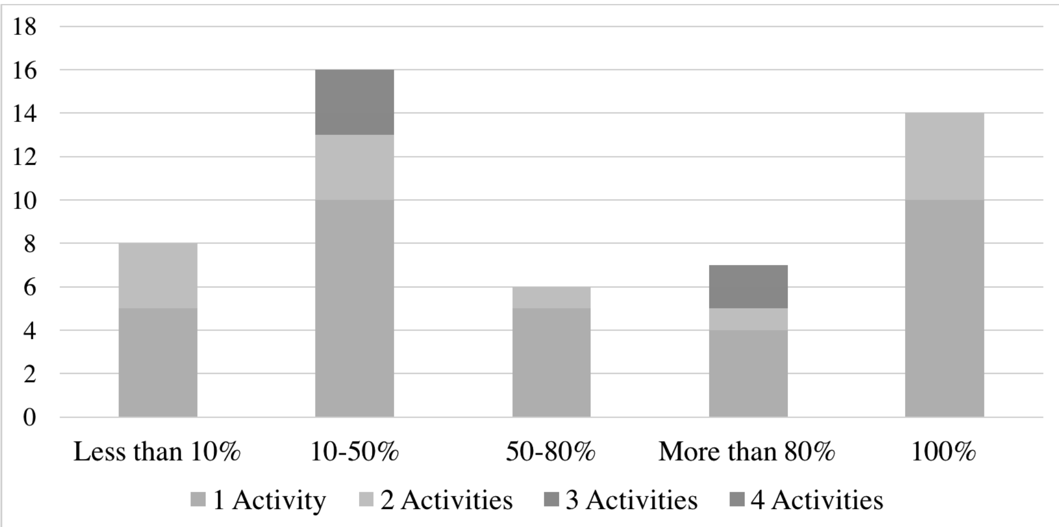
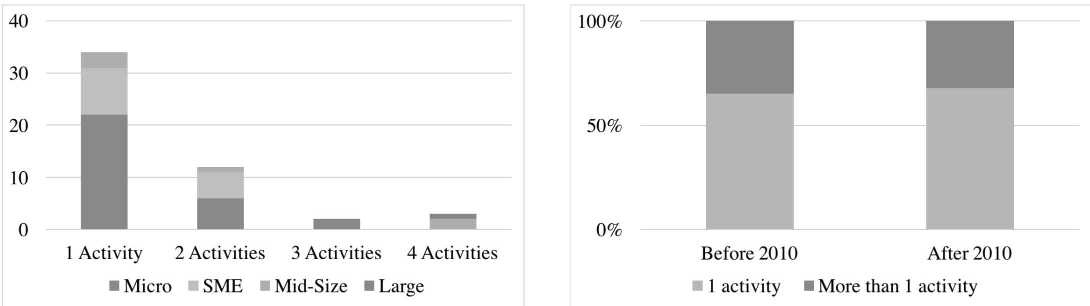


Figure 4.7: Number of downstream activities conducted by circle of actors

Interestingly, the respondents whose revenue is 100% attributable to downstream activities are not those with the most diversified downstream operations. Three companies in the '10-50%' circle report having more than two downstream activities, and two also have mission ground segment activities. Furthermore, two companies from the 'More than 80%' circle indicate more than two downstream activities, with one also involved in mission ground segment activities. Therefore, the survey results make us question the exclusion of the mission ground segment from the downstream space sector. If we had included it as a downstream activity, the different circles of actors might be more homogeneous regarding respondent characteristics.

Analysis of respondents' diversification into downstream activity segments

The last element of the survey data that we analyze here is the degree of diversification in the downstream activities of the respondents. This indicator particularly interests us, as it provides insights into the structure of the downstream value chain. One-third of the survey respondents report engaging in more than one activity from the downstream value chain. The idea for future evaluations is to analyze this figure's evolution with details of diversified companies' characteristics.



(a) (b)

Figure 4.8: Number of downstream activities of respondents, by category (a) and creation period (b)

Figure 4.8 presents the number of activities conducted by surveyed companies based on their size category (a) and the period of their creation (b). Companies with only one downstream activity are predominantly small-sized: 65% are micro-enterprises, and 24% are SMEs. Three companies, making up 8% of all companies with only one downstream activity, are mid-size. This distribution of company sizes is almost equivalent among companies with two downstream space activities. Moreover, the two companies with three downstream space activities are micro-enterprises. On the other hand, we note that companies involved in all four downstream chain segments are large companies (two intermediate-sized companies and one large company). Lastly, when looking at the creation period (Figure 4.8 (b)), we observe that there is a similar proportion of young companies (established after 2010) and older companies (created before 2010) that undertake more than one downstream activity.

Finally, the last measure of diversification we propose is the number of activities per activity segment (Figure 4.9). An interesting result emerges: most of the surveyed companies that sell and lease satellite capacities and those that sell space data have at least one additional downstream space activity. It means that companies at the beginning of the value chain are likely to be involved in multiple downstream segments. Subsequently, companies with a single downstream activity and multiple downstream space activities are equally proportioned in the 'software and equipment' and 'value-added services' segments.

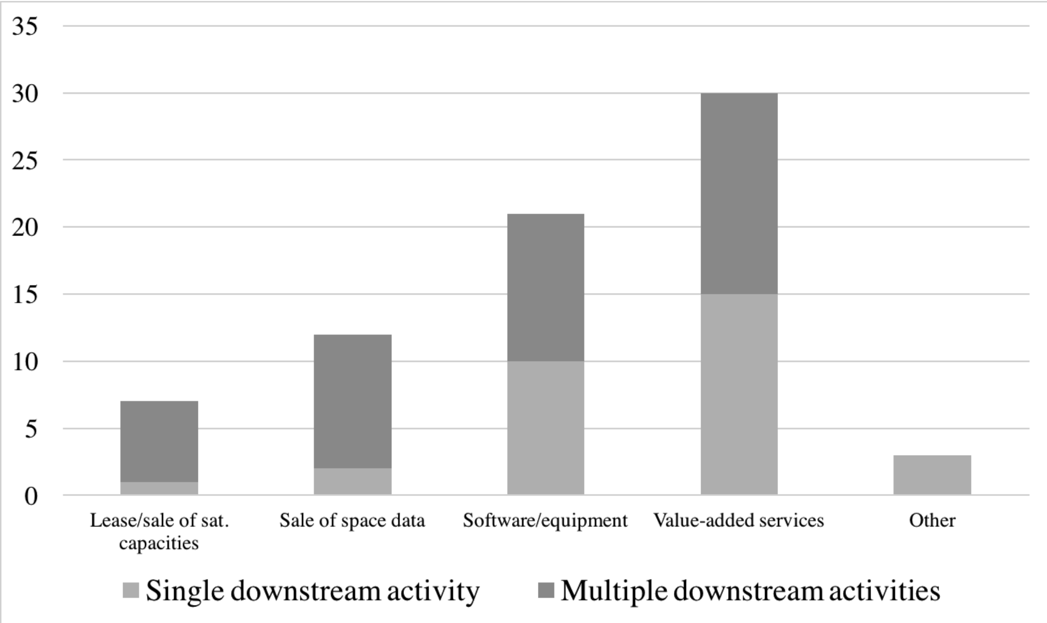


Figure 4.9: Diversification of respondents by activity type

4.3 Generalization of the Results and Assessment of Revenues Generated by the Downstream Space Sector in France in 2021

In this section, we aim to generalize the survey findings and provide an assessment of the revenues generated by the entire French downstream space sector in 2021. Specifically, this evaluation stage involves estimating the downstream space revenue of companies that do not participate in the survey based on common characteristics with the survey participants. We therefore focus on our primary indicator, the downstream space revenue, in the remainder of the assessment.

First, we present imputation methods used to generalize our findings to the broader sector. Second, we present the results and introduce the key economic indicators for the entire downstream space sector.

4.3.1 Imputation methods

We developed two imputation procedures for missing data. The first aims to estimate the 2021 sales revenues of downstream companies for which information is unavailable. Once we have the known and estimated 2021 revenue figures for all companies, the second procedure involves estimating which circle of actors non-responding companies belong to. That is, determining what share of their total revenue corresponds to downstream space activities. We detail the methodologies used for each imputation in the following sections.

4.3.1.1 Group-based median imputation method for 2021 missing revenue estimation

The procedure aims to estimate missing downstream companies' 2021 total revenues. We started by collecting available data on the total 2021 revenue for the 278 downstream space companies that did not respond to our survey. To do this, we used the Amadeus database[46] to compile missing revenue data. In total, we obtained the 2021 revenues for 152 downstream companies through our survey and using the Amadeus database.[47] This left us without 2021 revenue information for an additional 177 downstream companies.

We implemented a three-step imputation method to estimate the missing 2021 revenues. First, we divided our downstream companies into groups using a three-dimensional matrix: the size category of the company, the NAF group, and whether or not the 2021 total revenue is known. We simplified the 'company size' variable, which takes on two values: 'Small' for micro-enterprises and SMEs and 'Large' comprising mid-sized and large companies. This categorization increases the chances of having known 2021 revenue data in each group. The second variable, 'NAF group,' previously used in our survey results, is a three-digit code representing the company's industry

[46]See Section 4.2.1.3 for details on the Amadeus database.

[47]We could not collect any information for the year 2021 in the FARE files (a financial database on French companies presented in Section 4.2.1.3) since the most recent data available was from 2020.

sector per the French Activity Nomenclature. Finally, the third variable indicates whether or not the 2021 revenue data is known. At the end of this stage, companies sharing the same size category and the same sector of activity are grouped together.

The second step of the imputation procedure involves calculating the revenue growth rate between 2019 and 2021 for each group of companies with known 2021 revenues. We selected the base year as $n - 2$ instead of $n - 1$ because the 2020 pandemic shock could have significantly impacted the downstream space sector's revenues. Therefore, we collected the 2019 revenue data from the FARE 2019 file and supplemented it with the 2020 revenue data from FARE 2020 for the ten downstream space companies in our database established in 2020. After calculating the sales growth rate for each company with known 2021 sales, we calculate the median growth rate for each group.

The final step entails imputing each group's median revenue growth rate to the companies within the same group that do not have known 2021 revenues. This method is based on the strong assumption that companies follow a similar growth pattern to their peers. We used the median rather than the average for our estimates to mitigate the influence of outliers. Indeed, we encountered cases where groups had average growth rates of four or even five digits, resulting from a company with extremely low or even zero revenue in 2019 but with several hundred thousand euros in 2021. The median is more suitable than the mean as we are dealing with groups with a small number of companies and extreme values.

To illustrate our methodology, let's consider a medium-sized company operating in the '61.3: Satellite Telecommunications Activities' sector, for which we wish to estimate the 2021 revenue. We have access to its 2019 revenue data. Within the same group as this company (i.e., SMEs within the NAF 61.3 group), we possess 2019 and 2021 revenue data for 30 companies. We calculate the revenue growth rate for each company between 2019 and 2021 and subsequently derive the median growth rate within the group. We finally apply this median growth rate to the 2019 revenue baseline of our target company to arrive at an estimated revenue for 2021.

Finally, 18 downstream companies belonged to a group where no 2021 revenue was known. Seven had the same NAF group as companies with known 2021 revenue but did not belong to the same size category. To address this issue, we prioritized the size criterion. We applied the median growth rate of all downstream companies of the same size for which 2021 revenues are known across all industry sectors.

The remaining 11 companies did not share their NAF group with any company with known 2021 revenues. Given their small number, we examined the activities conducted by each of these companies and found a 'twin' company within a group with known 2021 revenues. We then attributed the median growth rate corresponding to the group of their 'twin' company. For instance, we had one small company with unknown revenue information in the NAF group '18.1: Printing and service activities related to printing.' We learned after research that this was a mapping company. We found a 'twin' company in the '58.2: Software publishing' group, falling under the 'Small' category and providing mapping solutions. We then assigned the median growth rate of the '$Small \cap 58.2$' group to the initial company. We repeated this process for all companies in this situation.

4.3.1.2 Downstream revenue share imputation using K-nearest neighbors method

The objective of the second imputation is to quantify the share of downstream revenue for companies that did not respond to the survey. Essentially, the idea is to assign a 2021 revenue range corresponding to downstream space activities for each company with missing information. In other words, we attempted to determine the circles to which the non-responding downstream companies belong: less than 10%, between 10% and 50%, between 50% and 80%, over than 80% or 100%. The challenge here arises from the small proportion of respondents relative to non-respondents. Indeed, we have data on the downstream revenue share for 51 companies, whereas this share needs to be determined for 278 companies.

We employed the k-nearest neighbors (KNN) method for imputing the share of downstream revenues of the non-responding companies. KNN is a machine learning algorithm typically used for classification tasks. It is a non-parametric technique [Ehsani and Drabløs, 2020], meaning it doesn't require us to make assumptions about the nature of the relationship between the variables, unlike linear regression methods, for example. The basic principle of a KNN model is to determine the value of a missing observation based on the k closest known observations. The proximity, or similarity, between two observations, is measured in a multi-dimensional space where each dimension corresponds to a characteristic of the observations. The value of the missing observation is determined by its k-nearest neighbor, i.e., the shortest distance to a known observation.

We chose the KNN imputation method for our study because it offers several advantages. First, in addition to being easy to implement [IBM, 2023], the KNN algorithm is a powerful and precise prediction tool compared to other existing methods (e.g., mean and median imputation, linear regression, iterative model) [Jadhav et al., 2019, Kenyhercz and Passalacqua, 2016]. Second, it can deal with both continuous and categorical variables [Song et al., 2008]. In our analysis, we are only using categorical variables.

The performance of a KNN model relies on two aspects. First is the selection of the optimal distance measure [Ehsani and Drabløs, 2020], and the second is the choice of k, which refers to the number of neighbors used to determine the value or class of the missing observation [IBM, 2023]. The choice of the optimal distance measure depends on the type of variables used. In our imputation model, we selected the same variables as we did for our initial 2021 revenue estimation procedure: the company size category and the NAF group. Our two qualitative variables thus define the dimensions of our space. The size category variable takes on values 'Micro,' 'SME,' 'Mid-size,' and 'Large,' classifying it as an ordinal or hierarchical qualitative variable. The second variable of the model, the NAF group, is of the nominal categorical type. Its values are numerical but unordered.

We used the **R** package *VIM* developed by Kowarik and Templ [2016] to implement our KNN model. This package is particularly suited to handle categorical variables. The distance measure used to determine the k-nearest neighbors is called *Gower's distance*. Formally, we compute Gower's distance between an observation i and an observation j as follows:

$$d_{i,j} = \frac{\sum_{k=1}^{p} w_k \delta_{i,j,k}}{\sum_{k=1}^{p} w_k} \tag{4.1}$$

where w_k is the weight assigned to the difference for variable k, and $\delta_{i,j,k}$ is the distance between individuals i and j for variable k. In cases where variables are categorical, we calculate the distance with a binary measure, assigning 0 if the two observations have the same value for variable k, 1 otherwise:

$$\delta_{i,j,k} = \begin{cases} 0, & \text{if } x_{i,k} = x_{j,k}, \\ 1, & \text{if } x_{i,k} \neq x_{j,k}. \end{cases} \tag{4.2}$$

In other words, each qualitative variable is transformed into a matrix form and normalized on a scale from 0 to 1 to compute Gower's distance. Subsequently, the average relative distances for each variable are calculated [Kowarik and Templ, 2016].

The second performance criterion for our KNN model, used to impute the downstream revenue share to companies that did not participate in the survey, is the choice of the value of k. This value refers to the number of neighbors considered when calculating the distance between the missing and non-missing observations. To determine this value, we implemented a commonly used technique called cross-validation [Kramer and Kramer, 2013]. This approach involves dividing our company dataset into subsets called 'folds,' which serve as training and testing sets. We defined ten folds, so the algorithm was trained and tested ten times, each using a different fold as the test set and the other nine combined as the training set. For each fold configuration, the imputation model was trained on the training set and tested on the test set. An error measure for the model was calculated for each configuration. At the end of the process, we selected the value of k that produced the lowest average error across the test folds. We tested the imputation model for k ranging from 1 to 30. The lowest error was for $k = 5$.

Therefore, we applied this number of neighbors for the distance calculation. The model accuracy graph is in the appendix 4.5.3. We observe that the accuracy curve is the highest (lowest average error) for neighbors equal to 5. Finally, the model allows us to obtain a table with imputation of the circle of downstream actors (downstream share revenue range) for each observation, i.e., each company with missing information.

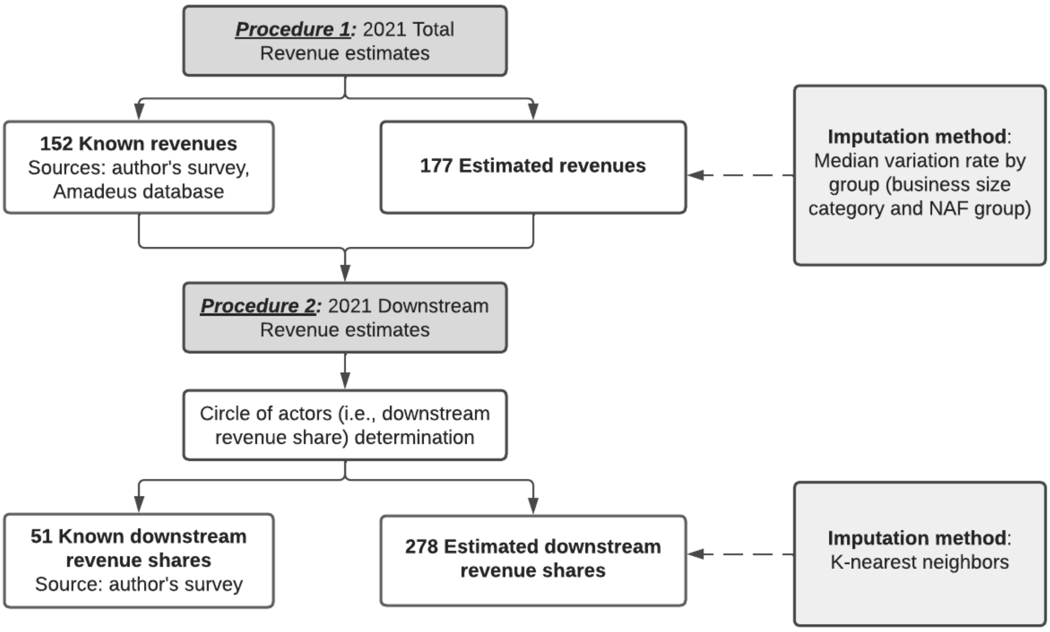


Figure 4.10: Methods of imputing missing revenues - Summary

In this section, we described the two imputation procedures we implemented to generalize our measures of the downstream space segment. Figure 4.10 summarizes the goals and methods used. Our focus was on measuring the downstream revenue in 2021. First, we estimated the missing total revenues in 2021 by applying the median growth rate of companies from the same group for which the information is known (Procedure 1 in Figure 4.10). Second, we implemented a KNN model to impute downstream revenue shares for companies that did not respond to the survey (Procedure 2 in Figure 4.10).

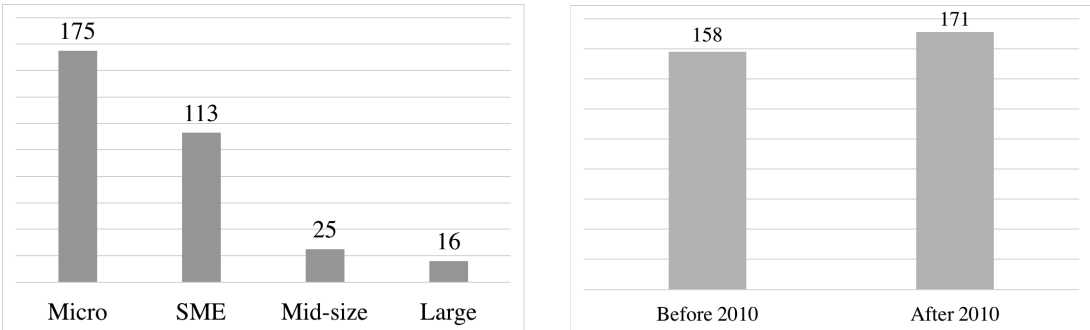
The following section presents the imputation procedures results and the primary indicators for the entire downstream space segment in France in 2021.

4.3.2 Imputation results and general assessment of the French downstream space sector in 2021

This final part presents the general findings of our evaluation of the downstream space sector in France in 2021. We start by providing information on the general characteristics of the 329 companies that make up the downstream sector. Subsequently, we present the estimation results of the downstream space revenue derived from our implemented imputations.

4.3.2.1 General characteristics of French downstream space companies

Our evaluation results start with examining economic indicators that allow us to understand the structure of the downstream space sector. Figure 4.11 describes the distribution of all downstream space companies by size category (a) and by period of creation (b). The downstream space sector comprises mainly small companies: 175 micro-enterprises (less than ten employees and less than 2 million euros in turnover) and 113 SMEs (between 10 and 250 employees and less than 50 million euros in turnover). In addition, 25 companies identified as belonging to the downstream space sector are of intermediate size, i.e., they employ between 250 and 4,999 people and generate less than 1.5 billion euros in turnover. Finally, 16 companies are large enterprises, representing 5% of downstream space companies.



(a) (b)

Figure 4.11: Number of downstream companies by size category (a); Number of downstream companies by creation period (b)

The second indicator (4.11 (b)) refers to the period of creation of the downstream companies. As we explained at the beginning of this chapter, we made this categorization to distinguish downstream companies established before the New Space period in France from those created during and after the New Space wave. The distribution of downstream companies between the two periods is almost equal, with 48% of companies created before 2010 and 52% after 2010.

The size and creation date indicators are related, and further analysis shows that three-quarters of the micro-enterprises were founded after 2010. Conversely, 93% of large downstream space companies (mid-size and large) were established before

2010.[48]

Finally, Table 4.4 summarizes the NAF groups held by at least ten downstream space companies. This information, which refers to the sector of activity of the companies defined in the official activity classification, is essential because we use it at several stages of our identification and evaluation method. For the evaluation, we use it as a variable to perform imputations of revenue growth rates and determine the circle of actors (downstream revenue share) to which companies that did not participate in the survey belong.

Code	**Description**	**N**
26.5	Manufacture of instruments and appliances for measuring, testing and navigation; watches and clocks	10
46.5	Wholesale of information and communication equipment	12
58.2	Software publishing	21
61.3	Satellite telecommunications activities	18
62.0	Computer programming, consultancy and related activities	76
63.1	Data processing, hosting and related activities; web portals	13
70.2	Management consultancy activities	10
71.1	Architectural and engineering activities and related technical consultancy	72
72.1	Research and experimental development on natural sciences and engineering	19
74.9	Other professional, scientific and technical activities n.e.c.	28
Other	-	50
Total		329

Table 4.4: Main NAF groups of downstream space companies

Two sectors stand out from the others regarding representation among downstream companies: 'Computer programming, consultancy and related activities (62.0)' with 76 companies, and 'Architectural and engineering activities and related technical consultancy (71.1)' with 72 companies. When reading the descriptions of these two groups of activities, it is clear that it would have been challenging to affiliate them with activities related to the space sector. Moreover, without our identification and evaluation tool, it would have been impossible to distinguish downstream companies

[48]We did not provide a graph to visualize this information to avoid overloading our study.

from the others within these NAF groups. The only NAF group that explicitly refers to downstream activities is: 'satellite telecommunications activities (61.3)' associated with 18 companies in our database.

