Algorithmic Transparency in the Public Sector

Case studies of repositories of public algorithms in Chile, the EU and the UK

November 2024

GPAI / THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

Acknowledgements

This report was developed in the context of the *"Algorithmic Transparency in the Public Sector"* project, with the steering of the project Co-Leads and the guidance of the Project Advisory Group, supported by the GPAI Responsible AI and Data Governance Working Groups. The GPAI Responsible AI and Data Governance Working Groups agreed to declassify this report and make it publicly available.

Co-Leads:
Juan David Gutiérrez[*], Universidad de los Andes
Alison Gillwald[*], Research ICT Africa/University of Cape Town

The report was written by **Arturo Arriagada**[‡] and **Romina Garrido**[‡] (Universidad Adolfo Ibáñez) with the collaboration of case study colleagues from **Chile's Algoritmos Públicos UAI**[†], **EU's Public Sector Tech Watch**[†] and **UK's Algorithmic Transparency Recording Standard**[†]**.**

GPAI would like to acknowledge the meaningful contributions of **Sarah Muñoz-Cadena**[‡], Universidad del Rosario; **Antoine Glory**[†], International Centre of Expertise in Montréal on Artificial Intelligence (CEIMIA) and **Matías Cifuentes**[†] (Universidad Adolfo Ibáñez).

GPAI would also like to thank the members of the project's advisory group:**Avik Sarkar*,** Indian School of Business; **Przemyslaw Biecek*,** University of Warsaw; **Seydina M. Ndiaye*,** Université numérique Cheikh Hamidou Kane – UN-CHK; **Abdul Majeed**[†]**,** Gachon University; **Juliana Sakai* ,** Transparência Brasil; **Seong Oun Hwang*,** Gachon University; **Thierry Warin*,** HEC Montréal; **Dafna Feinholz**,** UNESCO; **Rosanna Fanni**,** UNESCO; **Irakli Khodeli**,** UNESCO; **Beatriz Botero Arcila**[†], Sciences Po; **Cesar Renteria Martin**[†], Centro de Investigación y Docencia Económicas; **Danielle Zaror**[†], Facultad de Derecho en Universidad de Chile; **Jorge Luis Ordelin**[†], Derecho y Nuevas Tecnologías; and **Violet Turri**[†], Carnegie Mellon University Software Engineering Institute.

Finally, GPAI would like to acknowledge the tireless efforts of colleagues at the International Centre of Expertise in Montréal on Artificial Intelligence (CEIMIA). We are particularly grateful for the support of **Arnaud Quenneville-Langis** and **Antoine Glory** from CEIMIA, and for the dedication of the Working Groups Co-Chairs **Amir Banifatemi**[*], **Francesca Rossi**[*], **Bertrand Monthubert**[*] and **Shameek Kundu**[*].

* Expert of GPAI's Responsible AI Working Group or Data Governance Working Group
** Observer of GPAI Responsible AI Working Group or Data Governance Working Group
† Invited Specialist
‡ Contracted parties by the CEIMIA

Citation

GPAI (2024). "Algorithmic Transparency in the Public Sector: Case studies of repositories of public algorithms in Chile, the EU and the UK". Report, November 2024, Global Partnership on Artificial Intelligence.

Table of Contents

Introduction

Artificial intelligence (AI) is becoming integral to effective public sector delivery. Instruments for algorithmic transparency can help disclose and disseminate information and processes regarding the use of algorithms in the public sector to a broader audience and user base. Public repositories of algorithms are one of these instruments, which serve as “windows” and “channels” for individuals to understand system operations, decision-making processes (or ‘explainability’), and accountability while also providing clarity on data sources and potential outcomes (‘traceability’) (Gutiérrez & Muñoz-Cadena, 2023).

As part of a broader project aimed at studying algorithmic transparency in the public sector, with an emphasis on assessing both reactive and proactive transparency instruments that help governments comply with algorithmic transparency principles, standards, and rules, this report focuses on the development and management of public algorithm repositories (GPAI, 2024). These repositories are essential for fostering "meaningful transparency" and serve as a key tool in ensuring accountability to independent stakeholders, including citizens and government agencies (Suzor et al., 2019). This report explores the strengths, weaknesses, and challenges involved in building these instruments, as well as their diverse uses, users, and associated costs. It examines how transparency tools complement one another and contribute to goals such as explainability and accountability. To further understand these objectives, this report scrutinises three case studies of public algorithm repositories, providing a comprehensive overview of key findings. The selected cases represent a diverse range of factors, including the type of organisation managing the repository and the scope of its database. This will allow us to evaluate how these tools contribute to public value creation and enhance transparency in governments and societies' use of algorithms and AI technologies.

