Redefining Data Governance: Insights from the French University System

Guy Melançon[1] [a], Nathalie Pinède[2,3] [b] and Ugo Verdi[1,3] [c]

[1]*Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, F-33400 Talence, France*
[2]*Univ. Bordeaux, CNRS, Bordeaux INP, IMS, UMR 5218, F-33400 Talence, France*
[3]*MICA - Médiation, Information, Communication, Art, MSHA – 10 Esplanade des Antilles, 33607 Pessac, France*

Keywords: Data Governance, French Universities, Collegiality, Subsidiarity Principle, Inter-Organizational Collaboration, Multi-Tiered Governance.

Abstract: This position paper highlights the distinctive features of French universities that render data governance within these institutions a particularly challenging endeavor. These universities inherently operate in an exceptionally open milieu, necessitating the conceptualization of governance as a dynamic and adaptable framework that converges seamlessly with the governance structures of other institutions. The principle of collegiality further mandates a distributed approach to data governance, encompassing responsibilities, rules, and procedures across various levels of management. Moreover, it is essential to reevaluate the conventional viewpoint that segregates administrative tasks from research and teaching functions. Our findings underscore the necessity for developing and executing a dynamic, multi-tiered data governance model that integrates the three fundamental missions of universities. Given the intrinsic nature of French universities, it is imperative to envisage governance as an evolving ecosystem of agents who assume complementary responsibilities in a harmonized manner.

1 INTRODUCTION

In the current era, characterized by a data-intensive landscape (Bouchez, 2014) and an innovation-driven knowledge economy (Wessels et al., 2017), universities emerge as pivotal contributors to research. Their role is further accentuated by the imperative of sustainability (Musselin, 2006). Like all organizations, they must anchor their strategic and managerial decisions in the analysis of data that encapsulates their operations. This goal, propelled by the generation and utilization of data, necessitates the development of robust data governance within these institutions.

While some Anglo-Saxon universities have already established data governance frameworks[1], their French counterparts appear to be in a state of catch-up, often contingent upon centralized regulatory di-

[a] https://orcid.org/0000-0003-3193-7261
[b] https://orcid.org/0000-0002-5381-5524
[c] https://orcid.org/0000-0002-0562-6032

[1]Examples include University of Wisconsin-Madison, University of Michigan, New York University, Oxford University, University of San Francisco, University of Toronto and University of Washington.

rectives. The adoption of "corporate" models of data governance, frequently based on DMBOK guidelines (DAMA International, 2017), might not be entirely suitable for French universities. This mismatch underscores the ongoing efforts to revamp their managerial structures (Mignot-Gérard et al., 2023) and highlights the need for a tailored approach to data governance in this specific milieu.

A deeper understanding of the context and operations within academic settings will lay the groundwork for developing a dual-purposed methodology: firstly, to establish data governance, detailing its architecture and operations, and secondly, to assess it by evaluating the maturity level of the institution's data governance practices.

This position paper delves into the unique characteristics of French universities, positioning data governance as a significant and distinct case study. Our analysis is grounded in the experience of the University of Bordeaux, which is currently implementing data governance strategies. We pay particular attention to the ANR ACT initiative (ACT stands for "Augmented university for Campus and world Transition", the project being funded by the French Na-

tional Research Agency), a pioneering project that encompasses a suite of strategic institutional projects, all dedicated to the principles and practices of data governance.

2 THE FRENCH CONTEXT

In our examination, we critically analyze French universities as distinct entities, underscoring traits that not only set them apart from the corporate sector but also render the French university system unique in its essence. Despite recent significant transformations within these institutions, French universities continue to navigate the delicate balance between striving for autonomy and adhering to the mandates of a centralized state (Vinokur, 2008; Musselin, 2017). This intricate dynamic results in specific demands for data governance that are unique to the French context (Pierronnet and Blanchard, 2021). This duality of independence and state directives shapes the framework within which these universities must operate, influencing their approach to managing and utilizing data.

2.1 Interconnected Governance: The Open and Collaborative Structure of French Universities

The organizational complexity of French universities often astonishes those unfamiliar with the system. A hallmark of these institutions is their collaboration with a variety of supervisory bodies, both local (such as Institut Polytechnique) and national (like CNRS[2], Inserm[3], or INRAE[4]). While research laboratories are physically situated within university campuses, a substantial portion of the staff utilizing these facilities, operating local infrastructures, or managing administrative processes are not university employees. This openness fosters a deep entanglement of personnel and daily operational protocols from various institutions, a complexity that transcends simple contractual arrangements and is ingrained in the very fabric of the university system.