We now turn our attention to reviewing the imputation results and presenting the estimation of the downstream space revenue, which is the central indicator of our analysis.

4.3.2.2 Estimated 2021 downstream space revenue in France

Summary of imputation results

The imputation results of downstream revenue shares using k-nearest neighbors are summarized in Figures 4.12 and 4.13. As a reminder, the imputation procedure we used is as follows: for each company we did not have information about, we assigned the share of downstream revenue from the 'known' company that is closest among its five nearest neighbors. In other words, we used the data from the least distant company to impute the share of downstream revenue for companies for which this information was missing. The distance measure relies on two qualitative variables: the NAF group, corresponding to the company's sector of activity defined by the official classification, and the company size category (micro, SME, mid-size, large).

The graph in Figure 4.12 depicts the percentage of companies in each circle of actors, i.e., companies with the same share of downstream revenue. It compares the 'Known' group for which we have this information – the 51 companies that responded to the survey – and the 'Imputed' group for which we implemented KNN to impute this information – the remaining 278 downstream companies. While the **uniformity of distribution between the known and imputed groups does not indicate the model's performance**, it was informative to make this comparison. In general, the distribution of the number of companies per circle is quite similar between the group of known companies and the group for which KNN has imputed this information. Nevertheless, we note that a more significant proportion of companies in the 'Imputed' group belong to the '10-50%' circle than the 'Known' group. Additionally, the share of companies in the '50-80%' circle is smaller (around 3%) in the 'Imputed' group compared to the 'Known' group (11%).

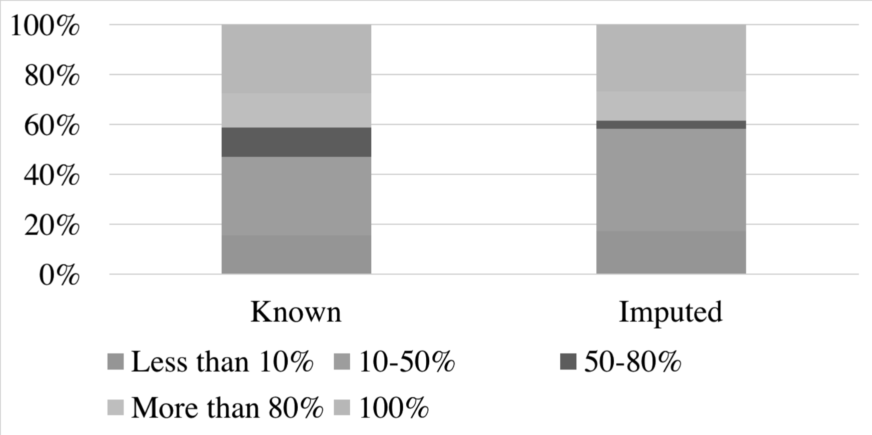


Figure 4.12: KNN Imputation results - Comparison of downstream revenue shares: Known vs. Imputed

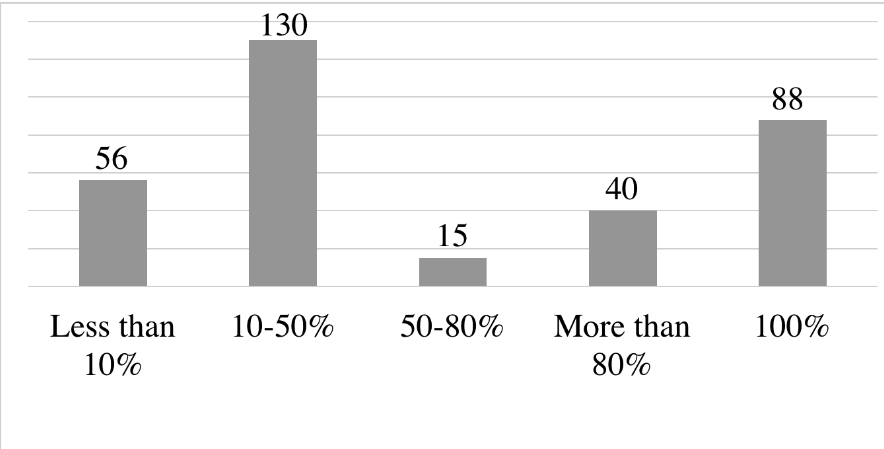


Figure 4.13: Total number of companies by circle of downstream actors (downstream revenue share)

Figure 4.13 provides information on the total number of downstream companies per circle. Each circle thus includes companies for which the share of downstream revenue is known and companies for which this share has been imputed using KNN ('Known' and 'Imputed'). The circle with the highest number of downstream companies is '10-50%' with 130 downstream companies. Following this is the '100%' circle with 88 companies. In other words, for 88 companies, over a quarter of the companies in our downstream database, revenues generated are attributable solely to downstream space activities. Next, we estimated that 56 companies belong to the 'Less than 10%'

circle and 40 to the 'More than 80%' circle. The '50-80%' circle contains the fewest downstream companies, with only 15 entities falling within this category.

Measure of total 2021 downstream space revenue in France

Lastly, we present our estimates of the total downstream space revenue in France for 2021. Table 4.5 details the figures obtained after both imputation procedures:

1. The total 2021 revenue of downstream space companies.
2. The minimum and maximum 2021 revenues from downstream space activities.

For the total revenue, the 'Known' row of the table are the figures obtained from our survey data and the Amadeus database (N=152 companies). The 'Imputed' row represents the total downstream revenue estimated from the median growth rate in procedure 1 (N= 177). Next, for the minimum and maximum downstream revenues, the 'Known' row corresponds to downstream space revenues calculated from our survey data (N=51). The 'Imputed' row refers to the data estimated via imputation procedures 1 and 2 (N=278). The 'Total' row provides the estimates for the total revenue, the minimum downstream space revenue, and the maximum downstream space revenue for the 329 downstream space companies.

	N	Total Revenue	N	Min. Downstream Revenue	Max. Downstream Revenue
Known	152	22,621,275,770	51	1,452,410,036	1,739,261,805
Imputed	177	2,837,875,770	278	2,229,820,450	5,162,533,160
Total	329	25,459,151,540	329	**3,682,230,486**	**6,901,794,965**

Table 4.5: Estimated Total Revenue and Downstream Space Revenue in 2021 (in euros)

Consequently, **French downstream space companies generated 25.5 billion euros in total revenue in 2021** (Table 4.5). Breaking down this total revenue into

'Known' and 'Imputed' groups, we find that the 152 companies for which we collected information via the Amadeus database and our survey represent the majority of this total revenue, at 22.6 billion euros. The 177 companies for which we applied the median growth rate of their group account for only 11% of the total revenue.

Upon closer examination of these figures, we notice the presence of an outlier, that is, a company whose total revenue is considerably higher than that of other downstream companies. This single company constitutes 70% of the total estimated revenue for the sector. Therefore, our subsequent measurements addressed this specific case with a dedicated approach. A detailed explanation of this approach will be provided in the presentation of subsequent indicators.

The final indicator in our general assessment of the downstream space sector is the **2021 downstream space revenue** generated by all companies identified by our identification tool. To provide a more comprehensive and nuanced estimation, we propose two variants of this indicator: a minimum downstream revenue and a maximum downstream revenue. This approach allows us to define a revenue range, highlighting the potential variability in our estimation results.

We obtain these figures by following the same approach we used for calculating the downstream revenue of survey respondents in Section 2. For each company in our database, we have the total revenue in 2021 (collected via the Amadeus database or estimated through Imputation Procedure 1) and a range representing the percentage of revenue attributed to downstream activities (imputed via Imputation Procedure 2). We multiply the known or estimated downstream revenue by the lower and upper bounds of the range to get the minimum and maximum downstream revenue. We then sum up all the downstream revenues to obtain an aggregate measure. Therefore, **downstream space activities in France generated between 3.7 billion and 6.9 billion euros in revenue in 2021**.

Finally, let us return to the treatment of the outlier. Our initial step was to investigate the downstream activity carried out by this company. It is a large company specialized in digital services. It has diversified activities, and space is not its primary market. However, the company is involved in the development of space applications and the provision of value-added services based on the use of satellite images. The KNN model assigned it a downstream revenue percentage of 'Less than 10%', which aligns with the company's description and our expert opinion. Thus, we decided to include this company in the assessment by applying the lower bound of its revenue – i.e., 1% – for both the minimum and maximum downstream revenue measures. We justify this choice by the fact that downstream activities are very residual within this group. From a replication perspective, we suggest adopting this same approach for identifying atypical company cases. These can significantly influence measurement results.

4.4 Conclusion

Main contributions

This chapter was dedicated to assessing the economic weight of the downstream space sector, with two objectives. First, we aimed to establish a comprehensive measurement approach, detailing the tools and steps involved alongside relevant indicators. Second, we applied this methodology to the French downstream space sector in 2021, presenting the initial results.

On the methodological front, we adopted a two-step approach. Firstly, we used a survey to collect detailed information about downstream space companies with specific and recent data unavailable in existing databases. Besides, the survey provided qualitative insights about a sample of downstream companies and helped us understand the sector's structure.

Using the survey data, we elaborated a series of economic indicators, including a new categorization by **circles of actor** that reflects the revenue range directly associated with companies' downstream space activities. It is an essential contribution of our research since it led us to the **downstream space revenue indicator**, which accurately reflects the value generated by exploiting space-based infrastructure, data, signals, and information. We offer two measures for this indicator (minimum and maximum downstream revenue), providing a nuanced estimation window. We also determined downstream revenue per activity, measuring the revenue generated by each downstream space value chain segment. These new indicators supplement conventional ones, such as company creation date, size, and NAF activity sector, providing insights into the sector's structure.

Finally, we outlined our strategy for generalizing the assessment to all downstream space sector companies. This step necessitated estimating the downstream space revenue for all companies. We hence had to collect 2021 revenue data for companies not included in the survey to calculate total downstream space revenue. We executed two imputation procedures as existing databases did not provide complete information. The first imputed the 2021 revenue data, while the second imputed a revenue range corresponding to downstream activities for non-participating companies, using the k-nearest neighbors imputation algorithm.

Limitations and Recommendations for Replication

This evaluation is designed to be replicated for tracking the development dynamics of the downstream space sector, so we provide the main avenues for methodological enhancement.

First, our findings are heavily reliant on survey data. We would have preferred more respondents for more comprehensive results and a lesser proportion of imputed data. We recommend that future evaluations maximize efforts to disseminate the survey extensively to collect information from as many respondents as possible. One idea might be modify the survey's title (and the subject of the email sent to respondents), initially 'Survey of companies in the downstream space segment.' Despite our clarification that companies with even marginal downstream activities were relevant, some may not have identified themselves as falling within the survey's scope.

Second, we highly recommend incorporating the ground mission services segment into downstream activities in future evaluations. Treating this activity separately in the survey was a source of bias in our measures. For instance, some downstream companies stated that their downstream revenue was '10 to 50%', and that their revenue associated with the ground mission segment was also between 10 and 50%. These were clearly identified downstream space companies, yet they fell into the '10 to 50%' circle at the time of our evaluation.

Lastly, an important aspect is that we did not use the standard definition of downstream activities in our analysis: Earth observation, telecommunications, and navigation. It was a deliberate choice because we believe this categorization is restrictive and does not capture evolutions in New Space. For example, some digital companies develop services using space infrastructures and data but do not identify as Earth observation specialized companies. Nevertheless, we should have introduced a question to better understand for what purpose space infrastructures and data are used, e.g., in climate, mobility, surveillance. We encourage this aspect to be considered in future evaluations.

4.5 Appendices

4.5.1 Survey questionnaire (First part, English translation)

Survey of companies in the downstream space segment

This survey is being conducted as part of a thesis work carried out at the University of Strasbourg and supported by CNES. Its objective is to measure the economic weight of the downstream space segment in France.

As part of this survey, the **downstream space segment** covers all activities based on **the exploitation of satellite systems or the production and use of data and/or signals of space origin** to provide **added-value products or services** (including user equipment).

Mission ground segment services, on the other hand, cover activities related to operations of satellites in orbit (e.g. tracking, telemetry and control of satellites, uplink and downlink for signal processing, etc.).

Your response is important, even if you consider that your company only marginally belongs to these activity segments.

Individual responses to the questionnaire will be accessible only by the doctoral student who is conducting this survey. Only anonymous and aggregated information may be disseminated.

Average duration of the questionnaire: 5 minutes

Note: This questionnaire aims to collect information only for companies located on French territory (including subsidiaries of foreign groups located on French territory)

In this survey, the **downstream space segment** encompasses all activities based on **the exploitation of satellite systems or on the production and use of data or signals of space origin** to provide **added-value products or services** (including user equipment).

These activities notably include:

- **Sale or lease of satellite capacities,**
- **Production and sale of space data**,
- **Provision of software and/or equipment for the storage, processing, and use of data, signals, and information provided by satellites**,
- **Value-added services** dependent on space infrastructures, data, or signals,
- ...

Mission ground segment services, on the other hand, **cover the activities of satellite operations in orbit** (e.g. tracking, telemetry, and control of satellites, uplink and downlink for signal processing, etc.)

Mandatory questions are indicated by an asterisk (*)

*What is the name of your company ?

*Do you have any activities in the **downstream space segment**?

Yes No

*Do you have any activities in the **mission ground segment**?

Yes No

*What were your company's sales revenues in 2021?

*What percentage of your sales corresponds to activities in the **downstream space segment**?[1]

- o Less than 10%
- o Between 10% and 50%
- o Between 50% and 80%
- o More than 80%
- o 100%

*What percentage of your sales corresponds to activities in the **mission ground segment**?

- o Less than 10%
- o Between 10% and 50%
- o Between 50% and 80%
- o More than 80%
- o 100%

*In which **downstream space segment** activity(ies) is your company active?

- ☐ Sale or lease of satellite capacities,
- ☐ Production and sale of space data,
- ☐ Provision of software and/or equipment for the storage, processing, and use of data, signals, and information provided by satellites,
- ☐ Value-added services dependent on space infrastructures, data, or signals,
- ☐ Other activity(ies). If so, which one(s): ____________

*What is the breakdown of your company's **downstream space** sales revenues?

Sale or lease of satellite capacities

0 100

Production and sale of space data

0 100

Provision of software and/or equipment

0 100

[1] This question is displayed if the respondent checks 'Yes' to the question about their downstream space activity.

Value-added services 0 100

Other 0 100

The sum should be approximately 100
If you do not wish to answer, please leave the sliders on position 0

4.5.2 Survey results: Revenue from Mission Ground Services Segment Reported by Respondents

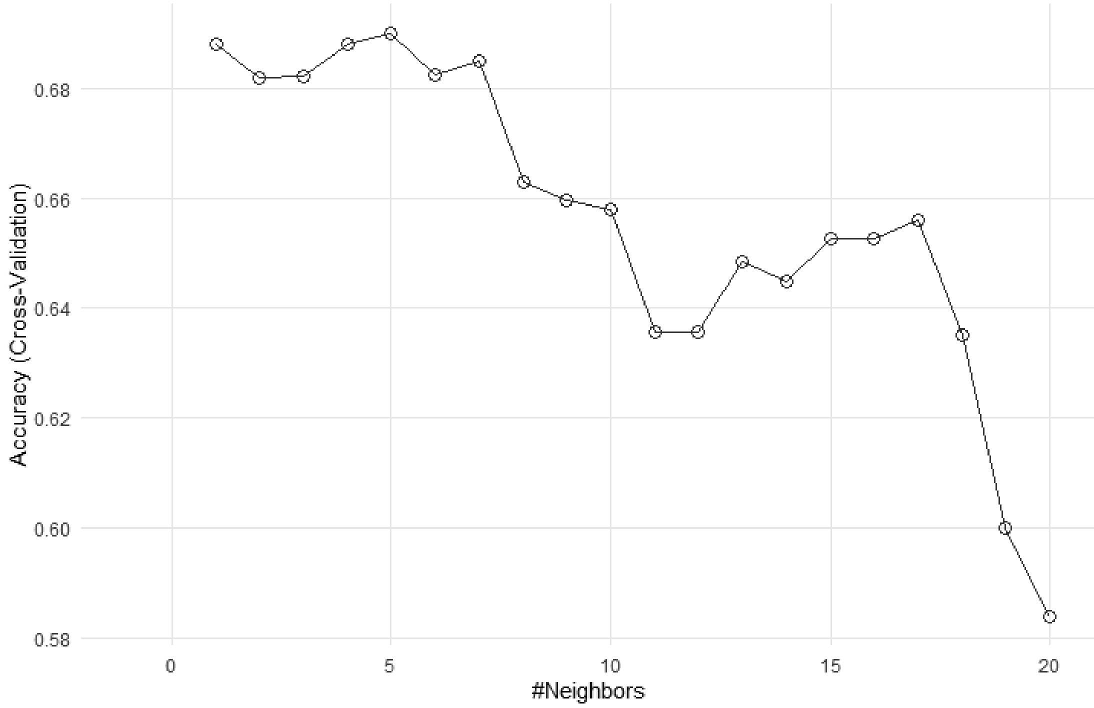
Share of Revenue from Mission Ground Services*	N	Total Revenue**	Min. Mission Ground Segment Revenue	Max. Mission Ground Segment Revenue
Less than 10%	6	1,235,345,000	12,353,450	111,181,050
10% - 50%	8	673,574,483	53,677,448	269,251,497
50% - 80%	0	-	-	-
More than 80%	1	96,000,000	76,800,000	95,040,000
100%	1	2,300,000	2,300,000	2,300,000
Total	16	2,007,219,483	145,130,898	477,772,547

* Revenue shares were converted into the following intervals for ground segment revenue computation: [0, 10), [10, 50), [50, 80), [80, 100), and 100%.

** Two respondents did not disclose their 2021 revenue.

Table 4.6: 2021 Revenue from Mission Ground Services Segment Reported by Respondents (in euros)

4.5.3 K-Nearest Neighbors Imputation: Cross-validation Test Result



Part III. Theoretical Analysis of Value Creation Dynamics in the Downstream Space Sector

Chapter 5

Data, Information and Value Creation in the Digital Age

5.1 Introduction

In this thesis, we developed a methodology for assessing the economic importance of the downstream space sector. To this end, we identified the players in the downstream space value chain and estimated the sector's importance based on revenue indicators. However, our identification and evaluation approaches do not consider the economic properties of the goods and services produced in this segment. Downstream space activities are based on valorizing data of both space and non-space origin to provide information-based services. This final chapter seeks to enrich our empirical developments through an analysis relying on economic approaches to information. Our study will specifically concentrate on the properties of information as an economic good expounded in seminal works [Arrow, 1962], as well as the characteristics that align it with the notion of a public or quasi-public good, as defined by Samuelson [1954]. Harris and Miller [2011] explore the concept of public good applied to the case of Earth observation. The article highlights that satellite images yield positive externalities even though they are not classified as pure public goods. It concludes on the necessity of considering the characteristics of Earth observation data (low rivalry and appropriability) to avoid underestimating the socio-economic benefits of Earth observation. Our theoretical ambition aligns with these authors, except that it targets the entire downstream space sector. We aim to explore and reevaluate old concepts derived from the literature in information theory in light of the new digital context. The purpose is to situate our evaluation tool within a suitable conceptual framework and provide enhancements. This reflection appears indispensable to ensure the long-term relevance of indicators, particularly in an evolving industry with the emergence of new users of satellite data and information.

Indeed, we have observed structural changes in the space sector within a broader context of technological change (Chapter 1). Hey et al. [2009] describe the entry into a fourth paradigm starting from the 2010s, succeeding to the era of information and communication technologies. This new paradigm is characterized by the development of artificial intelligence (AI) technologies such as machine learning, their application segments (e.g., Internet of Things, digital platforms), and the digital infrastructures upon which they rely (e.g., computing infrastructures, data centers). The digitization process, which accounts for the widespread adoption of digital technologies across all economic sectors, science, and public services, involves a massive production of data and an intensification of their usage [OECD, 2022a, Bukht and Heeks, 2017]. The increasing capabilities of digital tools enable the generation, storage, processing, analysis, and dissemination of large volumes of data in multiple formats. Data, becoming increasingly complex in their structure [Einav and Levin, 2014], fuel the models upon which AI systems are trained. These models require vast amounts of data to produce structured information or knowledge. For instance, image recognition involves systems identifying objects, individuals, or activities from images. It is used in various industries, including transportation, agriculture, and surveillance.

Data are a fundamental element of this new techno-economic paradigm. They are considered the primary factor or input of production in the digital era [Li and Li, 2021, Sadowski, 2019]. In their broadest sense, data refer to the information of any nature and form (data collected by sensors, personal data about users, videos, images, texts, sounds) encoded as sequences of 0s and 1s to be readable and exploitable by computers [Goldfarb and Tucker, 2017]. Their omnipresence is manifested first by their large quantity and heterogeneity [Feijóo et al., 2016]. From these two characteristics arises the increased potential for data combination, giving rise to increasingly complex services (e.g., multi-level data services). The actors involved in the data industry chain are also diverse. For instance, Tang [2016] distinguishes three economic entities: data resource providers, data product providers, and data product terminal customers. The roles of these agents in the data industry are often mixed. Large technology companies, for example, are data integrators (data storage, cleaning, transformation), providers of multi-level data products, and clients of data services. It should also be noted that the agents' positions in the chain may evolve depending on the time horizon considered.

The digital transformation, the widespread adoption of AI techniques, and the intensification of data flows resulting from this evolution are crucial factors driving industrial changes in New Space [Bockel, 2018, Frischauf et al., 2018]. In the upstream segment and satellite operations, the impact of digitization is evident through technological innovations such as automated production lines, robotized assembly of solar panels, and autonomous satellite operations [OECD, 2019, Pelton et al., 2017]. In the downstream segment, this new paradigm expands the scope of satellite exploitation and increases its intensity [Evans et al., 2005]. Moranta [2022] describes satellite systems as "becoming a component of the digital infrastructure [...] where space-based data are mingled in data streams, processes, and products as another input in the wider data value chain." In other words, the convergence between the space and digital sectors leads to the democratization of satellites and their outputs. The range of space uses is extended within this new context. Earth observation data, geolocation information signals, and other satellite-based information are integrated into mass-produced products. Similarly, space infrastructure is used as telecommunications infrastructure integrated into terrestrial networks to provide services based on the exchange of information and data.

In light of these elements, we understand that multiple configurations are possible for the downstream part of the space sector. The downstream value chain is neither unique nor static; it undergoes constant changes due to the emergence of new actors, their shifting positions in the value chain, and the formation of new connections. Apprehending the value of information becomes more complex and context-dependent. The economic properties of information, such as non-rivalry, imperfect appropriability, obsolescence, and the existence of asymmetries that distinguish it from physical industrial goods, are modified – perhaps amplified – in this new paradigm.