The first case is the Public Algorithm Repository (Repositorio Algoritmos Públicos) developed by Universidad Adolfo Ibáñez (UAI) in Chile, a supply-driven algorithmic transparency initiative characterised by a user-friendly website that allows for easy visualisation of 101 algorithmic tools implemented within the country's public administration and provides information on 17 to 19 different variables[1]. The repository is also available as CSV and XLS documents for data science compatibility.

The second case is the Public Sector Tech Watch Observatory, previously AI Watch, a repository of public algorithms developed by the European Commission’s Directorate-General for Digital Services (DIGIT) in collaboration with the Joint Research Centre (JRC). The project comprises 1666 cases of artificial intelligence tools and other emerging technologies across the European Union and categorises them into 46 variables. The website allows users to download the repository in CSV and XLS format. It also provides a link to interactively visualise some of its data through the Public Sector Tech Watch Cases Viewer & Statistics platform.

The third and last case is the repository based on the UK’s Algorithmic Transparency Recording Standard (ATRS), a framework designed to assist UK public institutions in adhering to the recent mandate issued by the Responsible Technology Adoption Unit (RTA), which requires the disclosure of algorithmic tool usage in the public sector. At the time of writing, the Algorithmic Transparency Records repository compiles extensive information (in some cases, with over 70 variables)[2] on nine algorithmic tools. It is presented through written text on individual pages on the UK.GOV website.

[1] The number of variables present in the Algoritmos Públicos website differs between each algorithm entry (12 to 13) and from the information available on the downloadable files for the repository (19).
[2] The number of variables detailed varies greatly from one entry to another, from 37 in the first cases to 72 in the latest.

This report examines various aspects of the repositories' construction, governance, and functionality. The repository development process includes details on who built it, how it was designed, and the key challenges faced during its creation. Information on how and by whom the repository can be accessed is also provided. Governance of the repository is explored by analysing its organisational structure and the decision-making process regarding which automated decision-making (ADM) systems are included. The report describes the types of computational systems contained in the repository, their tasks, and their outputs, including whether humans are involved in the system's control. Additionally, the types of metadata disclosed for each system and any existing standards for repository classification are discussed.

In this report, data lifecycle management is reviewed by assessing compliance with the rules, laws, and codes of conduct of the repository, along with its data management framework. The report also considers the types of users and their feedback on the repository's usefulness and identifies potential new user types. Broader contributions of the repository are assessed, especially regarding its role in enhancing transparency and value creation for various stakeholders. Furthermore, the report investigates key performance indicators (KPIs) used to measure the repository's outcomes and impact. A comparative analysis is conducted to evaluate the repository against other transparency instruments worldwide, considering aspects such as value creation, cost, and usage. Lastly, the repository's alignment with GPAI's data governance principles, including data justice, is examined to identify points of convergence or divergence.

The first section explains the concepts of "algorithmic transparency", "algorithmic transparency instruments", and "repositories of public algorithms. The second section describes the methodology used in this study. The third section presents the main findings of the analysis of the three case studies. The final section discusses the main findings, how these repositories work, and their contribution to algorithmic transparency.

1. Algorithmic transparency instruments: Repositories or registers of public algorithms

Transparency is a central principle of democratic accountability and public interest regulation. Algorithmic transparency is a multifaceted concept that can be understood simultaneously as a capacity, a principle, a standard, a rule, and an instrument (for a detailed description, see GPAI, 2024). As the first GPAI report of this project suggests (2024, p. 2), "these dimensions are not mutually exclusive and, in some cases, are associated". In this report, we focus on instruments developed to achieve algorithmic transparency, which are dependent on transparency principles and are operationalized through a set of standards and rules. We understand algorithmic transparency instruments as mechanisms that "provide information about algorithmic systems to the general public (e.g., affected persons, media, or civil society) so that individuals or groups can learn that these systems are in use and demand answers and justifications related to such use" (Ada Lovelace Institute et al., 2021, p. 18).