French universities, as state entities, must harmonize their strategies not only with the state's objectives but also operate within a strict regulatory framework. A significant number of the degrees they confer

---

[2]CNRS is the Centre National de la Recherche Scientifique

[3]Inserm is the Institut national de la santé et de la recherche médicale

[4]INRAE is the Institut national de recherche pour l'agriculture, l'alimentation et l'environnement

are national diplomas. The Ministry of Higher Education exercises oversight and approval of university curricula to maintain nationwide uniformity. This ensures that the value of a diploma is consistent, irrespective of the issuing university.

However, France's state-led research policy is delegated to specific research organizations (state operators) like the CNRS, with other bodies overseeing research in particular domains (Inserm for health, INRAE for environmental sciences, etc.). These operators' policies significantly influence the universities' research strategies, necessitating alignment with national priorities. Furthermore, the state's data policy shapes the policies of these organizations and universities (Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, 2021).

This open and interwoven organizational structure profoundly affects the intellectual property rights of research outcomes and the generated (research) data. Many projects produce their data while also utilizing external data from national databases or from institutional and industrial partners. Given the diverse institutional affiliations of team members, data must be regarded as a shared asset. This complexity extends to administrative data related to project management, hiring processes, and more.

Consequently, this mode of operation necessitates a conceptualization of university data governance that is intricately connected with that of organizations like CNRS, Inserm, and partner entities. Thus, when establishing governance structures, it is crucial to consider not only university stakeholders but also broader ecosystem actors at both decision-making and organizational levels. Governance rules and procedures must be harmonized with those of ecosystem participants, mirroring challenges akin to those posed by the European Data Protection Regulation (GDPR). While harmonization might sometimes be project-specific, involving *ad hoc* partners, achieving consistency at an institutional level with organizations such as CNRS or Inserm, for instance, can be essential.

This scenario stands in stark contrast to governance models prevalent in Anglo-Saxon universities (Jim and Chang, 2018) or those proposed in academic literature (Abraham et al., 2019), highlighting the unique challenges and considerations in the French context.

2.2 Balancing Governance: Collegiality and Subsidiarity in University Structures

Universities are deeply rooted in a tradition of collegiality, which inherently precludes a top-down,
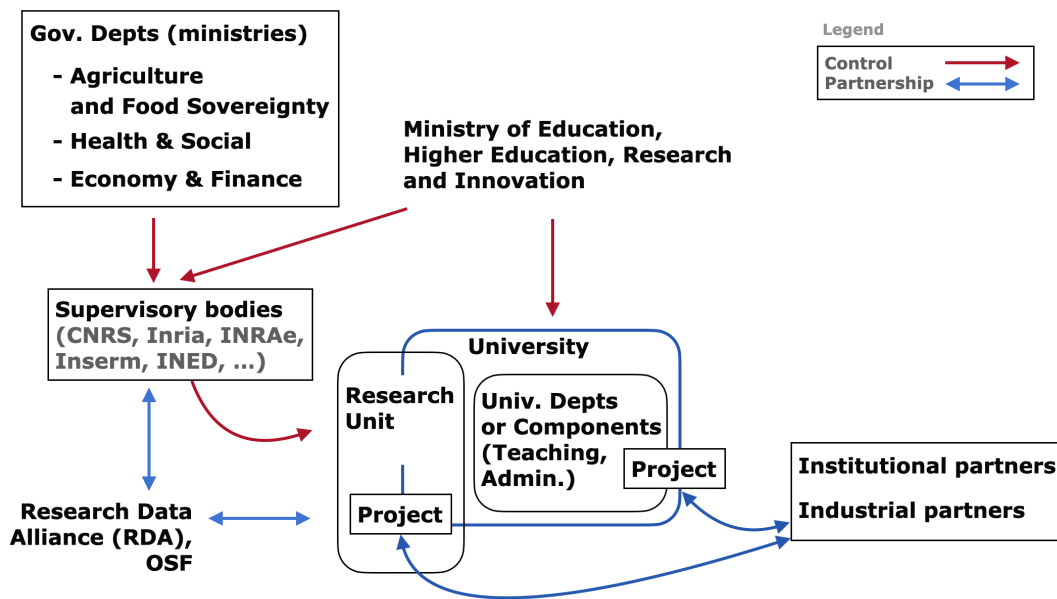
Figure 1: A look at the French academic ecosystem. Multiple supervisory entities shape data governance through their scientific policies and operational connections with research endeavors. This influence extends to collaborations with local authorities or industrial partners, driven by institutional mandates or specific projects. The interconnectedness of university projects, units, and components necessitates a nuanced, multi-tiered approach to governance, necessitating a tailored application of the subsidiarity principle to accommodate this complex structure effectively.