The value of space-based information has been extensively explored in the literature dealing with the socio-economic benefits or impacts of the downstream space sector. In Earth observation particularly, numerous works aim to evaluate the economic value of satellite imagery. Some studies grasp this value in terms of impact through proxies derived from econometric and statistical estimations. They quantify the costs saved by using satellite imagery in water monitoring [Papenfus et al., 2020] or the beneficial effects of using remote sensing information on land productivity [Nogueira et al., 2018] and urbanization dynamics analysis [Yang and Liu, 2005]. Other works evaluate space data's value in improving the performance of a decision model utilizing information. For instance, Bouma et al. [2009] demonstrate how using information reduces uncertainty in water quality management in the Netherlands. Finally, some analyses focus more on estimating the monetary value of satellite information. They employ survey-based methods to determine the value agents derive from using space data based on their willingness to pay [Eom and Hong, 2013, Jabbour, 2019b, Loomis et al., 2015].

The mentioned references share the common characteristic of evaluating information value from an end-user perspective. They seek to understand how economic actors, decision-makers, and managers can effectively and beneficially utilize space information. Through this approach, they consider the singular nature of space information, and the idea that information value is associated with considerations related to public goods, aiming to guide public investments allocated to space programs [Le Pellec-Dairon, 2013, Tassa, 2020]. A second common trait of this literature is adopting a static approach to assess the economic value of Earth observation data. The studies primarily focus on examining the immediate benefits or returns from agents' immediate or time-specific exploitation of space information, thus not considering the potential evolution of this value with a long-term perspective.

In this study, we propose approaching the value creation process in the downstream space sector from a more dynamic perspective through theoretical reflection. We aim to consider the passage of time to analyze how the stages of data and information production, dissemination, integration, recombination, recycling, etc., are interrelated. More specifically, the research question guiding this final development can be stated as follows: How does the evolving value of data and information, in the context of mass data, transforms the value creation dynamics in the downstream commercial space sectors? This entails two main aspects. Firstly, examining the extent to which the digital context modifies the characteristics of information as an economic good. Secondly, assessing the impact of various configurations resulting from data massification on the value chain in the downstream space sector. Unlike the aforementioned studies, we apply our research question to all downstream space domains, not solely focusing on Earth observation. Thus, we consider all types of content that pass through satellites, whether generated by satellites or not. We provide further justification for this choice in the subsequent development of the chapter.

The remainder of the chapter is organized as follows. In section 2, we define data and information using the literature on information theory. We apply these definitions to the downstream space segment. Section 3 examines the changing dynamics of data in the digital transformation era. We analyze how the characteristics of data and information are evolving in this new context. In section 4, we consider the role of data and information in a dynamic perspective within the value creation process. Finally, the last section is dedicated to applying the theoretical elements developed to our object of study, the downstream space sector.

5.2 Characterizing data and information as economic commodities

5.2.1 Conceptual differences between data, information, and knowledge

The concepts of data, information, and knowledge, while closely related, are often conflated. As a preliminary basis for our exploration, it is essential to delineate these concepts clearly. We deliberately focus on the cognitive dimension of this distinction, viewing it as a relationship between agents and the physical world. This allows us to remain within the scope of our initial research objective: to understand how information and its evolving properties impact the value creation process in sectors or firms that use information as a source of value. In Boisot and Canals [2004], the relationship between information and the other two notions is postulated as follows:

> ”Information is an extraction from data that, by modifying the relevant probability distributions, has a capacity to perform useful work on an agent’s knowledge base.” [Boisot and Canals, 2004, p. 47]

The authors describe data as the material of information. They are *observations* of different states of the physical world, whose properties are represented using three dimensions: time, space, and energy [Ackoff, 1989, Boisot and Canals, 2004]. In the context of organizations, Davenport et al. [1998] define data as a ”set of discrete, objective facts about events [or] structured records of transactions.” Based on these insights, we note that data are elements that are *given*, perceived by the senses or by instruments. They depict the sensory reality about objects, phenomena, or events in a raw form without containing any inherent interpretation or judgment. They result from stimuli originating from physical reality and are structured to be received by agents. Therefore, data is a representation or a construct presented in various formats (numbers, symbols, images, bits, waves, signals) to be understood by the agents [Kitchin, 2014]. It is also noteworthy that data are apprehended only in the plural, as a set.

Information, on the other hand, is described as an interplay between agents and the data they process. It stems from *regularities* or *differences* that agents extract from data for specific purposes. Consequently, it inherently carries context and meaning [Ackoff, 1989, Bateson, 1972, Boisot and Canals, 2004]. In information theory, initiated with the foundational work of Shannon [1948], the concept of information is assessed not by its content, but rather quantitatively in terms of transmission within a communication system [Jabbour, 2019a]. Information is described as a message, in the form of a signal in telecommunications, transmitted by a sender. This message has meaning for its receiver, but may also have an impact on his judgment or expectations [Davenport et al., 1998]. In this regard and as suggested by the quote at the start of our discussion, the idea of a relationship between physical reality and the agent appears important in defining information [Bates, 2010, Boisot and Canals, 2004]. It marks a notable difference with data. A given set of data, representing a state of the world at time t without processing or interpretation filters, remains unchanging. Conversely, two agents will not necessarily extract the same information from a dataset according to their personal characteristics, the way they think, their current needs, etc.

The concept of information is one of the pillars of economic theory [Stiglitz, 2000]. Generally, information is perceived as a commodity traded on the market, whose acquisition or non-acquisition influences the behavior of economic actors. Citing only the pioneering works, Stigler [1961] characterizes information as a scarce economic resource whose procurement incurs costs in terms of time and resources. He refers to the expenses related to information as "search costs" that justify adopting the imperfect information hypothesis to describe the behavior of firms and consumers. Furthermore, Stigler establishes a link between information and economic market efficiency: the rarer (and therefore more costly) the information, the greater information asymmetries, and the less efficient the market. In Akerlof [1970], information is defined as knowledge regarding the quality of economic goods. It is incorporated into market signals to mitigate information asymmetry and adverse selection phenomena. In summary, information in economic theory is considered a variable or a resource that helps to reduce uncertainty in economic transactions. In the initial quote, Boisot and Canals [2004] draw inspiration from the work of Hirshleifer [1973] in the economics of information. The latter envisages uncertainty as the distribution of probabilities that an individual assigns to different possible states of the world. Information is thus a resource that influences the dispersion of subjective probabilities. This capacity of information to reduce uncertainty underscores its importance as a key factor in market efficiency and the optimization of economic decision-making.

Finally, still drawing from the distinction formulated by Boisot and Canals [2004], the concept of knowledge refers to the sum of individuals' expectations, which can be challenged by the information they receive. Knowledge emerges from the ability to convert information into instructions [Ackoff, 1989] and can therefore be defined as processed information [Fransman, 1994]. To go further, Cohendet and Llerena [1999] draw a distinction between the concepts of information and knowledge within the firm. They emphasize that viewing the firm as an information processor focuses on the allocation of resources within the organization. In contrast, considering the firm as a knowledge processor allows to shift the emphasis to the production of resources. More generally, knowledge is acquired through a learning process via instructions transmitted by another agent or through personal experience. In firm theory, a fundamental distinction separates *explicit knowledge* and *tacit knowledge* [Nonaka and Takeuchi, 1995]. Explicit knowledge is formal, and agents can codify and transmit it without bias in the form of data, text, or specifications. In contrast, tacit knowledge is personal and informal, stemming from experience and context. Since agents embody this type of knowledge, its transmission requires their interaction. A last relevant aspect worth mentioning is the effect of information on the agents' *knowledge base*. This base constantly evolves, undergoing corrections and expansions as new information about the physical world is integrated. From this perspective, the dynamic approach is motivated by the fact that information, as an output of a system, acts upon this system to enhance its efficiency. This mechanism is called a *feedback loop* and is investigated in various branches of literature such as that dealing with information systems [Bajaj and Nidumolu, 1998] or knowledge management [Akbar et al., 2018].

From these considerations, it becomes clear that data, information, and knowledge are interrelated. Ackoff [1989] develops a sequential and hierarchical relationship among these concepts and adds the concept of wisdom. He posits that data, information, knowledge, and wisdom represent four categories within the content of the human mind. He establishes a hierarchy both quantitatively and in terms of value, with data at the base and wisdom at the pinnacle. This approach, later formalized as the DIKW (Data-Information-Knowledge-Wisdom) pyramid or model, illustrates an interdependency among the categories: each category emerges from the transformation of the preceding one [Rowley, 2007]. Consequently, there can be no knowledge without information, and no information without data. This modeling has been extensively adopted and reinforced in information systems and knowledge management analyses, further emphasizing the consequential linkages among these conceptual components.

In the remainder of the chapter, we will not go any further into the concept of knowledge and will concentrate on the concepts of data and information. We aim to specifically analyze the value-creation process in transforming space data into space-based information.

In the next section, we return to our case study in an attempt to draw parallels between the conceptual elements discussed and the outputs of satellites.

5.2.2 Data and information in the space sector

Before going further in our study, it seems appropriate to precisely characterize the contents generated and transmitted by satellites. Based on the definitions previously provided, we aim to determine whether these contents correspond to data, information, or knowledge. We limit our analysis to commercial satellites, i.e., those operated to provide value-added products or services. Satellites in orbit differ by various functions: Earth observation, navigation, communications, scientific research, security, etc. The type of data or information they generate and transmit will depend on their function and the instruments they carry.

In the field of space, the terms *space data* and *satellite data* are specifically used in the context of Earth observation. They refer to images and measurements of Earth taken by Earth observation satellites. They are obtained through remote sensing technologies from the instruments on the satellite payloads in orbit. Space data allows determining the physical properties of natural or artificial objects from a distance, based on the electromagnetic waves they emit or reflect. Satellite images, which are a type of satellite data, are characterized by their spatial resolution: low, high, and very high. The spatial resolution refers to the size of the smallest observable element in the raw images acquired by the satellite sensor. A higher spatial resolution means that smaller objects or details can be discerned in the images. More generally, the onboard radar and optical instruments generate images and measurements that carry information about maritime surfaces [Mateo-Pérez et al., 2021], climate, atmosphere and cloud compositions [Hollmann et al., 2013], land use [Kaul and Sopan, 2012], or population movements [Welch, 1980].

In satellite navigation, satellites emit *positioning signals*, which allow the calculation of *geographical coordinates*. These coordinates refer to information about an object's position as provided by satellite systems operating on the principle of triangulation [Bhardwaj et al., 2020]. They result from calculating the distance and time between at least three orbiting satellites and a ground-based receiver using signals emitted by the satellites. These coordinates represent the latitude, longitude, and altitude of the location on Earth's surface where the object equipped with the receiver is positioned [Welch, 1980]. Satellite navigation systems enable real-time position tracking: as satellites continuously broadcast their exact time and position, they allow for tracking objects or individuals over time and space. Additionally, an unlimited number of users can use the system simultaneously, as the position calculation is not interactive: satellites broadcast a signal, and receivers calculate the geographical coordinates of objects without broadcasting them in turn.

Finally, communication satellites serve a different function from the previous two types of satellites, as they do not generate data but serve as relay stations in space [Pelton, 2013] to transmit pre-existing content. Communication systems send and receive signals between a ground-based transmitter and one or several orbiting satellites, which then redirect these signals back to Earth. They facilitate the transmission of data or information over long distances, supplementing terrestrial infrastructure, and in geographical *white areas* where terrestrial infrastructure is poor or nonexistent. The content transmitted by satellites is varied, including voice (satellite telephony services), video and radio content (broadcasting services), text, images, and other types of data (connectivity services). Our aim is not to delve into technical considerations, but analyzing how different contents are transformed to be exchanged via satellites seems particularly interesting for our object of study.

The principle is as follows: any content (video signal, radio, text, image, data) is converted to digital format and then modulated onto a wave to be sent to the satellite in orbit [Maral et al., 2020]. Once received, the signal is processed and then transmitted back to Earth, where it is reconverted into a desired format. In sum, satellite communications can transmit any type of content, often information. However, their transmission requires encoding, a conversion into a data form.

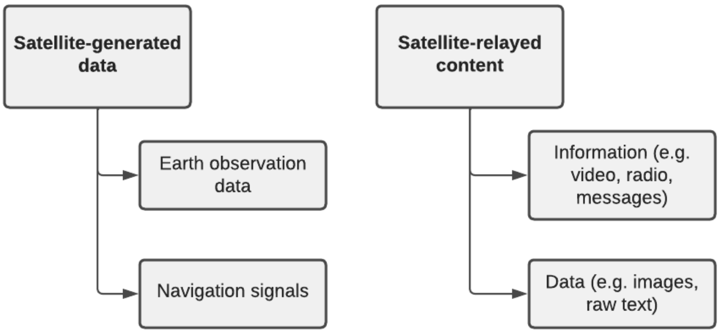


Figure 5.1: Categories of inputs in the downstream space sector: generated vs. relayed (Source: author)

In light of these elements, we propose in Figure 5.1 a categorization of the data exploited in the downstream space sectors. The first category, referred to as *satellite-generated data*, designates data originating from satellites. In other words, these are data whose production involves satellites and would not exist without them. This category includes Earth observation data, i.e., satellite images and all data on Earth acquired via remote sensing instruments. In our approach, satellite-generated data also encompass positioning coordinates derived from signals emitted by navigation satellites. We draw on the definitions provided in section 2.1 to characterize them as data, not information. Indeed, position coordinates represent measurements acquired via satellites in their raw form and without context. However, when these measurements are processed and used for a specific purpose, e.g., in navigation systems, to locate an object, they become information.
The second category we introduce is the *satellite-relayed contents*. Unlike space-generated data, these contents are not necessarily satellite-originated. They refer to content of any type that is broadcasted and transmitted via satellite communication systems. This second category includes information (e.g., television programs, radio programs, messages, telephone conversations) and data (e.g., images). It is worth

noting that all contents are encoded in a specific format to be transmitted on the communication channel. During this process, information is converted into data.

Now that we have defined the data generated and collected in the downstream sector, we propose to analyze the properties of these data as economic goods. This step will allow us to examine the influence of these goods' properties in the value-creation process within the downstream space sector.

5.2.3 The nature of data and information as economic goods

In the previous section, we identified that the primary inputs of the downstream sector are data and information, either generated by satellite systems or only transmitted via satellites. The next step is to delve deeper into the fundamental properties of data and information as economic commodities, drawing on the literature in information economics. This field posits that information is a unique type of good with characteristics that differentiate it from conventional goods [Arrow, 1996]. Information and data being intimately linked concepts, we seek to determine whether the main attributes of information also apply to data.

First, data and information share a common characteristic: they are intangible or immaterial assets [Arrow, 1962, Li et al., 2018]. Unlike conventional industrial goods, they have no physical reality per se. They are incorporated within various tangible and intangible mediums that facilitate their access and exchange, such as books, hard drives, databases, and videos [Hill, 1999]. In their contemporary understanding, data almost exclusively refer to digitized, and therefore intangible, content [Goldfarb and Tucker, 2017, Rosenberg, 2013].

A second property of information highlighted by economic theory is its non-rivalrous nature [Arrow, 1996, Varian, 1999]. An economic good is non-rival when agents' simultaneous and repeated use does not decrease its quantity or value [Romer, 1990]. This characteristic stems from the intangible nature of the information. When information is exchanged between individuals, it is actually duplicated. Consequently, the parties involved in the exchange can continue using it indefinitely and without any utility loss. Moreover, the cost associated with the duplication of information, that is, its marginal cost, is either zero or near-zero [Arrow, 1996] but the cost of producing the first unit of information is often remarkably high. Data share the same properties of non-rivalry and zero marginal cost as information [Carrière-Swallow and

Haksar, 2019, Jones and Tonetti, 2020, Kitchin, 2014]. For instance, the fixed costs associated with building launch vehicles and observation satellites are very high in the space sector. However, once satellite images are collected, their reproduction cost is negligible.

Arrow [1996] and Romer [1990] emphasize the idea that information, as a non-rival good, can be a source of increasing returns to scale.[49] The logic behind any industrial process involving tangible and rival goods is that the greater the investment in inputs, the larger the production volume will be. Arrow emphasizes that this reasoning cannot be applied to information. Indeed, a given piece of information is produced only once, and its repetition adds nothing in value. However, it can be used by the same producer or several different producers multiple times without losing value. The same information can be exploited regardless of the production scale, leading to increasing returns to scale. This assumption is fundamental in economics as it refers to the idea that information is a crucial factor in economic growth. Furthermore, increasing returns to scale enable economies of scale. In this sense, the nature of this good can influence industrial structures where production is technology-intensive, with a tendency towards concentration. Jones and Tonetti [2020] develop an endogenous growth model in which they examine the joint use of data and information (which they refer to as "ideas"). They conclude that data as a non-rival commodity is associated with strong increasing returns.

Furthermore, information is an imperfectly excludable good [Arrow, 1962, Dosi et al., 2006]. In other words, it is either impossible or costly to prevent someone from using it. Unlike conventional industrial goods, the agents who initially produced them may find it challenging to appropriate their value fully. Data, like information, are inherently imperfectly appropriable [Carrière-Swallow and Haksar, 2019]. However, Jones and Tonetti [2020] points out that, unlike ideas or knowledge, data can be easily monitored and secured to be made exclusive. Consider the example of a downstream space company providing a mapping service to the agricultural sector for fertile area detection. Once clients access the service, they can easily appropriate the information about the most fertile lands and sell it to other firms. However, raw satellite images are more challenging to appropriate for at least three reasons. First, interpreting Earth observation data requires specialized skills, and their format is not accessible

[49] Within Arrow's framework, it should be noted that 'information' is employed broadly, encompassing both knowledge and technology in the sense of technical information. Romer's model refers to technology as a non-rival good.

to everyone. Second, data providers (e.g., satellite operators) can restrict and secure their access via encryption. Lastly, unlike information, data do not directly emanate from the goods and services they are embedded in. Using the service directly provides access to information about soil fertility. However, the data used to generate this information are not directly and completely accessible through the service. Unless the database is publicly disclosed, its appropriation by agents not involved in its production is more complex than for information.

An issue regarding information as an economic good relates to the question of obsolescence. The depreciation of physical capital typically arises from wear and tear over time and use. However, information can depreciate and lose its practical value despite remaining in good condition. As an *experience good* [Varian, 1999], the simple act of using the information can lead to its self-depreciation [Dou and Liu, 2013]. Consider the case of journalistic information: once it is disclosed to readers, it immediately loses the initial value derived from its novelty. The situation of technological obsolescence is another example [Bosworth, 1978]. When a firm introduces a new technology, it risks its competitors assimilating the technical knowledge to develop superior technology, rendering the original technology obsolete. Technology, viewed as technical instructions stemming from information, is highly susceptible to depreciation.

Li et al. [2018] argue that data, unlike other intangible assets such as R&D, are not affected by depreciation due to obsolescence. They suggest that data can be aggregated and recombined to create new value. Based on the definition of data discussed in the previous section, different agents can extract diverse information from the same dataset. Consequently, a firm may have access to data and incorporates it into its production. The value of this data depreciates for the firm following its use. However, the same data can be employed by other firms for different purposes, creating new value from the initial dataset. Thus, data obsolescence is never definitive. Consider the example of meteorological data used by an app to generate weather forecasts. Once the information "It will be 20 degrees in Paris on May 1, 2023" is known to the application's users, it loses its usage value. The meteorological data that generated this information becomes obsolete for that specific purpose (predicting the weather on May 1, 2023). However, the same dataset could be reused in 2050 to provide a study on climate change in France since the 19th century. A new value is created from the same data, combined with others, aggregated, transformed, and so on.

Moreover, we assert that the property of indivisibility of information, as stated in Arrow [1962], does not apply to data. Arrow underlines that information is an indivisible good, implying that it cannot be divided into several parts while keeping the same value or utility. Specifically, half of the information does not hold half the value of the complete information. In fact, its value is zero. To illustrate this point, consider again the case of the weather in Paris. If we divide the information such that one party obtains the fragment "20 degrees" and another the fragment "in Paris on May 1, 2023", it would be as if both sides had no information. Now, consider a dataset of global climate data extending over a long period that contains multiple variables such as temperatures, precipitation, wind speed, and humidity rate. Suppose we divide this database into sub-datasets (e.g., all climate variables in a year, all climate variables in a specific country over the years, or one particular variable over several years). In that case, the subsets will still hold value and utility. This aspect could be due to the discrete nature of data [Davenport et al., 1998].

The concept of truthfulness is a final aspect that distinguishes data from information to consider when assessing its value. As [Rosenberg, 2013, p. 18] underlines, "the existence of a datum has been independent of any consideration of corresponding ontological truth. When a fact is proven false, it ceases to be a fact. False data is data nonetheless." The value of information is influenced by whether it is true or false. It is not the case for data, as they do not derive from the judgment of individuals.

Characteristic	Information	Data
Intangible	Yes	Yes
Non-rival	Yes	Yes
High fixed costs - Low/zero marginal costs	Yes	Yes
Excludable	Partially	Partially
Obsolescence	Yes	Temporary
Indivisible	Yes	No
True/False	Yes	No

Table 5.1: Properties of information and data as economic commodities

Table 5.1 provides a summary of the characteristics of information and data as economic goods we have described and compared. The properties of non-rivalry and imperfect appropriability classify them as public or quasi-public goods, as defined by Samuelson [1954]. They refer to situations of market failure in the neoclassical sense with information asymmetry. Therefore, public goods are often associated with underproduction issues: agents have low incentives as they cannot appropriate the benefits of their production. Moreover, public goods generate externalities. In this context, assessing the value of information and data must consider these different aspects.

5.3 The influence of digital on data and information properties

In the introduction of this chapter, we noted a new techno-economic paradigm that Brynjolfsson and McAfee [2014] refer to as the *second machine age*.[50] This new age

[50]According to the authors, the first machine age corresponds to the Industrial Revolution and the invention of the steam engine.

is associated with technological advances in the digital domain involving computers, software, and communication networks. Technology tools such as artificial intelligence, machine learning (a sub-field of AI), and robotics dominate this paradigm. They rely on system learning from experience to perform or automate tasks. Another characteristic is the unprecedented development of communication infrastructures and increased connectivity. The Internet, cloud computing, and the Internet of Things extend the communication capabilities between individuals, firms, and objects. Therefore, technologies in the second machine age confer another scale to the data economy. Data is the main input of production involving these technologies, giving rise to what we refer to as the *data-driven economy* [Ciuriak, 2018]. As these technologies continue to evolve, the way we conceptualize and analyze data and information may need to adapt. Since data and information have taken on a central role in economic and industrial processes, we believe it is essential to consider the implications of this new technological context on their inherent properties. In other words, we investigate whether digital technologies alter the nature of data and information outlined in the previous section, and if so, how.