Algorithmic transparency instruments can help to provide information and processes related to algorithmic uses in the public sector to broader audiences and users. Valderrama et al. (2023) classified algorithmic transparency mechanisms into three types: 1) disclosure (e.g., information requests, model cards, and AI Registers), 2) explanation, and 3) evaluation (e.g., audits and impact assessments). Public algorithm registers can share information in various formats and are seen as "proactive transparency". Model sheets are used to document algorithmic systems.

The previous GPAI report of this project (2024, p. 8-12) proposes a new classification of algorithmic transparency instruments with two broad categories based on the source of the algorithmic transparency initiative: 1) supply-driven transparency instruments and 2) demand-driven transparency instruments. *Supply-driven instruments* allow governments to proactively disclose information about algorithms used in decision-making processes without being prompted by specific requests from individuals or organisations. Thus, three types of instruments under this category were identified: 1) online repositories or registers of public algorithms, 2) user-driven proactive publications, and 3) automated responses triggered by interaction with certain processes.

Online repositories of ADM systems,[3] also called *repositories of public algorithms*, "are a supply-driven transparency instrument that is gaining prominence among governments (Ada Lovelace Institute et al., 2021; OGP, 2022; Valderrama et al., 2023)" (GPAI, 2024, p. 8). These repositories serve as "windows" and "channels" through which "individuals can access information about the functioning of ADM systems, the data used for training and operating these systems, the decision-making processes supported by them, and the procedures for challenging those decisions, among other aspects" (GPAI, 2024, p. 8).

"*User-driven proactive publications* are a transparency tool employed by public entities to proactively release information to the public, informed by previous user requests" (GPAI, 2024, p. 9). *Automated responses* activated by user interaction with certain processes constitute an algorithmic mechanism in which a public entity's platform automatically delivers relevant information about an ADM system. This disclosure occurs without the user explicitly requesting it (GPAI, 2024).

Demand-driven transparency instruments involve the public sector providing information in response to specific individual, group, or authority requests (GPAI, 2024). This differs from supply-side transparency, in which the public sector voluntarily discloses information without any external prompting or obligation. These instruments range from responses to requests for information to judge-ordered disclosure (GPAI, 2024).

2. Research design and methodology

The study behind this report sought to answer the following research questions: 1) How are repositories built and operated? and 2) Do these instruments contribute to algorithmic transparency, and -if so- what type of value is created and for whom?

The research design to achieve these objectives involves selecting three case studies. Through these cases, we explore a range of key dimensions that enable meaningful comparisons across the examples. A case study involves an in-depth examination of a singular case's uniqueness, specificity, and intricacy (Ragin & Becker, 1992; Stake, 1995; Yin, 2018).

The research examines several key dimensions across the case studies, including:

Building, design, and challenges: Who built and designed the repository, how it was developed, challenges faced during its creation, and user accessibility.

[3] Automated decision-making systems are "any systems, software, or processes that use computation to aid or replace government decisions, judgments, and policy implementation that impact opportunities, access, liberties, rights, and safety. Automated decision systems can involve predicting, classifying, optimising, identifying, and/or recommending." (Richardson, 2022, p. 795).

Governance: The organisation's governance structure that manages the repository and the decision-making process for the inclusion of specific automated decision-making (ADM) systems.

Type of information included: Types of computational systems listed, their tasks and outputs, involvement of human oversight, metadata provided for each system, and potential for establishing repository standards.

Data lifecycle management: Rules for data management, compliance with laws, regulations, and codes of conduct, and the framework governing data use.

Users: Types of users and their feedback on the repository's usefulness, along with identification of potential new users.

Contribution: The repository's contribution to transparency, the value it creates, and the beneficiaries of this value.

KPIs (Key Performance Indicators): Use of KPIs to measure the impact and outcomes of the repository and an evaluation of their effectiveness.