strictly hierarchical governance model (Musselin, 2006). A notable challenge arising from this tradition is that, despite the coherence and intelligence demonstrated by top-level management teams in collaboration with central services, directors of lower-level units (such as teaching departments and research units) often play a lesser role in decision-making or in endorsing the institution's policies (Chatelain-Ponroy et al., 2012).

The appointment of a data steward for specific domains or research areas is complicated by researchers' tendency to act independently, and for instance freely select data storage infrastructures (for example, recherche.data.gouv.fr at a national level, or Zenodo at European level). This independence extends to course materials, which may be hosted either on local platforms or national portals like France Universités Numériques (FUN).

Project leaders generally establish their data policies within their own areas, regardless of the project's strategic significance to the university. Incidentally, these projects often include staff from various organizations and might even receive joint funding from these entities. This scenario directly affects how upper-tier data governance bodies should tackle governance issues to ensure alignment between overarching institutional rules and specific project-level best practices.

In this light, the principle of subsidiarity is indispensable, providing a framework for allocating competences between these two governance tiers. The challenge lies in devising a harmonious integration that is applicable across the university while still being pertinent to individual projects.

Moreover, the principle of subsidiarity might be the optimal strategy for incorporating the multi-site aspect of university campuses. Local administrators and elected officials may resist policies that mandate the centralization of data, favoring instead a distributed model that grants them some measure of control and influence over the data they generate or utilize.

2.3 The University, The Three-Headed Cerberus of Education

From a business standpoint, data governance is typically geared towards enhancing efficiency and profitability. However, in the context of universities, while efficiency and sustainability remain important, there is also a strong emphasis on social goals such as democratizing access to science and ensuring its affordability for the public. Additionally, universities are bound by the requirements of transparency in public administration.

Universities' objectives span three primary mis-

sions: research, teaching, and administration (with universities acting as implementers of state education policy). At first glance, this tripartite mission might suggest a need for a three-pronged approach to data governance.

- Research data is notably diverse, varying in origin, use, and particularly in terms of its producers and users, who are often affiliated with different supervisory bodies and come from various disciplinary backgrounds. This data also reflects the practices of different research communities. Moreover, national and European open science initiatives play a significant role in shaping the governance of research data at the institutional level.

  While universities might leverage insights from the Research Data Alliance (RDA) which primarily aligns with open science (Madison, 2020), the approach here diverges, focusing on integrating data governance within the strategic framework of universities as socio-economic entities. Our approach critically examines the segmentation of research data management as an isolated university function.

- Teaching-related data, in contrast, is more localized, originating from educational activities involving university students. This data type is subject to multiple regulatory frameworks: the European GDPR, due to its inclusion of personal data (like student profiles and academic records), and national regulations regarding the preservation and archiving of educational data, among others.

- Institutional data, usually sourced from the university's information system, most closely resembles corporate data. However, this comparison is somewhat superficial, as we will explore further in the subsequent sections.

In conclusion, while the framework of data governance in universities may draw some parallels with corporate models, it is distinct in its complexity and scope. Universities must navigate a unique landscape where data governance intersects with diverse academic disciplines, regulatory environments, and social responsibilities. This intricate matrix demands a bespoke approach to data governance, one that harmoniously integrates the nuances of research, teaching, and administrative data while aligning with broader societal and educational objectives.

3 REDEFINING DATA GOVERNANCE IN FRENCH UNIVERSITIES: A MULTIFACETED APPROACH

The insights gathered from our observations underscore the necessity for data governance in French universities to be conceptualized as a multi-tiered framework. Notably, existing literature on data governance in primarily corporate contexts, such as the work of (Alhassan et al., 2018) and (Abraham et al., 2019), or in the realm of higher education as examined by (Jim and Chang, 2018), does not adequately address the requirement for a distributed governance model that can effectively accommodate collegiality. The principle of subsidiarity is critical here, as it must empower the university to manage its collegial structure, while simultaneously ensuring that individual projects are consistent with the university's overarching data strategy.