In the data-driven economy, the non-rivalry nature of economic commodities is amplified. This can primarily be attributed to a scale effect: an increasing number of goods are represented in digital format and a growing number of users are equipped with technologies enabling access to data and information. Moreover, replicating content in the digital format is extremely low-cost, if not entirely free, and immediate. The exponential progress in the digital sector [Brynjolfsson and McAfee, 2014] allows for the increase in digital power, i.e., the capabilities of storing, processing, and transmitting data on the one hand, and a decrease in the cost of equipment on the other hand, which amplifies the non-rival characteristic. The evolution of technical and technological capabilities facilitates many users' access, sharing, and simultaneous use of a large volume of data and information without compromising their quality or utility. In our initial definition of non-rivalry, we spoke about how, at the end of exchanging information or data between two agents, they were not held solely by the agent receiving them but by both sides. Digital tools allow the involvement of many agents in the exchange without the marginal cost of reproduction increasing accordingly. Moreover, non-rivalry is reinforced by the abundance of data generated by digital interactions. With the adoption of connected devices and the Internet of Things, vast data are continuously produced and made available for different users simultaneously. Take the example of public transportation equipped with Internet of

Things devices that collect data on bus locations, passenger numbers, and accidents. The real-time data serve several purposes simultaneously and for many users: the agency operating the transport network (maintenance, fleet management, scheduling) and the users (trip planning, consultation of crowd levels).

The non-rivalry nature of data has implications for its appropriability [Goldfarb and Tucker, 2017]. The concept of property used for private goods does not truly apply in the digital context, so much so that we speak instead of access to data [Varian, 2018]. Indeed, data is typically stored in warehouses, on cloud servers, or through network systems that allow distribution across multiple devices [Carrière-Swallow and Haksar, 2019]. Even though access to this data is limited to authorized users, these systems primarily rely on internet networks and are vulnerable to cyber-attacks and hacking risks. Ensuring data security and managing access to it requires substantial investments. Additionally, the non-rival characteristic of information makes maintaining the private nature of digital information goods challenging. Despite access being restricted to a select few, information replication's simplicity and low cost mean that it can be rapidly and widely disseminated.

Another point we discuss here related to the divisible nature of data is the question of combination, constituting a decisive dimension in the value-creation process in the digital context. Indeed, Brynjolfsson and McAfee [2014] discuss the combination capabilities offered by digital technologies as one of the engines of technological progress in this new paradigm. What we call *big data*, i.e., the manipulation and analysis of massive datasets [Kitchin, 2014], is an illuminating example. The characteristics of big data commonly cited are volume (large amounts of data generated), variety (multiple sources and multiple formats of data), and velocity (speed of creation and data flows) [Al-Mekhlal and Khwaja, 2019]. Big data is the primary resource of computer analysis systems like machine learning. Broadly speaking, these techniques account for a computational and automated learning process from massive and complex data to extract value. This process aims to learn from experience, that is, to analyze data to identify particular patterns and trends. The models detected from the data are then used to describe or predict phenomena or to make decisions. Some data mining tasks allow analyzing and linking data within large datasets [Miller and Han, 2009]. One of these tasks is associations, which consists of finding interesting relationships between several variables or observations in one or more databases. Another technique is the detection of deviations or outliers, which identifies elements in the data that have non-standard characteristics compared with expectations.

Thus, data analytics via notably machine learning tools highlights the paramount importance of combinations allowed by data's discrete and divisible nature. An observation or subgroup of data is extracted, cross-referenced, grouped with data from other sources, and combined to identify a particular pattern. The more numerous and diversified the data, the better the learning capabilities and the more precise the model is in its analysis and prediction capacity. In short, artificial intelligence techniques aim to automate a cognitive task discussed in section 2: extracting information from raw data on physical reality, some would even say knowledge. Unlike the human mind, machines can assimilate and process large volumes of data faster, thus identifying much more complex information. The automation of transforming data into information has eminently important consequences on the information cycle, that is, the process of transforming data into information and then into knowledge, and gives rise to multiple and complex configurations in terms of industrial processes exploiting data. We also note that once produced, information can be transformed into data, in the digital object sense, to be reintegrated into the value-creation process for another use.

In this context, the concept of obsolescence must also be reconsidered. Data and digital information, as objects of massive databases, can undergo phases of obsolescence, but these can only be temporary. First, data have a dual use in the field of artificial intelligence. They are used to train learning models and to describe or predict a phenomenon from this model. In the learning phase, data can retain their value or utility since we noticed that the more incoming data there are, the better the model. In addition, data or information in data form can be subjected to obsolescence only contextually. They are obsolete for a particular use at a given time but can be used in another context.

The case of multi-sided or multi-level platforms also illustrates amplified combination processes in the data-driven economy. These platforms have two main characteristics: 1) they allow direct interactions between two or more agents or groups of agents (suppliers and demanders), and 2) each agent or group of agents is affiliated with the platform [Hagiu and Wright, 2015]. Moreover, they are complex architecture services involving several levels of software. Each level (user, business, data interfaces) gives rise to several types of exchange that induce multiple data combinations. The first type of interaction refers to intra-layer interactions between users and the platform. The user has access to the information provided by the service, and the platform collects data on customer profiles. A second type of interaction occurs between the different levels of the platform, referring to inter-level communications. For example, suppliers at the business interface retrieve information about customer preferences from the presentation interface and combine it with content from the data layer to offer a customized service to each user. Finally, data exchanges can take place between different platforms. Two different systems, located on two different platforms, communicate with each other and combine their respective information to improve their service. For example, a user of a travel booking application wants to filter destinations based on the weather. The booking platform can thus interact via specific interfaces with a weather application to retrieve data from countries where the weather is most favorable at the user's booking dates.

Therefore, these platforms generate data (including personal data on users they collect) and perform combination processes involving data created initially for other uses to provide customized services, thus carrying more and more value. In addition, multi-level services induce positive externalities due to network effects: their value depends on the number of users on both sides of the platform in the case of a two-sided platform [Anderson Jr et al., 2014]. From a provider's perspective, a platform with a significant customer base is more appealing. Conversely, users gravitate towards platforms offering various services and hosting a large user base.

In the present section, we have reconsidered the essential properties of data and information within an economy governed by digital technologies. In this paradigm, data are critical inputs, while outputs materialize as information-based digital goods and services. We highlighted the intensification of the effects of the non-rival nature of data and information given the manipulation of massive, large-scale data flows by various actors (platforms, service providers, users). Furthermore, we have emphasized the combinatory capacity and the unlimited reusability of data and information across various applications and contexts. This observation sheds light on the temporary obsolescence and the inherent divisible nature of data.

In the remainder of our exploration, we will consider the implications of these properties and underlying mechanisms on value creation processes. We will adopt a dynamic approach to analyze: 1) how to represent value chains in light of the nature of data and information and 2) how to assess the value of data within this context.

This section addresses how data and information properties shape value-creation processes. We aim to characterize data and information as elements of a dynamic process to understand how these affect the configurations of value chains within an industry that utilizes data to produce information goods and services.

5.4 Data and information dynamics in value-creation processes: a fund-flow approach

We now consider data and information from a dynamic perspective to account for the effects of their particular characteristics on value-creation processes. The notion of *process* suggests integrating the temporal dimension into the analysis of a productive system. It also refers to the aspect of change, of transformation allowed by a sequence of activities with a well-defined boundary leading to the production of new outputs. We propose to address these questions by borrowing insightful concepts developed in the economic theory of Nicholas Georgescu-Roegen (NGR).

His analysis, known as the bioeconomic approach, stands out for the parallel made between economic dynamics (such as production and consumption) and the dynamics of physical matter transformation [Ferrari, 2021]. Specifically, NGR argues that economic processes are governed by the entropy law, or the second law of thermodynamics, according to which, in a closed system, the degradation of energy is an irreversible phenomenon. He states that the irreversible dissipation of resources is one of the primary sources of scarcity in economics [Dulbecco and Garrouste, 2004, Georgescu-Roegen, 1971]. A second aspect that accounts for the originality of this approach and makes it a dynamic analysis is the conception of time. NGR distinguishes between mechanical time (t) and thermodynamic time (T) [Dulbecco and Garrouste, 2004]. The former refers to the traditional understanding of time as a cardinal variable that measures an interval between two moments. The latter refers to an ordinal variable that allows ordering moments and thus comparing states of the world between different Times T. NGR then introduces the notion of process to describe the periods of resource transformation. This transformation is not measured only in quantitative terms (increase or decrease in capital, labor, production) but also and especially in qualitative terms. We observe that considering the effects of time allows for a qualitative analysis of the economic process. This aspect of the analytical framework is of particular interest to us since we have seen through the properties of data and information that their value-creation process cannot be assessed solely in quantitative terms.

Like a biological system, every production process is characterized by physical and temporal boundaries [Georgescu-Roegen, 1970]. The focus of analysis in this process is a set of events involving the entry and exit of elements (inputs and outputs). NGR defines two categories of elements that are differentiated based on whether they appear only as inputs or outputs or whether they enter and exit the process in the same state. The first category is called *flow elements* and refers to the objects of transformation. Input flows are drawn from stocks and are transformed to produce output flows. In other words, they undergo a qualitative change, and it is impossible to recover the input flows in their original form at the end of the production process. Output flows refer to the objects that were not present at the beginning of the process and appear at the end. A typical example of input flow is the baker's flour, drawn from a flour stock. One of the output flows from the baker's production process is bread. More generally, input flows include natural resources (raw materials) and intermediate consumption. Output flows refer to production and waste associated with the process [Couix, 2020, Dulbecco and Garrouste, 2004].

The second category refers to the *fund elements*, characterized by their "economic invariableness" [Georgescu-Roegen, 1970, p. 4] during the production process. In other words, funds enter and exit the process unaltered and in an identical form. These are agents of production: they transform input flows into output flows. They provide a service without alteration and are maintained at a constant efficiency level. Unlike stocks, funds are not subject to accumulation or decrease but rather to the phenomenon of wear over time. Take the bakery example. The oven is a fund factor as it is used to bake bread, remains unchanged at the end of each bake, and is equally effective for the next batch. In contrast, the stock of flour decreases after each preparation. Other examples of fund factors are fixed capital (equipment, machinery, factories), labor, and agricultural land [Couix, 2020].

To which category of elements do data and information belong within the production process of information-based goods and services? In computer science and communications, we often refer to *data* or *information flows* to describe the continuous transmission of content from one source to another. Given the nature of data and information we have extensively detailed in this chapter, we are asking whether they correspond to flows in the sense defined in the flow-fund model.

In their research on the economics of digital platforms, Li et al. [2018] propose a representation of the data value chain comprising four segments: data collection, data storage, data analytics, and data-driven business models. The first segment comprises gathering data from users and third parties. The subsequent segments encompass the processing, analyzing, and integrating of this data ('data collection' and 'data analytics') to develop business plans (fourth segment). Business plans rely on two primary sources of value: delivering information-based value-added products and services and monetizing access to the data. In the former case, data enter the process, where they are processed, analyzed, and combined with other data to generate information. This information is subsequently incorporated into targeted value-added products and services. According to them, the data collected at the start of the process are no longer visible in their original form. In the latter case, raw data also undergo a series of transformation stages, including visualization, processing, and integration with other databases, before being made accessible through licensing. In both scenarios, data are input transformed during production. They experience a qualitative change during the process that alters their original form. In this respect, we categorize them as *input flow factors*. In their work, Li et al. [2018] make no distinction between data and information, focusing more on the concept of data and its valorization in the digital context. In our research on the space segment, we define data as sets of observations or elements recorded by instruments (e.g., radar) or signals emitted by satellites, for example.
Information, on the other hand, refers to a combination of data that has meaning. In the case of the downstream space segment, this corresponds to information relayed by telecommunications satellites (e.g., TV programs) or information-based products and services (e.g., cartography). Therefore, we define the information integrated into information-based products and services as *output flows*. Note that information can also be considered input flows if re-encoded into data in a given production process.

However, we note that some aspects of the dynamics of data and information within the production process make them different from standard flows. We mentioned earlier that flows are most often drawn from a stock subject to accumulation and depletion constraints. Consider the case of the production of electronic components. One of the input flows of the process is lithium, used in battery manufacturing. Lithium is a metal drawn from limited reserves: when a given quantity is used to produce a battery, the company's stock, and consequently the world's lithium reserves, decrease by the same amount. The transformation process of lithium (input flow) will allow the increase of one unit in the battery stock (output flow). Now, consider a set of satellite images collected to produce an information-based service for precision agriculture. The service provider acquires the data at the beginning of the process. These are stored, analyzed, extracted, and combined with other types of data (e.g., agronomic data, yield data, soil data) to generate information about agricultural lands. The satellite images are transformed during the information generation process and are only present in an altered form in the output. However, the stock of initial data acquired by the company does not reduce at the end of the process; it remains identical. The images used are still available in their original form in the database (i.e., the stock) of satellite images purchased by the company. This phenomenon is due to the non-rivalrous nature of data and information. In the same way, the company can provide the precision agriculture service (i.e., the information) to as many users as it wishes without its stock of outputs and the amount of information available to each user decreasing.

It does not imply that data and information stocks never decrease. Data and information are destroyed, lost, or damaged in companies and administration servers. Merely it is not their consumption that leads to the deletion or alteration of their initial form. Furthermore, data and information do not possess this particular characteristic due to their intangible or immaterial nature. For instance, time is considered a flow factor, even though it does not exist as stock [Georgescu-Roegen, 1971]. Nonetheless, it is impossible to recover the time spent on production at the end of the process. In this sense, data and information are distinct flows because, despite transforming, their quantity in stock and initial quality do not decrease when used as inputs in production processes.

If we describe data and information as particular types of flows, it is because some aspects of their dynamics bring them closer to fund elements. We have already established that due to their non-rivalry, even when they enter the process as input flows to be integrated into value-added goods and services, they still can be found in their initial state at the end of the process. This specific characteristic is reminiscent of the invariable nature of the fund elements, with the difference that data and information undergo a qualitative change during the process. We can also imagine some cases where data are not the object of change but rather a variable that improves industrial production processes or decision-making by agents. In this case, data will provide a service during the production process similar to fund factors while also undergoing a qualitative change. They will be transformed into information and then into instructions (or knowledge) so that, at the end of the process, they will appear as a byproduct, leading to the creation of organizational capabilities acquired by the company, for example [Davenport et al., 2001].

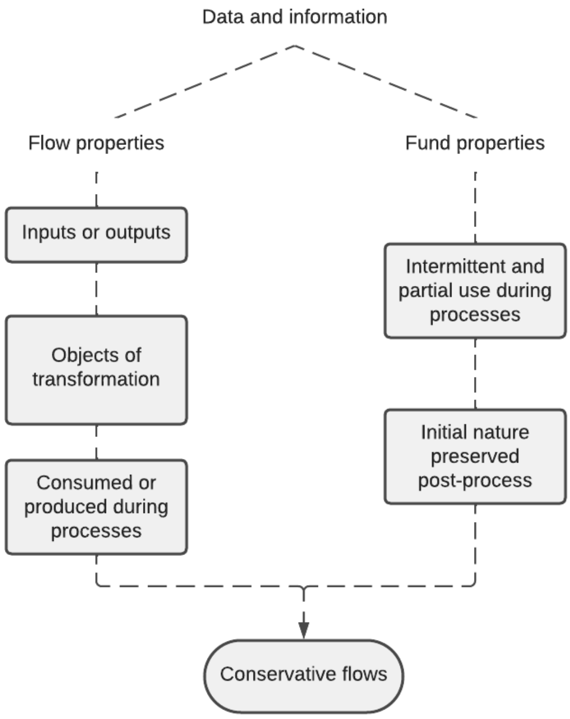


Figure 5.2: Data and information as conservative flows (Source: author)

Therefore, we introduce a new category of elements to describe the dynamics of data and information within the value-creation process that we call *conservative flows* (Figure 5.2). Data and information possess the characteristic of flows as objects of transformation in producing value-added information-based products or services. We define them as flow elements due to the flexibility in their usage. In the same way as raw materials, data can be fractionated, some observations extracted, others set aside, and aggregated. We qualify these flows as "conservative" by analogy with conservative forces in physics to illustrate the idea that regardless of the *transformation path* they take in processes, data and information retain their initial energy (i.e., form and quantity) in the stock. This characteristic distinguishes them from standard flows whose transformation is irreversible.

Figure 5.3 illustrates the dynamics of data and information as conservative flows in the production process of an information-based good or service. It attempts to summarize the properties of data and information discussed in this chapter from a dynamic perspective. It also takes into account the technological and technical capabilities acquired in the digital economy (increased flow speed and combination capabilities). Each segment corresponds to an *elementary process* that extends over a certain duration. NGR defines an *elementary process* as "the process defined by a boundary such that only one unit or only one normal batch is produced" [Georgescu-Roegen, 1984, p. 25]. In other words, an elementary process represents a sequence of transformation of input flows by a certain level of fund factors, resulting in a company obtaining an additional output unit. In our analytical framework, an elementary process corresponds to a unit or batch of information produced by the same or different firms. To simplify our analysis, production line (a) and production line (b) correspond respectively to the activities of firms (a) and (b). The duration of elementary processes is equal to T in firm (a) and $2T$ in firm (b).

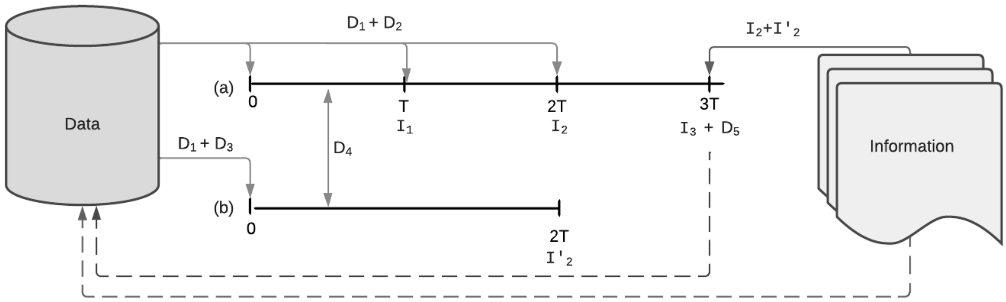


Figure 5.3: Data and information dynamics in production processes (Source: author's elaboration)

The two firms access databases with identical content to produce information goods of different natures. For its production, firm (a) combines the data or observations D_1 and D_2 for three successive elementary processes. We note several elements that stem from the conservative nature of data and information flows. Firstly, their use in the process $[0, T]$ does not lead to their degradation or disappearance. They remain identical in the initial database and are used in the following processes at different times, and their transformation does not result in a reduction of their initial stock. Secondly, each online elementary process of firm (a) necessarily involves the production of a different good (here, information). In the case of producing a tangible industrial good, say assembling cars, each elementary process arranged in series leads to the production of an additional car. Thus, at the end of $3T$, three cars are assembled. For a software developer, the period T is sufficient to produce three identical versions of the software since the reproduction time of information is immediate or quasi immediate. Each process thus leads to the production of different pieces of information I_1, I_2, and I_3, derived from the combination of D_1 and D_2 with other data (such as user data for personalized services) or integrated into different technologies.

Firm (b), on the other hand, extracts data D_1 and D_3 to produce information I'_2. We note that the processes of the two firms take place in parallel and that the data D_1, as a conservative input flow, is simultaneously exploited without loss of utility for the two agents. Moreover, D_4 indicates the exchange of data flows between the firms during two simultaneous processes. The exchange may encompass information regarding their clients or data from alternative sources.

Finally, we refer to the information produced (I_1, I_2, I_3 and I'_2 in Figure 5.3) as conservative output flows. They result from a transformation process involving data, information, and fund factors (e.g., satellites, computers, worker skills). Once generated and disseminated, the information can be integrated into a new elementary process as incoming flows, as shown by the arrow $I_2 + I'_2$ in period $3T$, to improve an existing value-added service, to generate a new one, or to enhance the production process. Once disseminated and used, the information can be converted into digital data and stored to be used in another value-creation process (dotted blue arrow in Figure 5.3). Similarly, an elementary process can generate data and information-based services ($I_3 + D_5$). If we take the example of platforms, they collect data on their users, which they monetize by providing access to it later on.

A last aspect of our analysis concerns the issue of data and information obsolescence. That is to say, the idea that the value or utility of data and information decreases and implies their non-use. We want to draw a parallel with the *idleness* of fund factors raised in the original flow-fund approach. NGR explains that fixed capital undergoes idleness phases during an elementary process, i.e., non-use or non-efficient use. This phenomenon occurs, for example, when a baker can only prepare the equivalent of 50 loaves of bread in an hour while his oven (fund factor) can bake 100 loaves in an hour. We understand that the fund's idleness is a technical problem whose source is fundamentally physical or material. Conversely, the obsolescence of data and information is contextual. A dataset may be of no value at a given time and for a given firm and may gain value at another time and for another use.

5.5 Implications for representing and assessing a data and information-driven sector: the case of downstream space

This final part examines the implications of our theoretical development for value-creation processes in the downstream space segment. Specifically, we apply the analytical model developed in the previous section to different cases of downstream space companies. This approach enables us to investigate how data and information, as conservative input and output flow factors, impact value chain configurations. We then examine how to consider the specific nature of these factors when assessing the economic importance of a sector, using our case study of the downstream space segment.

In this respect, we rely on the information gathered from the second part of the survey sent to companies operating in the downstream space sector in France in 2021.[51] This part of the survey was submitted to companies that declared one or both of the following responses regarding their downstream activities: 1) provision of software or equipment for the storage, processing, or use of space data, signals, and information; 2) provision of value-added services dependent on space infrastructure, data, or signals.

5.5.1 Main survey results (second part)

In the second part of the survey, the questions focused on two main aspects.[52] Firstly, we looked into the nature of these companies' offering and how data (both space and non-space) are integrated into their production processes and the products they deliver. Secondly, we asked about their customers to better understand what happens at the end of the space value chain.

In total, 43 companies indicated that they provide software and/or value-added services. Among these, 9 operate in both software provision and value-added service

[51]A detailed description of the survey can be found in Section 4.2.1 of Chapter 4
[52]For the second part survey questions, refer to Appendix 5.7

delivery. 12 companies provide software but do not offer value-added services. Finally, 23 companies deliver value-added services but no software or equipment for storing, processing, or using satellite data or signals.

The first result from this second part of the survey refers to companies that provide software or equipment for exploiting data, signals, and information supplied by satellites. Figure 5.4 shows that 57%, or 12 of the 21 companies involved in this segment, stated that the equipment and software they deliver are not exclusively dedicated to space data. This result points to two possibilities: either the software and equipment combine space data with data of other types, or the equipment and software can be used for other purposes that do not involve satellite data and signals. Among the 9 respondents who stated that their software and equipment were exclusively dedicated to satellite data and signals, 3 companies provide satellite navigation equipment.

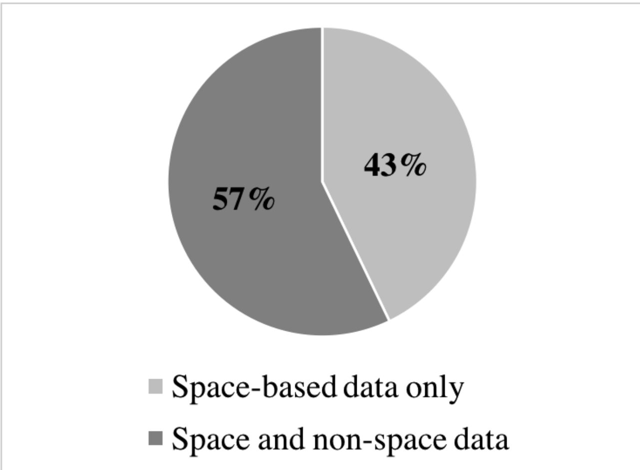


Figure 5.4: Distribution of respondents according to whether software and equipment are exclusive to space data

The subsequent results are based on the responses from the 32 companies that claimed to provide value-added services. Figure 5.5 (a) shows that 25 downstream space companies, or 78% of value-added service providers, combine space-based data with other data types. This figure underscores an important aspect that we discussed in our theoretical development: the combinatory nature of data. In addition, almost 60% of them also make use of big data (Figure 5.5 (b)). This result shows that the respondents handle large volumes of data in their service offering.