The case studies (three in total) included semi-structured interviews with members of the repository: a director/implementer for each case (N=3) and stakeholders (N=6). These interviews, which lasted 55 to 75 minutes, covered the key dimensions listed above. We focused on two types of stakeholders: "contributors" (Type 1) are institutions that have contributed information to the repository, and "users" (Type 2) are institutions or individuals that make use or have made use of the repository to achieve their goals (e.g., researchers). Some questions were exclusively designed for one of these two types. These categories may also overlap (for a detailed list of interviewees, see Annex A).

To build each case, we analysed the list compiled in the previous GPAI report (2024) of online repositories of public algorithms worldwide. That report identified repositories published by 1) supranational and 2) national public bodies, 3) subnational public bodies, 4) universities, 5) civil society organisations, and 6) private organisations. A list of 83 repositories of public algorithms was built based on those criteria.

Based on that typology, we defined a set of criteria for selecting three case studies in order to answer the research questions. First, we considered "location" as a selection variable referring to the location of the repository: Is the repository located in a country in the Global North or South? Second, we focused on the "type of organisation" leading the initiative and the scope of the collected data (supranational, national, subnational): What type of organisation is leading the repository? What is the scope of the collected data? Third, the variable "governance" refers to how the repository collects information and articulates different stakeholders: Is it an organisational effort, or is the repository the result of a partnership with different organisations/government areas? Thus, the case studies analysed for this report were the Algoritmos Publicos repository, led in Chile by the GobLab at the Universidad Adolfo Ibáñez (university case); the Public Sector Tech Watch latest dataset of selected cases, led by the European Commission (supranational case); and the UK Algorithmic Transparency Recording Standard (national case) (see Table 1).

Table 1: Summary table of the three repositories.

	Chile (Algoritmos Públicos UAI)	EU (Public Sector Tech Watch)	UK (Algorithmic Transparency Recording Standard)
Institution Type	University	Supranational Public Body	National Public Body
Start year	2021	2022	2021
Simple website	×	✓	✓
User-friendly website	✓	×	×
Number of Cases	101	1666[4]	9
Number of Variables	17-19[5]	46	37-72[6]
Focus of the Included Data	Wide scope, limited depth	Wide scope, moderate depth (supplemented by in-depth interviews with algorithm developers)	Narrow scope, extensive depth
Financing	Self-financing	Public (European Commission)	Public (UK Government)
Available Formats	CSV, XLS, Website	CSV, XLS, ODS	Website

Additionally, the case studies involved a systematic analysis of secondary documents from these organisations, such as annual reports on the repositories and direct interaction with each repository using the "walkthrough method" proposed by Light et al. (2016). This method allows for a detailed description of each website, its features, design, and potential uses to trace a platform's technological mechanisms and cultural references to understand how it guides users. We analysed each case's website for a month to understand the type of collected data, the description of each repository, and the type of imagined user, as well as to present and share information about algorithms.

Using the Dedoose software, we used open coding to identify emerging themes from the interviews and subsequently categorise them into broader codes. In addition to anticipated themes from deductive queries, we uncovered novel themes, such as discussions about the challenges and impact of repositories across stakeholders and institutions. After triangulating data from the websites and case studies (interviews and secondary material), we gained a comprehensive understanding of how algorithm repositories work and their impact on enhancing transparency in different political and social contexts.

One of the limitations of this methodology is related to the type of information we access. For instance, each of the repositories provides limited information about their aims, goals, and the type of information they offer to various audiences. That is a limitation when tracing repositories' history and evolution. In relation to the case study approach, interviewing different types of users, sometimes, as researchers, we had an opportunity to understand processes, and criteria in relation to the production and use of information

[4] As of 21 October 2024, 1666 cases were registered in this repository, of which 1342 corresponded to AI systems (81%) and the rest to other types of technologies such as blockchain, virtual worlds, quantum computing, AR/VR.
[5] The number of variables present in the Algoritmos Públicos website vary slightly between each algorithm entry (16 to 17) and differ from the downloadable files for the repository (19).
[6] The number of variables varies significantly across entries, ranging from 37 in the earliest cases to a maximum of 72 in the most recent ones.

within the institutions studied. However, it was not possible to trace the complete "journey" of data collection, production, and analysis in detail. We suggest that future studies should incorporate other techniques to explore and contrast processes of data collection, analysis, and communication. For instance, a focus group should be a good technique to gain in-depth access to this kind of process and evaluation by key actors across repositories.