3.1 Strategic Project Identification and Advancing Data Literacy

In a governance structure envisioned as multi-tiered, it is essential for projects to extend beyond simple data management plans to include comprehensive managerial activities. These activities should be in harmony with and contribute to the broader strategies of the affiliated institutions. The role of data literacy, as a spectrum of digital competencies, is increasingly recognized in universities (Verdi and Le Deuff, 2020). This literacy should be leveraged to integrate data competencies with strategic data governance, enabling project leaders to understand and align with the university's strategic data objectives. Governance structures, thus, need to be adept at identifying and focusing on projects with strategic significance.

3.2 Dynamic Governance and Inter-Organizational Collaboration

The convergence between a university's data governance and that of its ecosystem partners is an integral aspect of the inter-organizational governance dimension. Universities often excel in crafting multiparty agreements involving diverse partners. This expertise, encompassing strategic, legal or functional domains, should be extended to the realm of data governance. Data considerations should become a staple in long-term collaborative programs, ensuring continuous dialogue with both institutional and industrial partners.

3.3 Unifying Missions Through Data Governance

Data governance in universities should act as a unifier of the institution's primary missions: administration, research, and teaching. Current trends show universities harnessing their research capabilities to address critical issues like energy transition, sustainable development, and social responsibility.

The ANR ACT SmartMob/Datacampus project exemplifies this trend. This project, focused on mapping campus mobility to improve transport services. Three organizations fund this initiative: Université de Bordeaux, Bordeaux INP and Fondation Université de Bordeaux. Obviously, Bordeaux Metropole shows interest in the initiative and could possibly wish to reuse part of the research including data, and even share data produced by IoT devices installed on their facilities surrounding the campuses.

It demonstrates how administrative data can be transformed into research data and educational resources. Within the scope of this project, administrative data detailing user mobility on campuses is used for research, complementing scientific publications. This data is also customized for educational purposes, utilized in classroom settings, term projects or student internships. Furthermore, additional administrative data from Université de Bordeaux and Bordeaux INP, like course schedules is integrated into the project to aid in developing predictive models, such as those forecasting peak campus hours.

This multifaceted use of data underlines the need for a unified governance approach that accommodates administrative, research, and teaching objectives seamlessly.

In summary, this practical example vividly illustrates the convergence of administration, research, and teaching around a shared data-centric project. It also highlights the necessity for different organizations' governance models to synergize, facilitating the acquisition and utilization of data for comprehensive research programs. This approach marks a significant departure from traditional data governance models, advocating for a more integrated, collaborative, and flexible framework in French universities.

4 CONCLUDING REFLECTIONS: CHARTING THE PATH FOR DATA GOVERNANCE IN FRENCH UNIVERSITIES

The transformative shifts that have swept through the French university system over the past decade necessitate a more structured and managerial approach to data governance. However, these changes have not diminished the deeply-rooted collegial ethos of French universities (Mignot-Gérard et al., 2023). Consequently, the implementation of data governance must be conceptualized as a dynamic and responsive process, one that adapts to the evolving characteristics of universities amidst ongoing transformations.

Dynamicity should be pursued as a fundamental objective, endowing data governance with the versatility to navigate diverse contexts – economic, political, regulatory, or systemic. This adaptability is crucial for realizing the full potential of a data-driven strategy and management approach.

Dynamicity in data governance resonates with the concepts of *data lineage* and the varying needs of different *data audiences*. It is often necessary to tailor data access based on the intended data usage and user profiles. While data typically originates as a singular, well-defined set representing a specific phenomenon, it often evolves into multiple datasets through processes such as simplification, aggregation, or anonymization. These derived datasets are crafted for distinct audiences, which may not be fully identified at the data's inception. This evolving nature of data underscores the need for flexible governance that accommodates rules and procedures to varying data transformations and user requirements, possibly unknown at the origin of the data lineage.