The fact that most respondent companies generate information from high-volume

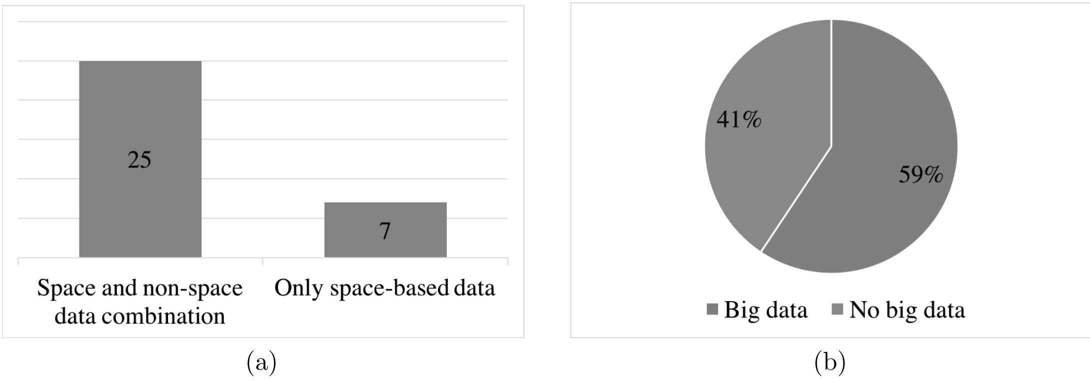


Figure 5.5: Survey results for value-added service providers – Combination of space and non-space data (a); Use of big data (b)

combinations of space and non-space data has implications for the representation of downstream activities. Indeed, we may wonder whether the value chain representation allows us to efficiently detect the multiple configurations of actors involved in producing information-based services. For two services that integrate the same satellite images but for different uses, we can imagine that the images are combined with different space and non-space data and, therefore, that the production process for each service involves different space and non-space actors.

We also surveyed value-added service providers about their customers. Figure 5.6 (a) presents the responses from survey participants regarding the type of clients they target with their value-added services. When asked, ”Your value-added service offering is primarily targeted at:”, respondents could choose between: businesses (BtoB), individuals (BtoC), or public administrations or associations. Multiple answers were permissible. It appears that the most represented client types are businesses and public administrations, with 28 and 26 responses, respectively. Only 5 out of 32 companies indicate they provide services to individuals. Moreover, all companies that deliver services to individuals (BtoC) also deliver services to businesses (BtoB). Thus, the fact that 88% of respondents state their services are primarily aimed at businesses led us to question the end of the value chain. Client businesses may use the information integrated within the value-added services to produce value. In this way, the downstream space segment might extend to these businesses.

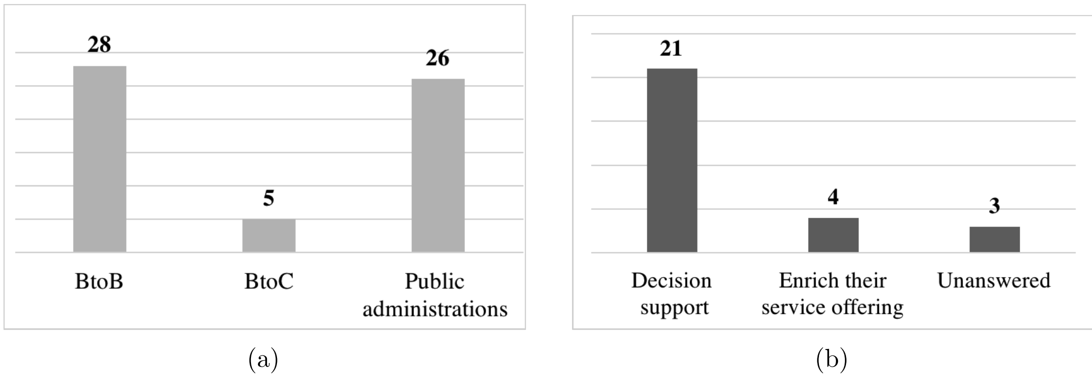


Figure 5.6: Survey results for value-added service providers – Customer types (a); Type of use of downstream space services (b)

In this regard, we asked the companies how their clients use their value-added services. This question was only for those who answered 'BtoB' to the previous question. For this question, only one choice was possible among three options: 'as a decision-making aid', 'to enrich their service offering', or 'declines to answer.' Most client companies use the respondents' value-added services as a decision-making aid (Figure 5.6 (b)). Additionally, 4 respondents indicated that their services were used to enrich the service offering of their clients. This result is interesting as it suggests that the downstream space value chain continues with these clients, who exploit the information contained within the service to generate additional value.

In addition, we asked the respondents if their services required information about their clients' needs. With this question, we intended to see whether the service offering in the downstream space segment was customized. Figure 5.7 shows that 22 out of 32 companies use information about their clients. In other words, nearly 70% of the surveyed downstream space service providers provide customized services.

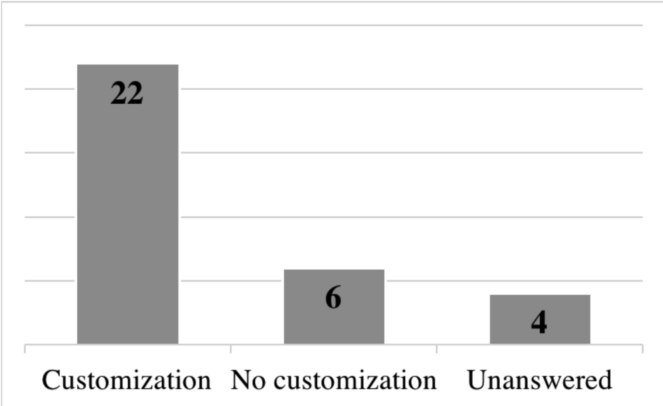


Figure 5.7: Distribution of respondents according to the customization of their services

With this question, we sought to explore the demand-pull hypothesis discussed in Chapter 1 of the thesis. Indeed, if downstream space services are not generic and they take into account users' needs, this could be an additional indicator illustrating the paradigm shift in the space sector.

Lastly, the survey explored two key areas: the reliance of services on satellite data and signals, and the cost of access to space infrastructures or the cost of acquisition of satellite data for respondents. For the former, we asked: "To what extent are the value-added services you offer critically dependent on space infrastructure, data, or signals?". Respondents could choose from three answers:

1. The services would not exist without space infrastructures, data, or signals.
2. The quality or performance of the services would be significantly degraded without them.
3. Their quality or performance would be marginally degraded without them.

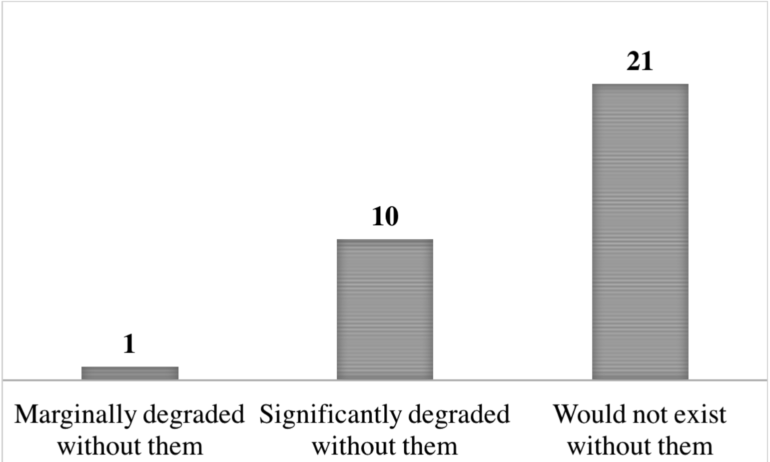


Figure 5.8: Distribution of respondents according to the dependence of their services on satellite infrastructure, data or signals

Most respondents reported a significant dependence on satellite infrastructure, data, or signals, as shown in Figure 5.8. Specifically, 21 out of 32 companies stated their services would not exist without them. Another 10 firms indicated that the quality and performance of their services would be significantly degraded without satellite resources. Only one company reported a marginal degradation.

Next, we asked respondents about the proportion of costs associated with accessing satellite infrastructure or acquiring satellite data/signals in relation to their total production costs for downstream services. They could choose from several ranges: 0% (data or signals are freely accessible); Less than 10%; Between 10% and 30%; Between 30% and 50%; More than 50%. They also had the option to decline to answer.

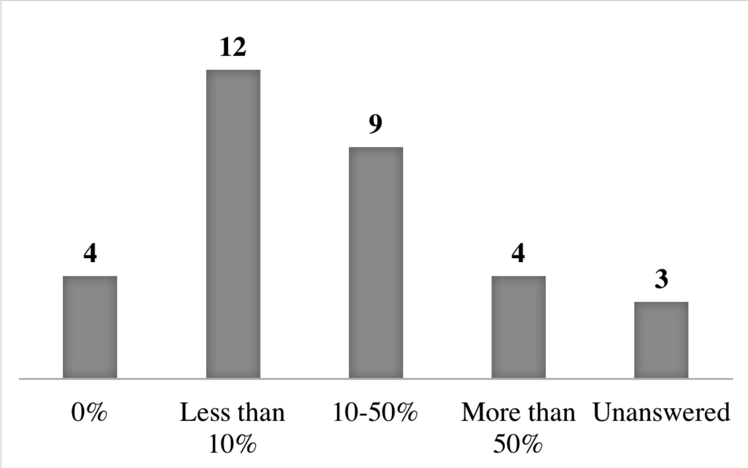


Figure 5.9: Distribution of respondents by share of infrastructure access costs/data or signal acquisition costs in total costs

Figure 5.9 displays the distribution of responses. For clarity, we aggregated the 'Between 10% and 30%' and 'Between 30% and 50%' responses. Four service providers reported that accessing satellite infrastructures or acquiring signals was free. One of these firms provides navigation services based on GNSS signals, while the remaining three likely rely on satellite imagery. They may exploit data from freely accessible online satellite imagery (e.g., Copernicus program data platform).

On the other hand, 12 companies stated that costs related to acquiring satellite data or infrastructure amounted to less than 10% of their total production costs. This suggests that for half of the respondents, access to satellite infrastructure or data is relatively low in cost, even though their services are highly dependent on satellite data and signals.

Finally, 9 respondents revealed that the costs associated with satellite infrastructure or data access made up 10 to 50% of their total costs, and for 4 respondents, this figure exceeded 50%. Notably, 70% of those who reported costs higher than 10% are also involved in capacity sales/leasing or data sales. These diversified companies operate satellites or produce data themselves, which explains their significant infrastructure and data access costs.

We will now consider two respondents' answers in detail and apply a portion of our analytical framework. This approach aims to explore the question of value chain configurations in the downstream space segment.

5.5.2 Implications on downstream space value chain configurations

We apply our analytical framework to two cases of downstream space companies to illustrate the different configurations of the downstream value chain in this sector at company level. However, we draw attention to an important point regarding the cases presented in this section. Given the nature of the data collected in the survey, we could not incorporate every element of our model. As such, we focus on a singular elementary process for each company, i.e., the successive production stages leading to an additional unit of an information-based good or service. Similarly, the survey results do not allow inference of inter-company connections; hence, we do not consider data exchanges between companies within the same period.

5.5.2.1 Company A: Diversified Earth observation downstream company

The first case we analyze is a company (Company A in the following) providing satellite services in the domain of Earth observation. Based on the downstream space company classification developed in Chapter 4, it belongs to the second circle of players. In other words, more than 80% of its total revenue depend on downstream activities. Additionally, Company A reports that the remaining 20% of its revenue is attributed to activities in the mission ground segment.

Figure 5.10 provides a simplified illustration of the data and information dynamics within a given production process for Company A. The company has diversified activities; it is involved in four segments of the downstream space value chain: lease and sale of satellite capacities, production and sale of satellite images, provision of software or equipment, and value-added services delivery (VAS_A in Figure 5.10). If we consider the activity of producing and selling data, the arrow with green D_A illustrates a data set generated and sold by Company A at a given period. Since the data – satellite images – are conservative input flows, we can imagine that D_A is sold and exploited by several companies simultaneously, without reducing Company A's stock of satellite-generated data.

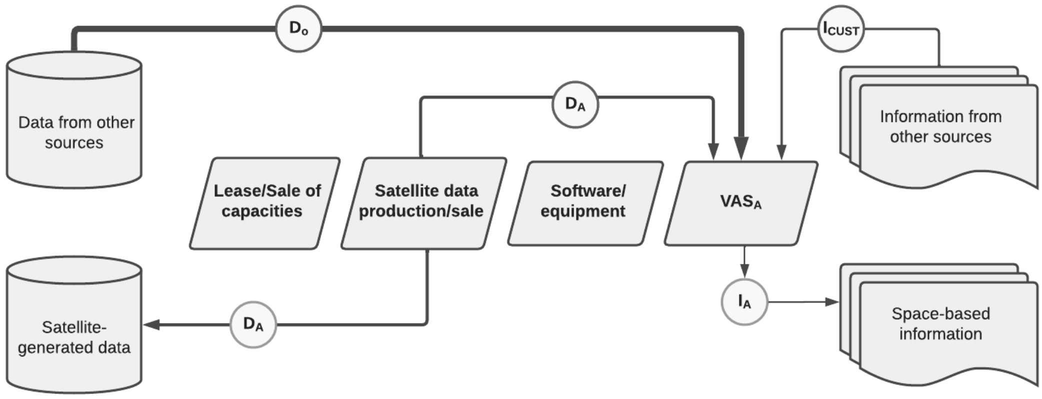


Figure 5.10: Company A: Diversified EO downstream company

The company supplies software or equipment and value-added services dependent on space infrastructures, data, or signals. To illustrate the non-rival nature of data at the company level, we have represented by the two arrows D_A that the same dataset can both be sold and exploited by Company A for its downstream services (red D_A) while retaining its initial form in the stock of satellite-generated data (green D_A). Again, this aspect of our analysis is fundamental and justifies the introduction of data as a particular type of factor in the production process. In the case of the production of a standard good such as a car, for example, a given quantity of input flow (e.g., aluminum for the bodywork) can under no circumstances be used to produce several cars at the same time, all without its initial quantity diminishing.

In addition, Company A indicated that for its value-added service offering (VAS_A), satellite-generated data were **combined** with other types of data, i.e., non-space data. It is represented by the arrow with red D_O in the figure. For example, the company delivers a solution to detect illegal fishing. We can imagine that this same service uses radar and optical images to track fishing activities, but also data on the regulation of these activities by the authorities (fishing quotas, etc.). Thus, we can assume that many space and non-space actors are involved in a value-added service in Earth observation, further complexifying the value chain representation within a single production period.
We have also introduced the influence of digital technologies on data and information flows within a production process in the downstream space segment. For Company A, this aspect is represented by the fact that arrow D_O is in bold in Figure 5.10. Indeed, the company answered ”Yes” to ”Do you use big data for your value-added

services?" This answer suggests that Company A combines large volumes of space and non-space data for value-added service.

A final aspect addressed by the survey is the position of the players in the chain. We sought to know whether the value-added services produced by respondents were aimed at individuals (BtoC), companies (BtoB), or public administrations or associations. If the company answered that its offering was aimed at businesses, we asked how its customer uses the service (as a decision aid or to enrich its service offering). Company A replied that its offer was aimed at both companies (BtoB) and public administrations or associations. In addition, it reported that companies use its services as a decision-making aid. Company A's offer is personalized: it has indicated that its services require information about its customers' needs (red I_{CUST} in Figure 5.10).

Finally, I_A refers to the information integrated in the value-added service VAS_A produced by Company A. This information, as a conservative output flow, could be delivered to multiple clients simultaneously, for example to several client companies of A, in various customized versions (e.g., I_A, I''_A, etc.) each incorporating different I_{CUST} based on the specific needs of the clients.

5.5.2.2 Company B: Satellite navigation and location services company

The second case we analyze (Company B in the following) is a company operating in satellite positioning and navigation. Company B indicated in the survey that its total revenue share corresponding to downstream activities was between 10% and 50%. Therefore, it belongs to the fourth circle of players in our classification, a more distant circle than Company A.

Figure 5.11 illustrates Company B's data-based production process for a given period. The company is involved in two downstream business segments: it supplies software and equipment and delivers value-added services.

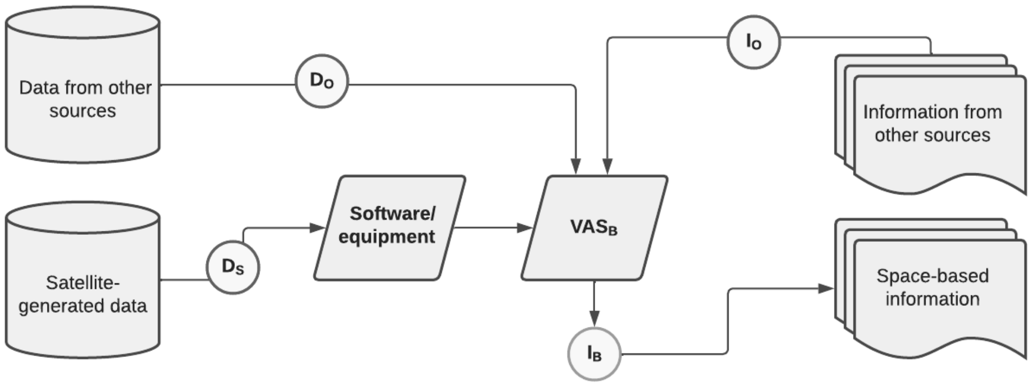


Figure 5.11: Company B: GNSS downstream space company

The supply of software and equipment is Company B's core business. It indicated that 80% of its downstream revenue corresponded to this activity, compared with 20% for value-added services.[53] When asked whether the software and equipment it supplies are dedicated solely to space data and signals, the company replied in the affirmative. We represent this information by the D_S arrow in red in Figure 5.11. This result is rather expected for receivers designed to receive satellite signals.

Additionally, Company B provides value-added services. Unlike its software and equipment, which are dedicated to receiving and processing GNSS signals, the company indicates combining these signals with non-space data (represented by arrow D_O in Figure 5.11). Hence, Company B's services integrate satellite signals with other data types to produce information denoted as I_B from the service VAS_B. Unlike the prior case, the company does not offer customized services, meaning they do not incorporate specific customer information.
Another distinction from Company A is that Company B's customers, primarily businesses, use its navigation services to **enrich** their own service offerings. Therefore, the production of information I_B does not signify the end of the value creation process. In a quest to identify downstream actors for sector assessment, it is indispensable to discern the relationships between the actors and to understand the nature of these connections. Ideally, we would have liked to identify Company B's business clients to incorporate them into our evaluation.

In this section, we aimed to illustrate and show the relevance of our analytical

[53]We have presented the results of total downstream revenues by activity segment in Chapter 4.

model for analyzing data and information dynamics in a production process with two company cases. We linked the nature of data and information as conservative flow factors and the multiple value chain configurations in the downstream segment. We could not study this aspect dynamically, as we only have data for one year. However, using company cases helped illustrate the complexity of defining a stable and unique value chain for the downstream space segment. The provision of downstream space information-based services involves different players, depending on the service's field of application, the user's needs, and the data combinations required to produce it.

5.5.3 Implications for measuring data value in value-added services

After discussing the role of data in the value chain configurations of the downstream space segment, we draw implications from our theoretical analysis regarding the economic value of data in downstream services. A central challenge of our assessment was to determine the portion of companies' revenues corresponding specifically to their downstream activities. To address this, we devised the downstream revenue indicator. Revenue measurement is commonly suitable for evaluating economic activities for producing standard industrial goods. It entails summing up the sales of each product or service of businesses, in our case, downstream space companies. However, our entire theoretical reflection revolved around the idea that data and information, specifically space-based data, signals, and information, do not possess the same properties as standard goods.

These distinct characteristics, particularly non-rivalry, non-excludability, and partial obsolescence of data and information, lead us to believe that the revenue indicator could be supplemented with a measure that acknowledges the peculiarities of data and information for value-added service providers. Let us initially consider the data acquisition cost as an indicator of its value. Take, for instance, a company providing natural resource management services using satellite imagery. This provider purchases a set of images once and can subsequently use this dataset for multiple services without repurchasing the images post their initial use. Now, think about a company that offers a mobile application leveraging its customers' geolocation data. While access to the geolocation signal is free, its significance in ensuring the service's performance is paramount. Similarly, questions arise concerning quantifying the volume of data integrated into a service to measure its importance. A service might incorporate a minimal amount of space data combined with other data types (as in the case of big data, for example). Yet, this data could be crucial in ensuring the service's effiency and performance.

We introduce a **Space Data Criticality Matrix** to address these inquiries. Figure 5.12 presents the matrix applied to the survey results presented in Section 5.5.1. This matrix relies on two features: the cost of accessing space infrastructure or acquiring data and the dependence of value-added services on space infrastructure or data.

Dependence on space infrastructure, data, or signals					
	Would not exist without them	4	6	5	4
	Significantly degraded without them	-	6	3	-
	Marginally degraded without them	-	-	1	-
		0%	Less than 10%	10-50%	More than 50%
		Cost of access to space infrastructure and data (% of total production costs)			

Figure 5.12: Criticality Matrix: Space data acquisition cost vs. dependence in downstream space services

The green zone of the matrix represents a low criticality of space data: it contains services of companies for which the acquisition cost of data is very low or zero and without which the quality and performance of the services would be marginally degraded. No companies responding to our survey belong to this zone.

The orange zone represents a medium criticality of space data. It includes services that strongly rely on space data for their quality and performance even though the acquisition cost is low **or** services with a significant share of space data acquisition costs in total costs, even if the service only marginally depends on space data. In the medium criticality zone, we find 6 respondents for whom the data acquisition cost is low (less than 10% of total costs) but whose services would be significantly degraded without this data. For 3 respondents, the cost of acquiring data is relatively high (10 to 50% of total costs), and their services would be significantly degraded without these data. Only one responding company indicated that even if the quality of its services depended marginally on space data, their acquisition accounted for between 10 and 50% of the total service production cost.

The last zone, shown in red in figure 5.12, represents the high criticality of space data. It corresponds to value-added services that would not exist without space data **and/or** services for which the acquisition costs of space data represent a very large share of the total production costs (more than 50%). One interesting survey result is that 4 companies claimed their services would not exist without space data or satellite signals even though their acquisition cost was zero. Similarly, 6 companies stated their services would not exist without data while their acquisition costs account for less than 10% of total costs. Another survey finding is that 4 companies offer services that wouldn't exist without space data, and the data acquisition cost is very high (more than 50% of total costs). 3 of these 4 companies deal in satellite capacity leasing/selling and/or data sales, which might explain this result.
Lastly, none of the companies responding to the survey indicated that their services were marginally or significantly dependent on space data while the data acquisition cost of data was over 50% of total costs. We do not have enough data to interpret this result, but finding companies in this part of the high criticality zone would be interesting. Indeed, we might think that if the acquisition of space data represents a significant expense for the company, its services would be highly dependent on the data in question.

We developed the Space Data Criticality Matrix to illustrate the singular nature of space data and signals integrated into value-added services. This matrix bridges the dependency of downstream space services on space data and signals with the proportion of their acquisition cost in the overall business expenses to deliver the service. With this criticality metric, we aim to provide a richer assessment of the downstream space sector by recognizing that the significance of space data and signals within a service cannot exclusively be understood through a quantitative measure of value (e.g., quantity, acquisition cost). Indeed, in the digital economy, access to data is cheap or even free with the open data trend. Yet, data stands as a central economic resource in this new paradigm. This is also true for the space sector, exemplified by the European program Copernicus which offers access to Earth observation imagery and services free of charge.

5.6 Conclusion

In this chapter, we conducted a theoretical exploration of the nature of goods and services produced in the downstream space sector. Starting with the observation that downstream activities generate value by producing, processing, and exploiting data and information delivered or transmitted through satellites, we highlighted that the economic assessment of this segment could be enriched by considering the fundamental properties of data and information as economic goods. Consequently, we leaned on the main concepts of information economics to describe these goods. We examined which properties of information also applied to data and discussed how the new techno-economic paradigm dominated by digital technologies amplifies some of these properties. Then, we adopted a dynamic approach to characterize data and information in value creation processes.

We elaborated an analytical framework adapted from the fund-flow approach to capture the dynamics of data and information in the production process. We defined these goods as particular flow factors. Data acts as a *conservative input flow factor*; it can be used in multiple simultaneous or successive processes and combined with other datasets without diminishing. Furthermore, it may undergo temporary or contextual obsolescence but regain value in the future. Information, on the other hand, is a *conservative output flow.* It is generated from a combination of data and various production factors, and its non-rivalrous property allows multiple agents to use it without any decrease in quantity. Information can be re-encoded as data and repurposed in production processes, particularly with digital technologies, while retaining its initial form in the company's information stock.

This theoretical development aims to supplement our assessment tool for the downstream space segment and to provide insights for future evaluations. It is interesting to examine the transformation of the downstream space sector in light of the dynamics of data and information. Thus, we've illustrated our theoretical intuitions using our survey of downstream space companies providing value-added software, equipment, and services. We posed questions about space data and signals, such as how they use data (in combinations, volume, and customization with client data) and how their customers employ space-based information. From these findings, we drew implications concerning the evolution of value chain configurations and assessing the importance of data in value-added services.

Finally, we formulate avenues to explore in future research. At the beginning of this chapter, we defined knowledge as a concept linked to data and information. However, as our discussion progressed, we concentrated on data and information. This was primarily because we aimed to analyze specifically the value-creation process wherein data are exploited to produce information. Moreover, since we considered the evolution of data dynamics in the digital paradigm, we emphasized that digital technologies produce, process, and leverage vast amounts of data to yield information. Yet, it would be of great interest for future analyses to integrate the notion of knowledge, which plays a pivotal role in value creation.
Secondly, for a more comprehensive and dynamic understanding of how data and information – especially in the context of the converging digital and space sector– modify the configurations of the downstream value chain, we advocate for a longitudinal approach. The proposition is to select a limited sample of downstream space companies and monitor the possible evolution of their value-creation processes over multiple periods. This method would ensure that if any shift occurs in the sector, it can be identified.

5.7 Appendix: Survey questionnaire (Second part, English translation)

*Are the software and/or equipment you provide solely dedicated to the data, signals, or information provided by satellites?[1]

Yes No

*To what extent are the **value-added services** you offer critically dependent on space infrastructure, data or signals?[2]

- The **value-added services** you offer would not exist without these infrastructures, data or space signals.
- The quality and/or performance of the **value-added services** you offer would be significantly degraded without them.
- The quality and/or performance of the **value-added services** you offer would be marginally degraded without them.
- Declines to answer

*What percentage does the cost of access to infrastructures and/or the cost of acquiring space data or signals represent in the total production costs of your **value-added services**?

- None, the data and/or signals are freely accessible
- Less than 10%
- Between 10% and 30%
- Between 30% and 50%
- More than 50%
- Declines to answer

* Are space data and/or signals combined with other types of data for your **value-added service** offering?

Yes No

*Do you use big data for your **value-added services**?

Yes No

[1]This question is displayed only if the respondent checks ‘provision of software and/or equipment […]’ to the question on the downstream activities conducted.
[2] The rest of the questionnaire is displayed if the respondent checks 'Value-added services dependent on space infrastructures, data or signals' to the question on the downstream activities conducted.

*Your **value-added service** offering is primarily targeted at:
Check one or more responses

- ☐ Businesses (BtoB)
- ☐ Individuals (BtoC)
- ☐ Public administrations or associations (ministries, local authorities, NGOs, ...)
- ☐ Declines to answer

*Do you know how your customers use your **value-added services**?[3]

- o As a decision-making aid
- o To enrich their service offering
- o Declines to answer

*Is your **value-added service** offering a customized solution (requiring information about your clients' needs)?

- o Yes
- o No
- o Declines to answer

[3] This question is displayed if the respondent checks « Businesses (BtoB) » to the previous question.

General Conclusion

The principal objective of this dissertation was to develop a tool for evaluating the size of the downstream space sector. This led us first to question how to identify downstream space companies and, consequently, the nature of the activities conducted in this sector. A fundamental methodological challenge has occupied us throughout this work: how to represent the downstream space value chain? In other words, what are the successive activities in downstream space, from satellite operations in orbit to providing value-added services to end users? Due to structural changes in the space industry (Chapter 1), we have observed difficulties defining the boundary between upstream and downstream segments, with downstream activities increasingly encroaching on upstream activities. These same changes, linked to the development of digital technology, are extending the downstream value chain with the involvement of a growing number of actors and the multiplication of commercial space applications for civil use (Chapter 2).

The exploitation of satellite infrastructures extends to an increasing number of sectors that are ostensibly distant from the space domain. For example, in precision agriculture, satellite imagery and satellite navigation are combined with other technologies (drones, sensors) for optimal management of cultivated land [Elijah et al., 2018, Delgado et al., 2019]. Should we, therefore, consider the provision of precision agriculture services as a downstream space activity? This question has arisen for many activities, notably those that use their customers' geopositioning information. To this are added specificities inherent to each application domain. For instance, satellite navigation services do not involve operations on satellites, unlike telecommunications and Earth observation. Geopositioning information is calculated without an intermediary from the distance between the user's receiver and the satellites in orbit.

With this in mind, we have adopted an exploratory approach to evaluate the downstream sector. The proposed method allows for detecting companies based on their citation context in a large volume of texts from the French-speaking press (Chapter 3). The companies correspond to a specific description of downstream activities determined by the expert in charge of applying the method. The latter builds or updates a query encompassing all known downstream space activities as much as possible at time t. This particular aspect of the method addresses the issue of potential structural changes in the space sector, as the keywords and phrases used in the query can be modified by experts over time. If the query remains unchanged at $t + 1$, the application of the method will identify newly created or cited downstream companies involved in activities stated in the query. If a significant structural change in downstream activities is observed when reproducing the method, it is possible to adapt the query by adding keywords that account for the newly emerged downstream activities. Consider the case of Earth observation. With the development of large-scale data management tools and facilitated access to satellite images, digital players offer new so-called 'intermediary' services for storing, processing, and disseminating Earth observation data [OECD, 2019]. If these activities are observed in newspapers, they can be easily integrated into the downstream query of the method to detect companies offering this type of service in Earth observation.

After the initial implementation of the method, we obtain a set of companies belonging to the downstream space sector in France. **The main contribution of this chapter is collecting a set of 334 companies with downstream space activities. Using our identification method, we detected 88 companies, representing over a quarter of the total downstream company set.** The first attempt demonstrated the method's ability to retrieve already known downstream companies and, more importantly, identify new ones unknown to sector experts.

Once the downstream actors identified, we continued the evaluation by considering the companies' revenue as one of the measures of value generated in the sector (Chapter 4). Compared to existing evaluations, one notable contribution of our work is the proposal of a downstream revenue indicator, i.e., the sector's sales that exclusively correspond to the companies' downstream activities. This indicator addresses questions regarding the measurement of value in this sector and contributes to the discussion on the measurement scope. By surveying companies regarding the portion of their activity that directly depends on satellite exploitation, space data, or signals, or the provision of equipment and software dedicated to satellite exploitation,

space data or signals, we outline an analysis in terms of *circles of actors*. If most of a company's total revenue relies on its downstream space activities, it belongs to the sector's first circle. Conversely, if the revenues generated by a company's downstream activities represent a small portion of its total revenue then it belongs to a more distant circle in the downstream sector.
From the data gathered from the 51 survey participants, we estimated the total downstream space revenue of the responding companies to be between 1.45 billion euros and 1.74 billion euros in 2021. Subsequently, we employed imputation methods to categorize the non-responding companies into a circle of actors, assigning them a share of the total revenue corresponding to downstream activities. After this imputation process, **we estimated the total downstream revenue for France in 2021 to range between 3.68 billion euros and 6.9 billion euros.**

The evaluation of the downstream sector presented in the fourth chapter of the thesis builds on the previously developed methodological framework. We aimed to test the reliability of our tool and achieve an estimation based on commonly accepted representations of the downstream sector, ensuring it can be easily interpreted and applied. Therefore, we deliberately narrowed the evaluation to a specific scope of activities. In the company identification phase, we formulated the query and selected sectors of activities based on similarity rules derived from known downstream actors. In the survey of downstream companies, we provided respondents with a traditional definition of the downstream sector and asked them to position their activities within a relatively strict segmentation, including sales and leasing of capacities, satellite operations, provision of software and equipment, and value-added services. These empirical choices enabled us to identify a group of companies engaged in activities related to exploiting space systems, satellite imagery, or signals, allowing us to estimate the value generated by these activities. From a methodological perspective, we have achieved the objective of developing a tool capable of detecting commercial downstream space actors, estimating the size of this segment, and adapting to potential evolutions.

Finally, we questioned whether the proposed evaluation method could evolve by considering in the unique nature of the goods and services offered by these downstream space entities (Chapter 5). Downstream activities primarily involve the production, transmission, and valorization of information. They exploit satellite systems and their outputs to create value-added goods and services. Within the Earth observation segment, these outputs include information collected from satellite imagery about agricultural lands, climate, water levels, or urbanization. In the navigation domain, satellite positioning systems provide continuous information about the geographic position of objects and individuals and their evolution. In addition to generating space-based information, space infrastructures can disseminate or exchange various types of information. Satellite communication systems enable the transmission of TV programs, radio broadcasts, images, or messages through the link between orbiting satellites and ground stations [Maral et al., 2020]. When dealing with commercial space applications, we refer to goods and services that incorporate space or non-space data and information in combination with other space or non-space information. To delve deeper into the evaluation question, it is crucial to analyze how information as an economic good with singular properties influences the value creation process in the downstream space sector.

In Chapter 5, we started a theoretical exploration of data and information dynamics in the digital age. We delved into literature that characterizes information as a public good, highlighting its fundamental properties and then applying these to data, the primary resource of the digital economy.

Beyond deepening theoretical concepts, this chapter aimed to tackle the role of data and information as economic goods with unique attributes. Their significance is increasing in the current techno-economic paradigm dominated by digital technologies, particularly regarding value-creation processes. To address this, we elaborated an analytical framework using the principal elements of Nicholas Georgescu-Roegen's flow-fund model. We introduced a new category of factors representing data and information within the economic process. Specifically, we characterized data as a conservative input flow factor and information as a conservative output flow. From a dynamic viewpoint, we investigated the role of these specific factors in production. Unlike standard flow factors, data and information can be used across multiple simultaneous or successive production processes without their quantity diminishing. Consequently, economic processes involving data and information lead to evolving and multifaceted value chain configurations. Furthermore, we observed that this distinctive nature of data and information should be considered when assessing sectors that produce value-added services incorporating information, such as the downstream space sector.

Finally, we applied these theoretical insights to our subject of study: the downstream space sector. We drew upon the results from the second part of our survey, which focused on software and equipment suppliers and value-added services providers, to illustrate specific elements of our analytical framework. We conclude the chapter by discussing avenues for enriching our method in two main areas: considering the evolution of value chain configurations and employing a criticality indicator for space data within value-added services.

Areas for improvement and future expansions

In this section, we address the areas for improvement and the main prospects of our work. This thesis is primarily a methodological development intended for future evaluation of the downstream space sector.

Regarding the identification of actors (Chapter 3), the tool presented could be improved using artificial intelligence techniques. Indeed, even if part of it is automated, the proposed method is highly evaluator-dependent. It relies on the automatic application of rules defined by the expert in charge of the evaluation and a final sorting of companies that can be time-consuming. The rules must be adapted to each evaluation context, to the specificities of the evaluated country's downstream sector, and to potential changes in the sector. However, we have seen that the space sector is undergoing a paradigm shift, and the downstream segment is one of the main drivers of this change. This aspect implies that the evaluator must have precise knowledge of the challenges and transformations of the sector at the time of the method's application. Therefore, an improvement avenue for future reproduction would be to use the downstream company dictionary we have built (i.e., the set of 334 companies) to train the machine to identify company names in a text discussing downstream space activities. This would involve using the Named-entity recognition technique with a statistical approach rather than a rule-based one. This would eliminate the tedious step of updating the rules and manually sorting the companies identified in the text after the rules' application.
Another avenue for improvement regarding the identification method would be to identify links between downstream companies. To achieve this, we have already begun a co-citation analysis, checking if known companies were cited near a company present in the text. However, this was inconclusive as the text still had too much noise. Therefore, we recommend exploring this route to detect connections between the identified companies (competitors, customer suppliers, etc.) and to identify additional

companies to expand the identification perimeter to actors more distant from the downstream sector.

We also suggest areas for improvement regarding assessing the downstream space segment and the proposed indicators. As we pointed out in the conclusion of Chapter 4, it seems necessary to define ground segment mission services as a downstream space activity in future estimations to align with existing definitions [OECD, 2022b]. Next, we focused on the downstream turnover indicator for this initial measurement. An enhancement for future evaluations would be to integrate an employment indicator. Although it seems complicated to estimate the number of people employed in the downstream sector without considering all company employees, this measure is interesting to gain a deeper understanding of the sector. Finally, we recommend adding a question to the survey about their field of activity (e.g., climate, resource management, agriculture, transport, telecommunications services).

Conclusion générale

Dans cette thèse, nous nous sommes attachés à développer un outil méthodologique afin d'évaluer le secteur spatial aval. Cela a conduit à nous interroger dans un premier temps sur la manière d'identifier les entreprises spatiales en aval et, conséquemment, sur la nature des activités conduites dans ce secteur. Un enjeu méthodologique fondamental nous a occupés tout au long de ces travaux : comment représenter la chaine de valeur du spatial aval ? En d'autres termes, quelles sont les activités successives qui composent le secteur spatial aval depuis les opérations de satellites en orbite jusqu'à la fourniture de services à valeur ajoutée aux utilisateurs finaux ? Du fait d'évolutions structurelles dans l'industrie spatiale (Chapitre 1), nous avons observé des difficultés à définir la frontière entre segments amont et aval, les activités aval rognant de plus en plus les activités amont. Ces mêmes évolutions, non sans lien avec le développement du numérique, étendent la chaine de valeur en aval avec l'implication d'un nombre croissant d'acteurs et la multiplication des applications spatiales civiles (Chapitre 2).

L'exploitation des infrastructures satellitaires s'étend à un nombre accru de secteurs a priori éloignés du domaine spatial. Dans le domaine de l'agriculture de précision par exemple, l'imagerie satellite et la navigation par satellite sont combinées à d'autres technologies (drones, capteurs) pour la gestion optimale des terres cultivées [Elijah et al., 2018, Delgado et al., 2019]. Doit-on pour autant considérer la fourniture de services d'agriculture de précision comme une activité spatiale aval ? Ce questionnement s'est posé pour un nombre conséquent d'activités, notamment celles qui utilisent les informations de géo-positionnement de leurs clients. A cela s'ajoutent des spécificités inhérentes à chaque domaine d'application. A titre d'exemple, la fourniture de services de navigation par satellite n'implique pas, contrairement aux télécommunications et à l'observation de la Terre, d'opérations sur les satellites. Les informations de géo-positionnement sont calculées sans intermédiaire à partir de la distance entre le récepteur de l'utilisateur et les satellites en orbite.

Nous avons adopté une approche exploratoire d'évaluation du secteur aval qui tient compte de ces différents constats. La méthode proposée permet de détecter des entreprises à partir de leur contexte de citation dans un volume important de textes issus de la presse francophone (Chapitre 3). Ces entreprises répondent à une certaine description des activités aval déterminée par l'expert en charge de l'application de la méthode. Ce dernier construit ou actualise une requête qui englobe autant que faire se peut l'ensemble des activités spatiales aval connues au moment t. Cet aspect particulier de la méthode répond au problème de changements de structure potentiels du secteur spatial en ce sens que les mots clés et expressions utilisés dans la requête peuvent être modifiés par les experts dans le temps. Si la requête est inchangée en $t+1$, l'application de la méthode permettra d'identifier les entreprises aval nouvellement créées ou citées dans la presse impliquées dans des activités énoncées dans la requête. Si l'on observe un changement structurel important dans les activités aval au moment de la reproduction de la méthode, il est possible d'adapter la requête en y ajoutant des mots clés qui rendent compte des nouvelles activités aval apparues. Prenons le cas de l'observation de la Terre. Avec le développement des outils de gestion des données en grand volume et la facilitation de l'accès aux images satellitaires, les acteurs du numérique proposent de nouveaux services dits 'intermédiaires' pour stocker, traiter et diffuser les données d'observation de la Terre [OECD, 2019]. Ces activités, si elles sont constatées dans les journaux, sont aisément intégrables dans la requête aval de la méthode afin de détecter les entreprises qui proposent ce type de services dans l'observation de la Terre.

A l'issue de la première application de la méthode, nous obtenons un ensemble d'entreprises appartenant au secteur spatial aval en France. **Le résultat principal du Chapitre 3 est la collecte d'un ensemble de 334 entreprises avec des activités spatiales aval. En utilisant notre méthode d'identification, nous avons détecté 88 entreprises, soit plus d'un quart de l'ensemble des entreprises en aval.** Le premier essai a démontré la capacité de la méthode à retrouver des entreprises aval déjà connues et, plus important encore, à en identifier de nouvelles non connues par les experts du secteur.

Une fois les acteurs aval identifiés, nous avons poursuivi la démarche d'évaluation en considérant le chiffre d'affaires des entreprises comme l'une des mesures de la valeur générée dans le secteur (Chapitre 4). L'un des apports notoires de nos travaux par rapport aux évaluations existantes réside dans la proposition d'un indicateur de chiffre

d'affaires aval, c'est-à-dire les ventes du secteur qui correspondent exclusivement aux activités aval des entreprises. Cet indicateur apporte une réponse aux questionnements sur la mesure de la valeur dans ce secteur et alimente également la réflexion sur le périmètre de mesure. En sondant les entreprises sur la part de leur activité qui dépend directement de l'exploitation des satellites, données ou signaux spatiaux ou sur la fourniture d'équipements et logiciels dédiés à l'exploitation des satellites, données ou signaux satellitaires, nous esquissons une analyse en termes de *cercles d'acteurs*. Si la majorité du chiffre d'affaires total d'une entreprise repose sur ses activités spatiales en aval, alors on considère que cette entreprise appartient au premier cercle du secteur. A l'inverse, si les recettes générées par les activités aval d'une entreprise représentent une faible part de son chiffre d'affaires total, alors elle appartient à un cercle plus éloigné du secteur aval.
À partir des données recueillies auprès des 51 participants à l'enquête, nous avons estimé que le revenu spatial aval total des entreprises répondantes se situait entre 1,45 milliard d'euros et 1,74 milliard d'euros en 2021. Ensuite, nous avons utilisé des méthodes d'imputation pour estimer le chiffre d'affaires 2021 des entreprises non-répondantes et les classer dans un cercle d'acteurs, en leur attribuant une part du revenu total correspondant à leurs activités spatiales aval. Après ces processus d'imputation, **nous avons estimé que le revenu total en aval pour la France en 2021 se situe entre 3,68 milliards d'euros et 6,9 milliards d'euros.**

L'évaluation du secteur aval présentée dans la seconde partie de cette thèse s'inscrit dans la continuité de la construction méthodologique. Il s'agissait de tester la fiabilité de notre outil et de parvenir à une estimation interprétable à l'aune des représentations du secteur aval communément admises. A cet égard, plusieurs étapes de l'application circonscrivent délibérément l'évaluation à un périmètre restreint d'activités. Dans la phase d'identification des entreprises, nous avons appliqué des règles de similarité avec les caractéristiques des acteurs aval déjà connus dans la formulation de la requête ainsi que dans le choix des secteurs d'activités retenus. Dans l'enquête auprès des entreprises aval, nous avons proposé aux répondants une définition traditionnelle du secteur aval, et leur avons demandé de positionner leur activité par rapport à une segmentation relativement stricte du secteur : vente et location de capacités, opérations sur satellites, fourniture de logiciels et d'équipements et services à valeur ajoutée. Ces choix empiriques nous ont permis d'obtenir un premier socle d'entreprises impliquées de manière certaine dans des activités d'exploitation des systèmes spatiaux, des images ou signaux satellitaires et ainsi d'estimer la valeur générée par ces activités. D'un point de vue purement méthodologique, nous avons rempli l'objectif de développement d'un outil qui détecte les acteurs du secteur spatial aval, qui donne une estimation de ce segment et qui est capable de s'adapter aux évolutions auxquelles ce dernier pourrait être soumis.

Enfin, nous nous sommes demandés si la méthode d'évaluation proposée pouvait évoluer en tenant compte de la nature unique des biens et services délivrés par ces entités spatiales aval (chapitre 5). En effet, les activités aval reposent sur la production, la transmission et la valorisation de l'information. Elles consistent à exploiter les systèmes satellitaires et leurs outputs pour en faire des biens et services à valeur ajoutée. Parmi ces outputs, on trouve dans le segment de l'observation de la Terre des informations collectées à partir des images prises par les satellites sur les terres agricoles, le climat, le niveau des eaux ou l'urbanisation. Dans le domaine de la navigation, les systèmes de positionnement par satellite fournissent des informations en continu sur la position géographique des objets et individus et l'évolution de cette position. En plus de générer de l'information, les infrastructures spatiales peuvent également avoir pour fonction de diffuser ou d'échanger des informations de toute nature. Les systèmes de communication par satellites permettent de transmettre des programmes TV, des émissions de radio, des images ou des messages grâce à la liaison entre satellites en orbite et stations terrestres [Maral et al., 2020]. Lorsque l'on traite des applications spatiales commerciales, on se réfère à des biens et services qui incorporent de l'information spatiale ou non spatiale, en combinaison avec d'autres informations d'origine spatiale ou non spatiale. Pour approfondir la question de l'évaluation, il nous paraît donc fondamental d'entamer une réflexion sur la manière dont l'information, en tant que bien économique aux propriétés singulières, influence le processus de création de valeur dans le secteur spatial aval.