Upon reviewing existing frameworks for data governance (e.g., (Alhassan et al., 2018); (Abraham et al., 2019); (Gagnon-Turcotte et al., 2021); (Seiner, 2014)), it becomes apparent that none sufficiently encapsulate the nuanced characteristics imperative for a dynamic, multi-level governance model. The intrinsic complexity of French universities necessitates envisioning governance as a fluid system, where agents assume complementary responsibilities in a harmonized fashion. Yet, within this distributed system, a regulatory framework dictates oversight by a higher authority, ensuring coherence and alignment, while still allowing for innovation and adaptation from various levels and peripheral players.

We are currently conducting interviews with various stakeholders, including project leaders and data-related service managers (like those in IT and Legal

Departments). These interviews serve a dual purpose. Firstly, they contribute to the development of a governance maturity model specifically tailored for the French academic context. This model is partly based on insights gained from these interviews. Secondly, the interviews are instrumental in identifying effective mechanisms for establishing a dynamic, multi-level governance architecture, a crucial aspect for evolving academic environments.

In the realm of governance, a pivotal focus should be placed on data literacy. This concept extends beyond mere skill acquisition. It encompasses a broader understanding and contextual application of data. It involves cultivating a mindset that recognizes the strategic value of data, encouraging individuals to think critically about how data can be utilized effectively within their specific roles and projects. This aspect of data literacy is about developing a deeper, more nuanced appreciation of data's role in decision-making, problem-solving, and innovation. It's about empowering individuals to not just use data tools and techniques, but to understand the implications of data in the broader context of their work, the organization's goals, and even societal impacts. This holistic approach to data literacy facilitates a culture where data is not just a tool, but a fundamental component of strategic thinking and planning.

ACKNOWLEDGEMENTS

REFERENCES

Abraham, R., Schneider, J., and vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49:424–438.

Alhassan, I., Sammon, D., and Daly, M. (2018). Data governance activities: a comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, 31(2):300–316.

Bouchez, J.-P. (2014). Autour de "l'économie du savoir": Ses composantes, ses dynamiques et ses enjeux. *Savoirs*, 34(1):945.

Chatelain-Ponroy, S., Mignot Gerard, S., Musselin, C., and Sponem, S. (2012). La gouvernance des universités françaises. Research report, HEC Montréal ; Centre de sociologie des organisations ; Conservatoire national des arts et métiers ; Université Paris-Est Créteil.

Gagnon-Turcotte, S., Sculthorp, M., and Coutts, S. (2021). Les partenariats de données numériques: mettre les bases d'une gouvernance de données collaboratives dans l'intérêt du public. Technical report, Laboratoire d'innovation urbaine de Montréal.

Jim, C. K. and Chang, H. (2018). The current state of data governance in higher education. *Proceedings of the Association for Information Science and Technology*, 55(1):198–206.

Madison, M. J. (2020). Data governance and the emerging university. In *In Research Handbook on Intellectual Property and Technology Transfer*, pages 364–390. Edward Elgar Publishing.

Mignot-Gérard, S., Chatelain-Ponroy, S., Sponem, S., and Musselin, C. (2023). Kaleidoscopic collegiality and the use of performance research metrics. the case of french universities. *Higher Education*, 85:887–918.

Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (2021). Politique des données, des algorithmes et des codes sources : Feuille de route 2021-2024. Technical report, MESRI.

Musselin, C. (2006). Are universities specific organisations? In Krücken, G., Kosmützky, A., and Torka, M., editors, *Towards a Multiversity? Universities between Global Trends and National Traditions*, pages 63–84. Bielefeld: transcript Verlag.

Musselin, C. (2017). *La grande course des universités.* Presses de Sciences Po.

Pierronnet, R. and Blanchard, A. (2021). La donnée : un enjeu managérial à saisir pour les universités. In *Proximités et territoires aux défis du management public*. Association Internationale de Recherche en Management Public (AIRMAP).

Seiner, R. S. (2014). *Non-Invasive Data Governance: The Path of Least Resistance and Greatest Success.* Technics Publications.

Verdi, U. and Le Deuff, O. (2020). La data literacy distribuée: Périmètres définitionnels, origines documentaires, perspectives réticulaires. *Les Cahiers du numérique*, 16:137 – 173.

Vinokur, A. (2008). Vous avez dit "autonomie" ? *Mouvements*, 55-56:72 – 81.

Wessels, B., Finn, R., Wadhwa, K., and Sveinsdottir, T. (2017). *Open Data and the Knowledge Society.* Amsterdam University Press.