Dans le chapitre 5, nous avons entamé une exploration théorique de la dynamique des données et de l'information à l'ère numérique. Nous nous sommes plongés dans la littérature qui caractérise l'information comme un bien public, en soulignant ses propriétés fondamentales et en les appliquant ensuite aux données, la ressource principale de l'économie numérique.

Au-delà de l'approfondissement des concepts théoriques, ce chapitre visait à aborder le rôle des données et de l'information en tant que biens économiques dotés de propriétés singulières. Leur importance augmente dans le paradigme techno-économique actuel dominé par les technologies numériques, particulièrement en ce qui concerne les processus de création de valeur. Pour ce faire, nous avons élaboré un cadre analytique en utilisant les éléments clés du modèle flux-fonds de Nicholas Georgescu-Roegen. Nous avons introduit une nouvelle catégorie de facteurs représentant les données et l'information dans le processus économique. Plus précisément, nous avons caractérisé les données comme un facteur flux conservatif entrant et l'information comme un facteur flux conservatif sortant. D'un point de vue dynamique, nous avons étudié le rôle de ces facteurs spécifiques dans la production. Contrairement aux facteurs flux standards, les données et l'information peuvent être utilisées dans plusieurs processus de production simultanés ou successifs sans que leur quantité ne diminue. Par conséquent, les processus économiques impliquant des données et de l'information conduisent à des configurations de chaînes de valeur évolutives et multiformes. En outre, nous avons observé que cette nature distinctive des données et de l'information devrait être prise en compte lors de l'évaluation des secteurs qui produisent des services à valeur ajoutée incorporant des informations, tels que le secteur spatial en aval.

Enfin, nous avons appliqué ce cadre analytique à notre sujet d'étude : le secteur spatial en aval. Nous nous sommes appuyés sur les résultats de la deuxième partie de notre enquête, qui portait sur les fournisseurs de logiciels et d'équipements et sur les fournisseurs de services à valeur ajoutée, pour illustrer des éléments spécifiques de notre cadre analytique. Nous concluons le chapitre en discutant des possibilités d'enrichissement de notre méthode sous deux aspects : la prise en compte de l'évolution des configurations de la chaîne de valeur et l'utilisation d'un indicateur de criticité pour les données spatiales au sein des services à valeur ajoutée.

Pistes d'amélioration et extensions futures

Dans cette section, nous abordons les axes d'amélioration et les principales perspectives d'enrichissement de notre travail. Cette thèse est avant tout un développement méthodologique destiné à l'évaluation du secteur spatial aval à des intervalles réguliers.

En ce qui concerne l'identification des acteurs (Chapitre 3), l'outil présenté pourrait être amélioré à l'aide de techniques d'intelligence artificielle. En effet, même si elle est en partie automatisée, la méthode proposée est fortement dépendante de l'évaluateur. Elle repose sur l'application automatique de règles définies par l'expert en charge de l'évaluation et un tri final des entreprises qui peut s'avérer chronophage. Les règles doivent être adaptées à chaque contexte d'évaluation, aux spécificités du secteur

aval du pays évalué et aux évolutions potentielles du secteur. Or, nous avons vu que le secteur spatial est soumis à un changement de paradigme et que le segment aval est l'un des principaux moteurs de ce changement. Cet aspect implique que l'évaluateur ait une connaissance précise des défis et des transformations du secteur au moment de l'application de la méthode. Par conséquent, une piste d'amélioration pour la reproduction future serait d'utiliser le dictionnaire des entreprises en aval que nous avons construit (c'est-à-dire l'ensemble des 334 entreprises) pour entraîner l'algorithme à identifier les noms d'entreprises dans un texte traitant des activités spatiales en aval. Cela impliquerait d'utiliser la technique de reconnaissance des entités nommées avec une approche statistique plutôt qu'une approche basée sur des règles. Cela permettrait d'éviter l'étape fastidieuse de la mise à jour des règles et du tri manuel des entreprises identifiées dans le texte après l'application des règles.
Une autre piste d'amélioration de la méthode d'identification serait d'identifier les liens entre les entreprises spatiales aval. Pour ce faire, nous avons déjà entamé une analyse des co-citations, en vérifiant si des entreprises connues étaient citées à proximité d'une entreprise présente dans le texte. Cependant, cette analyse n'a pas été concluante car le texte contenait encore trop de bruit. Nous recommandons donc d'explorer cette voie pour détecter des connexions entre les entreprises identifiées (concurrents, clients fournisseurs, etc.) et d'identifier des entreprises supplémentaires pour élargir le périmètre d'identification à des acteurs plus éloignés de la filière aval.

Nous suggérons également des pistes d'amélioration concernant l'évaluation du segment spatial en aval et les indicateurs proposés. Comme nous l'avons souligné dans la conclusion du chapitre 4, il semble nécessaire de définir les services de mission du segment terrestre comme une activité spatiale en aval dans les estimations futures afin de s'aligner sur les définitions existantes [OECD, 2022b]. Ensuite, nous nous sommes concentrés sur l'indicateur de chiffre d'affaires en aval pour cette première application. Une amélioration pour les évaluations futures serait d'intégrer un indicateur d'emploi. Bien qu'il semble compliqué d'estimer le nombre de personnes employées dans le secteur en aval sans prendre en compte l'emploi total, cette mesure est intéressante pour mieux comprendre le secteur. Enfin, nous recommandons d'ajouter une question à l'enquête sur leur domaine d'activité (e.g., climat, gestion des ressources, agriculture, transport, services de télécommunications).

Bibliography

R. L. Ackoff. From data to wisdom. *Journal of applied systems analysis*, 16(1):3–9, 1989.

H. Akbar, Y. Baruch, and N. Tzokas. Feedback loops as dynamic processes of organizational knowledge creation in the context of the innovations' front-end. *British Journal of Management*, 29(3):445–463, 2018.

G. A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics*, 84(3):488–500, 1970.

A. A. Al-Hassan, F. Alshameri, and E. H. Sibley. A research case study: Difficulties and recommendations when using a textual data mining tool. *Information & Management*, 50(7):540–552, 2013. ISSN 0378-7206. doi: https://doi.org/10.1016/j.im.2013.05.010. URL `https://www.sciencedirect.com/science/article/pii/S037872061300061X`.

M. Al-Mekhlal and A. A. Khwaja. A synthesis of big data definition and characteristics. *2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, pages 314–322, 2019.

E. G. Anderson Jr, G. G. Parker, and B. Tan. Platform performance investment in the presence of network externalities. *Information Systems Research*, 25(1):152–172, 2014.

W. Antweiler and M. Z. Frank. Is all that talk just noise? the information content of internet stock message boards. *The Journal of Finance*, 59(3):1259–1294, 2004. doi: https://doi.org/10.1111/j.1540-6261.2004.00662.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2004.00662.x`.

S. K. Arora, J. Youtie, P. Shapira, L. Gao, and T. Ma. Entry strategies in an emerging technology: a pilot web-based study of graphene firms. *Scientometrics*, 95:1189–1207, 2013.

K. J. Arrow. *Economic Welfare and the Allocation of Resources for Invention*, pages 609–625. Princeton University Press: Princeton, nelson, richard r. edition, 1962.

K. J. Arrow. The economics of information: An exposition. *Empirica*, 23(2):119–128, 1996.

Australian Space Agency. *Advancing Space: Australian Civil Space Strategy 2019-2028*. Australian Space Agency, 2019.

Australian Space Agency. *Economic snapshot of the Australian space sector: 2016-17 to 2018-19*. Australian Space Agency, 2021.

L. Bach, P. Cohendet, G. Lambert, and M. J. Ledoux. Measuring and Managing Spinoffs: The Case of the Spinoffs Generated by ESA Programs. In J. S. Greenberg and H. R. Hertzfeld, editors, *Space economics*, pages 171–206. American Institute of Aeronautics and Astronautics, 1992.

A. Bajaj and S. R. Nidumolu. A feedback model to understand information system usage. *Information & management*, 33(4):213–224, 1998.

P. Barbaroux. The metamorphosis of the world space economy: investigating global trends and national differences among major space nations' market structure. *Journal of Innovation Economics Management*, 2(20):9–35, 2016.

P. Barbaroux and V. dos Santos Paulino. Le rôle de la défense dans l'émergence d'une nouvelle industrie: le cas de l'industrie spatiale. *Innovations, Revue d'Economie et de Management de l'Innovation*, 2(42):39–58, 2013.

M. J. Bates. Information behavior. *Encyclopedia of library and information sciences*, 3:2381–2391, 2010.

G. Bateson. The logical categories of learning and communication. *Steps to an Ecology of Mind*, pages 279–308, 1972.

P. Besha and A. MacDonald. *Economic Development of Low-Earth Orbit*. Patrick Besha and Alexander MacDonald, NASA Headquarters - 300 E Street SW, Washington, DC 20546, 2016.

S. C. Bhardwaj, S. Shekhar, A. Vidyarthi, and R. Prakash. Satellite navigation and sources of errors in positioning: A review. *2020 International Conference on Advances in Computing, Communication & Materials (ICACCM)*, pages 43–50, 2020.

J.-M. Bockel. *The future of the space industry.* NATO Parliamentary Assembly Brussels, Belgium, 2018.

M. Boisot and A. Canals. Data, information and knowledge: have we got it right? *Journal of evolutionary economics*, 14:43–67, 2004.

A. Bondiou-Clergerie. Les chiffres clés de l'industrie spatiale française. In *Annales des Mines-Realites industrielles*, number 2, pages 38–43. FFE, 2019.

Booz & Company. *Evaluation of socio-economic impacts from space activities in the EU.* EU publications, 2014. URL https://op.europa.eu/en/publication-detail/-/publication/b3c64cf6-3caa-4f46-b6cc-a69c3b583cc5.

D. L. Bosworth. The rate of obsolescence of technical knowledge–a note. *The Journal of Industrial Economics*, pages 273–279, 1978.

J. Bouma, H. van der Woerd, and O. Kuik. Assessing the value of information for water quality management in the north sea. *Journal of Environmental Management*, 90(2):1280–1288, 2009. ISSN 0301-4797. doi: https://doi.org/10.1016/j.jenvman.2008.07.016. URL https://www.sciencedirect.com/science/article/pii/S0301479708001977.

Bryce Space & Technology. *Start-Up Space 2018: Update on Investment in Commercial Space Ventures.* Bryce Space & Technology, Washington DC, 2019a. URL https://brycetech.com/reports/report-documents/Bryce_Start_Up_Space_2019.pdf.

Bryce Space & Technology. *2019 State of the Satellite Industry Report.* Number 22. Prepared for the Satellite Industry Association, Washington DC, 2019b. URL https://brycetech.com/reports.

Bryce Space & Technology. *Start-Up Space 2022: Update on Investment in Commercial Space Ventures.* Bryce Space & Technology, Washington DC, 2022.

BryceTech. 2020 State of the Satellite Industry Report. Technical report, 2021.

BryceTech. 2021 State of the Satellite Industry Report. Technical report, 2022.

E. Brynjolfsson and A. McAfee. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company, 2014.

R. Bukht and R. Heeks. Defining, Conceptualising and Measuring the Digital Economy. In *International Organisations Research Journal*, 2017.

L. Bybee, B. T. Kelly, A. Manela, and D. Xiu. Business news and business cycles. Technical report, National Bureau of Economic Research, 2021.

F. Calvino and C. Criscuolo. Business dynamics and digitalisation. *OECD Science, Technology and Industry Policy Papers*, (62), 2019. doi: https://doi.org/10.1787/23074957.

B. Canis. *Commercial Space Industry Launches a New Phase*. CRS Report. Congressional Research Service, 2016.

Y. Carrière-Swallow and V. Haksar. The Economics and Implications of Data: an Integrated Perspective. *International Monetary Fund*, 2019.

L. Chiticariu, R. Krishnamurthy, Y. Li, F. Reiss, and S. Vaithyanathan. Domain adaptation of rule-based annotators for named-entity recognition tasks. In *Proceedings of the 2010 conference on empirical methods in natural language processing*, pages 1002–1012, 2010.

C. Christensen. *The innovator's dilemma: When new technologies cause great firms to fail*. Harvard Business Review Press, 2013.

D. Ciuriak. *The Economics of Data: Implications for the Data-driven Economy*. Center for International Governance Innovation, 2018.

A. Coad, A. Segarra, and M. Teruel. Innovation and firm growth: Does firm age play a role? *Research Policy*, 45(2):387–400, 2016. ISSN 0048-7333. doi: https://doi.org/10.1016/j.respol.2015.10.015. URL `https://www.sciencedirect.com/science/article/pii/S0048733315001687`.

P. Cohendet. Evaluating the industrial indirect effects of technology programmes: the case of the european space agency (esa) programmes. In *Proceedings of the OECD Conference "Policy Evaluation in Innovation and Technology", BETA,* *Université Louis Pasteur, Strasbourg, France. Available on-line http://www. oecd. org/dataoecd/3/37/1822844. pdf.* Citeseer, 1997.

P. Cohendet and P. Llerena. La conception de la firme comme processeur de connaissances. *Revue d'économie industrielle*, 88(1):211–235, 1999.

Q. Couix. Georgescu-roegen's flow-fund theory of production in retrospect. *Ecological Economics*, 176:106749, 2020.

Council of European Union. Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector Information, 2019.
`http://data.europa.eu/eli/dir/2019/1024/oj`.

CSA. *2020 State of the Canadian Space Sector Report: Facts and Figures 2019*. Canadian Space Agency, Saint-Hubert, 2020. URL `https://www.asc-csa.gc.ca/eng/publications/2020-state-canadian-space-sector-facts-figures-2019.asp`.

T. H. Davenport, L. Prusak, et al. *Working knowledge: How organizations manage what they know*. Harvard Business Press, 1998.

T. H. Davenport, J. G. Harris, D. W. De Long, and A. L. Jacobson. Data to knowledge to results: building an analytic capability. *California management review*, 43(2): 117–138, 2001.

K. Davidian. Definition of NewSpace. *New Space*, 8(2), 2020. doi: http://doi.org/10.1089/space.2020.29027.kda.

V. Dedieu, G. Fioraso, and L. Ménétrier. *L'ouverture comme réponse aux défis de la filière spatiale*. Ministère de l'Enseignement Supérieur et de la Recherche, 2016.

J. A. Delgado, N. M. Short, D. P. Roberts, and B. Vandenberg. Big Data Analysis for Sustainable Agriculture on a Geospatial Cloud Framework. *Frontiers in Sustainable Food Systems*, 3, 2019. ISSN 2571-581X. doi: 10.3389/fsufs.2019.00054. URL `https://www.frontiersin.org/articles/10.3389/fsufs.2019.00054`.

L. Delponte, J. Pellegrin, E. Sirtori, M. Gianinetto, and L. Boschetti. *Space market uptake in Europe*. European Parliament, 2016.

G. Denis, D. Alary, X. Pasco, N. Pisot, D. Texier, and S. Toulza. From new space to big space: How commercial space dream is becoming a reality. *Acta Astronautica*, 166:431–443, 2020. ISSN 0094-5765. doi: https://doi.org/10.1016/j.actaastro.2019.08.031. URL `https://www.sciencedirect.com/science/article/pii/S0094576519313451`.

G. Dosi, F. Malerba, G. B. Ramello, and F. Silva. Information, appropriability, and the generation of innovative knowledge four decades after arrow and nelson: an introduction. *Industrial and Corporate Change*, 15(6):891–901, 2006.

Y. Dou and T. Liu. Pricing the information goods in the presence of product depreciation and word-of-mouth effects. In *2013 10th International Conference on Service Systems and Service Management*, pages 142–148. IEEE, 2013.

P. Dulbecco and P. Garrouste. Théorie de la dynamique économique: une réévaluation de la tentative de nicholas georgescu-roegen. *Recherches Économiques de Louvain/Louvain Economic Review*, 70(1):5–29, 2004.

EARSC. *A Survey into the State & Health of the European EO Services Industry*. EARSC publications, Brussels, Belgium, 2013.

EARSC. *A Survey into the State & Health of the European EO Services Industry*. EARSC publications, Brussels, Belgium, 2017.

EARSC. *A Survey into the State & Health of the European EO Services Industry*. EARSC publications, Brussels, Belgium, 2019.

EARSC. *EARSC Industry Survey 2021*. EARSC publications, Brussels, Belgium, 2021.

R. Ehsani and F. Drabløs. Robust distance measures for k nn classification of cancer data. *Cancer informatics*, 19:1176935120965542, 2020.

L. Einav and J. Levin. The Data Revolution and Economic Analysis. *Innovation Policy and the Economy*, 14:1–24, 2014. doi: 10.1086/674019. URL `https://doi.org/10.1086/674019`.

O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and M. N. Hindia. An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges. *IEEE Internet of Things Journal*, 5(5):3758–3773, 2018. doi: 10.1109/JIOT.2018.2844296.

Y. S. Eom and J. H. Hong. Measuring the economic benefits of an environmental monitoring satellite project: The value of information approach. *Space Policy*, 29 (3):203–209, 2013. ISSN 0265-9646. doi: https://doi.org/10.1016/j.spacepol.2013.06.003. URL https://www.sciencedirect.com/science/article/pii/S0265964613000532.

EPO and EUIPO. IPR-intensive industries and economic performance in the European Union. Report, European Patent Office (EPO) and European Union Intellectual Property Office (EUIPO), 2022.

ESA. *Rapport intermédiaire sur l'économie de l'espace*. European Space Agency, Paris, France, 2016.

ESPI. *Space Policies, Issues and Trends in 2017-2018*. Number 65. European Space Policy Institute, Vienna, Austria, 2018.

ESPI. *Space Venture Europe 2018: Entrepreneurship and Private Investment in the European Space Sector*. Number 67. European Space Policy Institute, Vienna, Austria, 2019.

ESPI. *Space Venture Europe 2019: Entrepreneurship and Private Investment in the European Space Sector*. Number 73. European Space Policy Institute, Vienna, Austria, 2020.

ESPI. *Space Venture Europe 2020: Entrepreneurship and Private Investment in the European Space Sector*. Number 78. European Space Policy Institute, Vienna, Austria, 2021.

Eumetsat. The case for EPS/Metop Second-Generation: Cost Benefit Analysis. Technical report, EUMETSAT, 2014. URL https://www-cdn.eumetsat.int/files/2020-04/pdf_report_eps-sg_cost-benefit.pdf.

Euroconsult. Space Economy Report, 8th Edition. Technical report, Euroconsult, 2022.

European Commission. Eu space industrial policy: Releasing the potential for economic growth in the space sector. Technical report, European Commission, Brussels, Belgium, 2013.

European GNSS Agency. *GNSS Market Report 2015*. Number 4. Publications Office of the EU, 2015. doi: 10.2878/251572.

European GNSS Agency. *GNSS Market Report 2017*. Number 5. Publications Office of the EU, 2017. doi: 10.2878/0426.

European GNSS Space Agency. *GNSS Market Report 2010*. Number 1. Publications Office of the EU, 2010.

Eurospace. *Facts and Figures: the European Space Industry in 2019*. Number 24. ASD Eurospace publications, 2020.

EUSPA. *EUSPA EO and GNSS market report*. Number 1. European Union Agency for the Space Programme, 2022. doi: 10.2878/94903. URL https://www.euspa.europa.eu/european-space/euspace-market/gnss-market/eo-gnss-market-report.

B. Evans, M. Werner, E. Lutz, M. Bousquet, G. Corazza, G. Maral, and R. Rumeau. Integration of Satellite and Terrestrial Systems in Future Multimedia Communications. *IEEE Wireless Communications*, 12(5):72–80, 2005. doi: 10.1109/MWC.2005.1522108.

M. K. Evans. The economic impact of nasa r and d spending. Technical report, NASA, 1976.

C. Feijóo, J.-L. Gómez-Barroso, and S. Aggarwal. *Chapter 25: Economics of Big Data*, chapter 25, pages 510–526. Edward Elgar Publishing, Cheltenham, UK, 2016. ISBN 9780857939845. doi: 10.4337/9780857939852.00034. URL https://www.elgaronline.com/view/edcoll/9780857939845/9780857939845.00034.xml.

S. Ferrari. 8. éthique et bioéconomie chez nicholas georgescu-roegen. *Cahiers d'économie politique*, (1):213–242, 2021.

M. Florio, P. Castelnovo, V. Lupi, V. Morretta, D. Vurchio, L. Zirulia, S. Di Ciaccio, and M. Piermaria. The Socio-Economic Impact of Public Policies in the Space Sector in Italy. *Working Paper CIRIEC*, (01), 2022.

M. Fransman. Information, knowledge vision and theories of the firm. *Industrial and corporate change*, 3(3):713–757, 1994.

N. Frischauf, R. Horn, T. Kauerhoff, M. Wittig, I. Baumann, E. Pellander, and O. Koudelka. NewSpace: New Business Models at the Interface of Space and Digital Economy: Chances in an Interconnected World. *New Space*, 6(2):135–146, 2018. doi: 10.1089/space.2017.0028. URL https://doi.org/10.1089/space.2017.0028.

M. Gentzkow, B. Kelly, and M. Taddy. Text as data. *Journal of Economic Literature*, 57(3):535–74, 2019.

N. Georgescu-Roegen. The economics of production. *The American Economic Review*, 60(2):1–9, 1970.

N. Georgescu-Roegen. *The entropy law and the economic process*. Harvard university press, 1971.

N. Georgescu-Roegen. Feasible recipes versus viable technologies. *Atlantic Economic Journal*, 12:21–31, 1984.

C. Giannopapa, A. Staveris-Poykalas, and S. Metallinos. Space as an enabler for sustainable digital transformation: The new space race and benefits for newcommers. *Acta Astronautica*, 2022. ISSN 0094-5765. doi: https://doi.org/10.1016/j.actaastro.2022.06.005. URL `https://www.sciencedirect.com/science/article/pii/S0094576522002843`.

A. Gök, A. Waterworth, and P. Shapira. Use of web mining in studying innovation. *Scientometrics*, 102:653–671, 2015.

A. Goldfarb and C. Tucker. Digital Economics. *NBER Working Paper Series*, 2017.

P. J. Gorinski, H. Wu, C. Grover, R. Tobin, C. Talbot, H. Whalley, C. Sudlow, W. Whiteley, and B. Alex. Named entity recognition for electronic health records: a comparison of rule-based and machine learning approaches. *arXiv preprint arXiv:1903.03985*, 2019.

S. V. Gudmundsson. Blue origin: Riding the wave of disruption in the space industry. *Available at SSRN 3179135*, 2018.

A. Hagiu and J. Wright. Multi-sided platforms. *International Journal of Industrial Organization*, 43:162–174, 2015. ISSN 0167-7187. doi: https://doi.org/10.1016/j.ijindorg.2015.03.003. URL `https://www.sciencedirect.com/science/article/pii/S0167718715000363`.

R. Harris and I. Baumann. Open data policies and satellite earth observation. *Space Policy*, 32:44–53, 2015. ISSN 0265-9646. doi: https://doi.org/10.1016/j.spacepol.2015.01.001. URL `https://www.sciencedirect.com/science/article/pii/S0265964615000028`.

R. Harris and L. Miller. Earth observation and the public good. *Space Policy*, 27(4): 194–201, 2011.

H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, and M. R. Yeganegi. Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1):1, 2020.

A. Hein and C. Bruce Rosete. Space-as-a-service: A framework and taxonomy of-as-a-service concepts for space. In *Proceedings of the International Astronautical Congress 2022*, 2022.

L. Heracleous, D. Terrier, and S. Gonzalez. Nasa's capability evolution toward commercial space. *Space Policy*, 50:101330, 2019. ISSN 0265-9646. doi: https://doi.org/10.1016/j.spacepol.2019.07.004. URL `https://www.sciencedirect.com/science/article/pii/S0265964619300839`.

H. R. Hertzfeld. *Space economic data*. US Department of Commerce, Office of Space Commercialization, 2002.

A. J. Hey, S. Tansley, K. M. Tolle, et al. *The fourth paradigm: data-intensive scientific discovery*, volume 1. Microsoft research Redmond, WA, 2009.

T. Highfill, A. Jouard, and C. Franks. Preliminary estimates of the u.s. space economy, 2012–2018. *Survey of Current Business*, 100(12), 2020.

T. C. Highfill and A. C. MacDonald. Estimating the united states space economy using input-output frameworks. *Space Policy*, page 101474, 2022. ISSN 0265-9646. doi: https://doi.org/10.1016/j.spacepol.2021.101474. URL `https://www.sciencedirect.com/science/article/pii/S0265964621000667`.

P. Hill. Tangibles, intangibles and services: a new taxonomy for the classification of output. *The Canadian journal of economics/Revue canadienne d'Economique*, 32 (2):426–446, 1999.

J. Hirshleifer. Where are we in the theory of information? *The American Economic Review*, 63(2):31–39, 1973.

G. Hoberg and G. Phillips. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016. doi: 10.1086/688176. URL `https://doi.org/10.1086/688176`.

R. Hollmann, C. J. Merchant, R. Saunders, C. Downy, M. Buchwitz, A. Cazenave, E. Chuvieco, P. Defourny, G. de Leeuw, R. Forsberg, et al. The esa climate change initiative: Satellite data records for essential climate variables. *Bulletin of the American Meteorological Society*, 94(10):1541–1552, 2013.

T. Hong and I. Han. Knowledge-based data mining of news information on the internet using cognitive maps and neural networks. *Expert systems with applications*, 23(1): 1–8, 2002.

Y. Huang, Z. xin CHEN, T. YU, X. zhi HUANG, and X. fa GU. Agricultural remote sensing big data: Management and applications. *Journal of Integrative Agriculture*, 17(9):1915–1931, 2018. ISSN 2095-3119. doi: https://doi.org/10.1016/S2095-3119(17)61859-8. URL https://www.sciencedirect.com/science/article/pii/S2095311917618598.

IBM. What is k-nearest neighbors (knn)?, 2023. URL https://www.ibm.com/fr-fr/topics/knn. Accessed: June 16, 2023.

A. Jabbari, O. Sauvage, H. Zeine, and H. Chergui. A french corpus and annotation schema for named entity recognition and relation extraction of financial news. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2293–2299, 2020.

C. Jabbour. Spatial data infrastructures & information: An economic challenge. In *Essays in the Economics of Spatial Data Infrastructures (SDI): Business model, service valuation and impact assessment*, chapter 2, pages 42–89. Université Libanaise, 2019a.

C. Jabbour. How much would you pay for a satellite image? lessons learned from a french spatial data infrastructure. In *Essays in the Economics of Spatial Data Infrastructures (SDI): Business model, service valuation and impact assessment*, chapter 5, pages 154–193. Université Libanaise, 2019b.

A. Jadhav, D. Pramod, and K. Ramanathan. Comparison of performance of data imputation methods for numeric dataset. *Applied Artificial Intelligence*, 33(10): 913–933, 2019.

O. Jafari, P. Nagarkar, B. Thatte, and C. Ingram. Satellitener: An effective named entity recognition model for the satellite domain. In *KMIS*, pages 100–107, 2020.

R. Jiang, R. E. Banchs, and H. Li. Evaluating and combining name entity recognition systems. In *Proceedings of the Sixth Named Entity Workshop*, pages 21–27, 2016.

C. I. Jones and C. Tonetti. Nonrivalry and the economics of data. *American Economic Review*, 110(9):2819–58, 2020.

S. Jutz and M. Milagro-Pérez. Copernicus: the european earth observation programme. *Revista de Teledetección*, (56), 2020.

H. A. Kaul and I. Sopan. Land use land cover classification and change detection using high resolution temporal satellite data. *Journal of Environment*, 1(4):146–152, 2012.

M. W. Kenyhercz and N. V. Passalacqua. Missing data imputation methods and their performance with biodistance analyses. In *Biological Distance Analysis*, pages 181–194. Elsevier, 2016.

C. O. Kile and M. E. Phillips. Using industry classification codes to sample high-technology firms: Analysis and recommendations. *Journal of Accounting, Auditing & Finance*, 24(1):35–58, 2009. doi: 10.1177/0148558X0902400104. URL https://doi.org/10.1177/0148558X0902400104.

R. Kitchin. *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage, 2014.

A. Kowarik and M. Templ. Imputation with the r package vim. *Journal of statistical software*, 74:1–16, 2016.

O. Kramer and O. Kramer. K-nearest neighbors. *Dimensionality reduction with unsupervised nearest neighbors*, pages 13–23, 2013.

M. Lafaye. Benefit assessment of the application of satellite earth observation for society and policy: Assessing the socioeconomic impacts of the development of downstream space-based earth observation applications. In M. Onoda and O. Young, editors, *Satellite Earth Observations and their Impact on Society and Policy*, chapter 7. Springer Nature, 2017.

M. Landoni and dt ogilvie. Convergence of innovation policies in the european aerospace industry (1960–2000). *Technological Forecasting and Social Change*, 147: 174–184, 2019. ISSN 0040-1625. doi: https://doi.org/10.1016/j.techfore.2019.07.007.

URL https://www.sciencedirect.com/science/article/pii/S0040162518310023.

M. Le Pellec-Dairon. *Stratégies d'évaluation et de pilotage de la valeur des projets innovants - Le cas de la valeur environnementale des programmes spatiaux*. PhD thesis, Ecole Polytechnique, 2013.

A. Lebeau. Space: The routes of the future. *Space Policy*, 24(1):42–47, 2008.

A. Lebeau, J.-P. Contzen, R. Gibson, and I. Taylor. The integrated applications promotion: A new field of action for the european space agency. *Space Policy*, 29 (3):197–202, 2013. ISSN 0265-9646. doi: https://doi.org/10.1016/j.spacepol.2013.06.004. URL https://www.sciencedirect.com/science/article/pii/S0265964613000544.

J. Lerner, A. Leamon, and A. Speen. Venture capital activity in the low-earth orbit sector. In P. Besha and A. MacDonald, editors, *Economic Development of Low-Earth Orbit*, chapter 4. 2016.

C. Li and X. Li. Analysis on the Logic of Digital Economy Development in the Era of Big Data. *2021 2nd International Conference on Big Data Economy and Information Management (BDEIM)*, pages 185–188, 2021.

W. Li, M. Nirei, and K. Yamana. There's no such thing as a free lunch in the digital economy. In *Sixth IMF Statistical Forum, Washington, DC, November*, 2018.

D. Libaers, D. Hicks, and A. L. Porter. A taxonomy of small firm technology commercialization. *Industrial and Corporate Change*, 25(3):371–405, 2016.

I. Liu, E. Linck, B. Lal, K. W. Crane, X. Han, and T. J. Colvin. Evaluation of China's Commercial Space Sector. Technical report, Institute for Defense Analyses, 2019. URL http://www.jstor.org/stable/resrep22872.1.

X. Liu, H. Chen, and W. Xia. Overview of named entity recognition. *Journal of Contemporary Educational Research*, 6(5):65–68, 2022.

J. Loomis, S. Koontz, H. Miller, and L. Richardson. Valuing geospatial information: Using the contingent valuation method to estimate the economic benefits of landsat satellite imagery. *Photogrammetric Engineering & Remote Sensing*, 81(8):647–656, 2015. ISSN 0099-1112. doi: https://doi.org/10.14358/PERS.81.8.647. URL https://www.sciencedirect.com/science/article/pii/S0099111215302068.

S. K. Majumdar. The impact of size and age on firm-level performance: some evidence from India. *Review of industrial organization*, 12:231–241, 1997.

G. Maral, M. Bousquet, and Z. Sun. *Satellite Communications Systems: Systems, Techniques and Technology*. John Wiley & Sons, 2020.

V. Mateo-Pérez, M. Corral-Bobadilla, F. Ortega-Fernández, and V. Rodríguez-Montequín. Determination of water depth in ports using satellite data based on machine learning algorithms. *Energies*, 14:2486, 2021.

Mathematica Inc. Quantifying the Benefits to the National Economy from Secondary Application of NASA Technology. Technical report, For NASA, 1975.

P. Maurya, O. Jafari, B. Thatte, C. Ingram, and P. Nagarkar. Building a comprehensive ner model for satellite domain. *SN Computer Science*, 3(3):1–8, 2022.

M. Mazzucato and D. K. Robinson. Co-creating and directing innovation ecosystems? nasa's changing approach to public-private partnerships in low-earth orbit. *Technological Forecasting and Social Change*, 136:166–177, 2018. ISSN 0040-1625. doi: https://doi.org/10.1016/j.techfore.2017.03.034. URL https://www.sciencedirect.com/science/article/pii/S0040162517304122.

Midwest Research Institute (MRI). Economic Impact and Technological Progress of NASA Research and Development Expenditures. Technical report, For the National Academy of Public Administration, Washington DC, 1988.

W. H. Miernyk. *The elements of input-output analysis*. Regional Research Institute, West Virginia University, 2020.

G. S. Miller. The press as a watchdog for accounting fraud. *Journal of Accounting Research*, 44(5):1001–1033, 2006.

H. J. Miller and J. Han. *Geographic data mining and knowledge discovery*. CRC press, 2009.

Ministère de l'Économie et des Finances. France 2030 : grands axes de la stratégie spatiale dévoilés. https://www.entreprises.gouv.fr/fr/actualites/industrie/filieres/france-2030-grands-axes-de-la-strategie-spatiale-devoiles, 2021. Accessed: 2023-05-06.

S. Moranta. The space downstream sector: Challenges for the emergence of a european space economy. *Notes de l'Ifri*, 2022.

D. Mowery. Chapter 29: Military R&D and Innovation. In B. H. Hall and N. Rosenberg, editors, *Handbook of the Economics of Innovation*, volume 2, pages 1219–1256. North-Holland, 2010. doi: https://doi.org/10.1016/S0169-7218(10)02013-7.

K. Murthi, U. Sankar, and H. Madhusudhan. Organizational systems, commercialization and cost-benefit analysis of Indian space programme. *Current Science*, 93(12), 2007.

L. Nardon. New space: l'impact de la révolution numérique sur les acteurs et les politiques spatiales en europe. *Notes de l'Ifri*, (28), 2017.

S. Nogueira, M. A. Moreira, and M. M. Volpato. Relationship between coffee crop productivity and vegetation indexes derived from oli/landsat-8 sensor data with and without topographic correction. *Engenharia Agrícola*, 38:387–394, 2018.

I. Nonaka and H. Takeuchi. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, 1995. ISBN 978-0-19-509269-1.

OECD. *The Space Economy at a Glance 2007*. OECD Publishing, Paris, 2007. doi: https://doi.org/https://doi.org/10.1787/9789264040847-en.

OECD. *OECD Handbook on Measuring the Space Economy*. OECD Publishing, Paris, 2012. doi: https://doi.org/https://doi.org/10.1787/9789264169166-en.

OECD. *The Space Economy at a Glance 2014*. OECD Publishing, Paris, 2014. doi: https://doi.org/https://doi.org/10.1787/9789264217294-en.

OECD. *The Space Economy in Figures*. OECD Publishing, Paris, 2019. doi: https://doi.org/10.1787/c5996201-en.

OECD. Data in an Evolving Technological Landscape. *OECD Digital Economy Papers*, OECD Publishing(346), 2022a. doi: https://doi.org/https://doi.org/10.1787/ec7d2f6b-en. URL `https://www.oecd-ilibrary.org/content/paper/ec7d2f6b-en`.

OECD. *OECD Handbook on Measuring the Space Economy, 2nd Edition*. 2022b. doi: https://doi.org/https://doi.org/10.1787/8bfef437-en. URL `https://www.oecd-ilibrary.org/content/publication/8bfef437-en`.

OECD. *Harnessing "New Space" for Sustainable Growth of the Space Economy*. 2023. doi: https://doi.org/https://doi.org/10.1787/a67b1a1c-en. URL `https://www.oecd-ilibrary.org/content/publication/a67b1a1c-en`.

A. Orlova, R. Nogueira, and P. Chimenti. The present and future of the space sector: A business ecosystem approach. *Space Policy*, 52:101374, 2020. ISSN 0265-9646. doi: https://doi.org/10.1016/j.spacepol.2020.101374. URL `https://www.sciencedirect.com/science/article/pii/S0265964620300163`.

Oxford Economics. The case for space: the impact of space derived services and data. *Online: Oxford Economics http://www. oxfordeconomics. com/my-oxford/projects/129029. Paul Quirke, "African Space Programmes: Political or Scientific Endeavours*, 2009.

M. Papenfus, B. Schaeffer, A. I. Pollard, and K. Loftin. Exploring the potential value of satellite remote sensing to monitor chlorophyll-a for us lakes and reservoirs. *Environmental Monitoring and Assessment*, 192(12):808, 2020.

X. Pasco. *Le nouvel âge spatial. De la Guerre froide au New Space*. CNRS, 2017.

X. Pasco. L'évolution du contexte spatial américain. *Annales des Mines - Réalités industrielles*, Mai 2019(2):21–24, 2019.

V. D. S. Paulino. *Innovation trends in the space industry*. John Wiley & Sons, 2020.

W. Peeters. Evolution of the Space Economy: Government Space to Commercial Space and New Space. *Astropolitics*, 19(3):206–222, 2021. doi: 10.1080/14777622.2021.1984001. URL `https://doi.org/10.1080/14777622.2021.1984001`.

S. M. Pekkanen. Governing the new space race. *American Journal of International Law*, 113:92–97, 2019.

J. N. Pelton. Space telecommunications services and applications. In *Handbook of Satellite Applications*, pages 67–92. Springer, 2013.

J. N. Pelton, S. Madry, and S. Camacho-Lara. *Handbook of Satellite Applications*. Springer, 2017.

A. R. Prest and R. Turvey. *Cost-Benefit Analysis: A Survey*, pages 155–207. Palgrave Macmillan UK, London, 1966. ISBN 978-1-349-00210-8. doi: 10.1007/978-1-349-00210-8_5. URL `https://doi.org/10.1007/978-1-349-00210-8_5`.

PwC France. *Copernicus Market Report*. Publications Office of the EU, 2019. doi: 10.2873/011961.

E. Quintana. The new space age. *The RUSI Journal*, 162(3):88–109, 2017. doi: 10.1080/03071847.2017.1352377.

D. K. Robinson and M. Mazzucato. The evolution of mission-oriented policies: Exploring changing market creating policies in the US and European space sector. *Research Policy*, 48(4):936–948, 2019.

M. Rohera. Indian space policy for the private sector. Technical report, Aerospace Security - Center for Strategic and International Studies, 2021.

P. M. Romer. Endogenous technological change. *Journal of political Economy*, 98(5, Part 2):S71–S102, 1990.

D. Rosenberg. Data before the fact. *Raw data" is an oxymoron*, pages 15–40, 2013.

N. Rosenberg. Science, invention and economic growth. *The Economic Journal*, 84 (333):90–108, 1974.

J. Rowley. The wisdom hierarchy: representations of the dikw hierarchy. *Journal of information science*, 33(2):163–180, 2007.

J. Sadowski. When data is Capital: Datafication, Accumulation, and Extraction. *Big Data & Society*, 6(1):2053951718820549, 2019. doi: 10.1177/2053951718820549. URL `https://doi.org/10.1177/2053951718820549`.

P. A. Samuelson. The pure theory of public expenditure. *The review of economics and statistics*, pages 387–389, 1954.

J. Schmookler. *Invention and economic growth*. Harvard University Press, 2013.

C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

Q. Song, M. Shepperd, X. Chen, and J. Liu. Can k-nn imputation improve the performance of c4. 5 with small software project data sets? a comparative evaluation. *Journal of Systems and software*, 81(12):2361–2370, 2008.

Space IGS. *The Space Innovation & Growth 2014-2030: Space Growth Action Plan*. Space Innovation and Growth Strategy, 2014.

G. J. Stigler. The economics of information. *Journal of political economy*, 69(3): 213–225, 1961.

J. E. Stiglitz. The contributions of the economics of information to twentieth century economics. *The quarterly journal of economics*, 115(4):1441–1478, 2000.

A.-H. Tan et al. Text mining: The state of the art and the challenges. In *Proceedings of the pakdd 1999 workshop on knowledge disocovery from advanced databases*, volume 8, pages 65–70, 1999.

C. Tang. Data industry chain. In C. Tang, editor, *The Data Industry: The Business and Economics of Information and Big Data*, pages 41–58. John Wiley & Sons, 2016.

A. Tassa. The socio-economic value of satellite earth observations: huge, yet to be measured. *Journal of Economic Policy Reform*, 23(1):34–48, 2020. doi: 10.1080/17487870.2019.1601565.

Technopolis group. *Design of a Methodology to Evaluate the Direct and Indirect Economic and Social Benefits of Public Investments in Space*. Prepared for ESA, 2012.

E. Tobback, H. Naudts, W. Daelemans, E. J. de Fortuny, and D. Martens. Belgian economic policy uncertainty index: Improvement through text mining. *International journal of forecasting*, 34(2):355–365, 2018.

Y. Tong and A. Fortino. Using text data mining to discover the naics industry code for a business.

UK Space Agency. *Summary Report: The Size and Health of the UK Space Industry 2015*. Prepared by London Economics, London, United Kingdom, 2016. URL `https://londoneconomics.co.uk/blog/publication/size-health-uk-space-industry-2016/`.

UK Space Agency. *Summary Report: The Size and Health of the UK Space Industry 2018*. Prepared by London Economics, London, United Kingdom, 2019. URL `https://www.gov.uk/government/publications/uk-space-industry-size-and-health-report-2018`.

UK Space Agency. *Summary Report: The Size and Health of the UK Space Industry 2020*. Prepared by Know.space, London, United Kingdom, 2021. URL https://www.gov.uk/government/publications/uk-space-industry-size-and-health-report-2020.

UK Space Agency. *Summary Report: The Size and Health of the UK Space Industry 2021*. Prepared by BryceTech, London, United Kingdom, 2022. URL https://www.gov.uk/government/publications/the-size-and-health-of-the-uk-space-industry-2021/size-and-health-of-the-uk-space-industry-2021.

H. Varian. Artificial intelligence, economics, and industrial organization. In *The economics of artificial intelligence: an agenda*, pages 399–419. University of Chicago Press, 2018.

H. R. Varian. *Markets for information goods*, volume 99. Citeseer, 1999.

A. Vernile. *The Rise of Private Actors in the Space Sector*. Springer, 2018.

M. Vidmar. New space and innovation policy: Scotland's emerging "space glen". *New Space*, 8(1):31–51, 2020. doi: 10.1089/space.2019.0032.

M. Weinzierl. Space, the final economic frontier. *Journal of Economic Perspectives*, 32(2):173–92, 2018.

R. Welch. Monitoring urban population and energy utilization patterns from satellite data. *Remote sensing of Environment*, 9(1):1–9, 1980.

K. Whealan George. The economic impacts of the commercial space industry. *Space Policy*, 47:181–186, 2019. ISSN 0265-9646. doi: https://doi.org/10.1016/j.spacepol.2018.12.003. URL https://www.sciencedirect.com/science/article/pii/S0265964618300651.

X. Yang and Z. Liu. Use of satellite-derived landscape imperviousness index to characterize urban spatial growth. *Computers, Environment and Urban Systems*, 29(5):524–540, 2005. ISSN 0198-9715. doi: https://doi.org/10.1016/j.compenvurbsys.2005.01.005. URL https://www.sciencedirect.com/science/article/pii/S0198971505000049. Remote Sensing for Urban Analysis.

V. Zervos and D. S. Siegel. Technology, security, and policy implications of future transatlantic partnerships in space: Lessons from galileo. *Research Policy*, 37(9):

1630–1642, 2008. ISSN 0048-7333. doi: https://doi.org/10.1016/j.respol.2008.06.008. URL https://www.sciencedirect.com/science/article/pii/S004873330800142X.

List of Figures

List of Tables

Kenza BOUSEDRA

CHANGE IN COMMERCIAL SECTORS EXPLOITING SPACE INFRASTRUCTURE: ANALYSIS AND ECONOMIC INDICATORS

RÉSUMÉ

Cette thèse développe une méthode d'évaluation de l'importance économique des secteurs commerciaux utilisant les infrastructures spatiales, regroupés sous l'appellation de secteur spatial aval. La Partie 1 (Chap. 1 et 2) situe le contexte et identifie les enjeux méthodologiques. Le Chapitre 1 aborde le New Space comme changement structurel du secteur spatial, accélérant le développement commercial de l'espace et ouvrant de nouvelles opportunités de marché. Le Chapitre 2 réalise une revue de littérature pour souligner les enjeux liés à l'évaluation de la partie aval du secteur spatial. La Partie 2 (Chap. 3 et 4) se concentre sur le développement de la méthodologie d'évaluation, appliquée à la France en 2021. Le Chapitre 3 propose une approche basée sur la reconnaissance d'entités nommées pour identifier les entreprises spatiales aval. Le Chapitre 4 détaille la mesure des revenus du secteur spatial aval. Enfin, la Partie 3 (Chap. 5) enrichit la méthode par une analyse théorique de la nature économique des données et de l'information à l'ère numérique, en développant un cadre inspiré de l'approche flux-fond pour caractériser leur rôle dans la création de valeur.

Mots clefs: Secteur spatial aval, New Space, méthodologie d'évaluation, importance économique, named entity recognition, mesure des revenus, économie de l'information

SUMMARY

This thesis develops a methodology for assessing the economic importance of commercial sectors using space infrastructure, grouped under the term downstream space sector. Part 1 (Chapters 1 and 2) contextualizes and identifies methodological challenges. Chapter 1 discusses the New Space era as a structural change in the space sector, accelerating commercial space development and opening new market opportunities. Chapter 2 conducts a literature review to underline challenges in evaluating the space sector's downstream part. Part 2 (Chapters 3 and 4) focuses on developing the evaluation methodology applied to France in 2021. Chapter 3 proposes an approach based on named entity recognition to identify downstream space companies. Chapter 4 details the measurement of revenues from the downstream space sector. Lastly, Part 3 (Chapter 5) enriches the method with a theoretical analysis of the economic nature of data and information in the digital era, developing a framework inspired by the flow-fund approach to characterize their role in value creation.

Keywords: Downstream space sector, New Space, evaluation methodology, economic importance, named entity recognition, revenue measurement, information economics