3. Results

3.1 Building and design

In this section, we describe the motivations and contingencies that led to the creation of the selected repositories under study (Algoritmos Públicos, Tech Watch, and ATRS). Then, we will provide a general description of each repository in terms of its design and visual presentation. On the one hand, **the motivations behind constructing these repositories were primarily linked to the importance of algorithmic transparency as a key driver and contributor to citizens' trust in public services.** This value is particularly relevant in the cases of Algoritmos Públicos in Chile and the ATRS in the UK. On the other, the primary motivation for the EU case (Tech Watch) is to establish standard benchmarks and parameters that allow comparisons between AI applications across European nations. Although the regulatory framings of each country were mentioned in all three cases, they were never mentioned as the main incentives to develop these databases.

3.1.1 Chile: History of the repository

The Repositorio de Algoritmos Públicos project was launched in 2021 as an initiative of the public innovation lab of the School of Government (GobLab) at Universidad Adolfo Ibáñez[7]. Driven by the need to highlight the various potential applications of data innovation within Chile and its public sector, the project aims, on the one hand, to promote public innovation by providing visibility into the algorithms operating within the Chilean public system and, on the other hand, to characterise their use by public institutions. Considering the high costs and low response rates from data access requests made to public institutions, the Algoritmos Públicos team decided that the data collection methodology should focus on searching public information made available through press releases, social media, websites, public accounts, public bidding processes, etc. The initial data collection identified 25 algorithms used in the Chilean public sector, with the information initially made available only through an online platform. The main sources of funding for the repository were training and consulting projects carried out by GobLab, enabling it to be self-financed. This situation remains the same at the time of writing this report, except for occasional financing opportunities derived from collaborations with external institutions.

Throughout 2021, the repository benefited from a research project conducted by GobLab in collaboration with the Transparency Council (CPLT, Consejo para la Transparencia), a Chilean public service institution created in 2008 as a result of Law 20.285 on Transparency of the Civil Service and Access to Information of the State Administration enacted that same year. This made it possible to collect data on 90 ADM systems used in the Chilean public sector via information access requests. This approach ensured high-quality data,

[7] The idea for the repository of public algorithms first came about when academics at UAI's Public Innovation Laboratory were inspired by NYU's Open Data's Impact project, which consisted of deep qualitative research on multiple open data cases. Professors figured that building a repository could be an efficient strategy to communicate the usefulness of data science in the public sector and promote its use among operators of public institutions. The opportunity to develop the repository arose when GobLab's director at the time, María Paz Hermosilla, took on a thesis request from an engineering student in 2020 aimed at data science and the common good.

which would have been excessively expensive for GobLab to obtain through other means. The study revealed that 11 of the 90 algorithms had some form of publicly available information, further facilitating access to public data sources for the Algoritmos Públicos Team.

Based on the discovery that 90 ADM systems were being used in the public sector with little public information available on their operation, the CPLT decided to advance toward transparency by designing a regulatory instrument to be implemented among public institutions. After a thorough discussion regarding whether alignment with the instrument should be mandatory or voluntary, the CPLT opted for a set of recommendations to promote the disclosure of information used by algorithms from public institutions. These recommendations were developed over two years and were finally published on August 30, 2024, in the Official Gazette through Exempt Resolution No. 372 of 2024. The Algoritmos Públicos UAI repository has served as essential support for the CPLT in promoting active algorithmic transparency by providing evidence of the types of information contained in ADM systems used within the public sector and highlighting the potential implications of their use on fundamental rights.

This repository is the only one of the three cases studied that was designed and built by an academic institution. However, the main objective of Algoritmos Públicos UAI since its conception has been to provide the public sector with a centralised platform that gathers information on how algorithmic and AI instruments have been implemented by public institutions across the country while simultaneously serving the purpose of evidencing how data science can improve public management services. The Algoritmos Públicos website states that the repository is one of GobLab's initiatives to "promote the responsible and transparent use of algorithms, artificial intelligence, and automated or semi-automated systems in the Chilean public sector, aiming to support and drive innovative public institutions" (GobLab Universidad Adolfo Ibáñez, n. d.). Aside from the repository, the website also offers a Tools for Ethical Algorithms section that comprises:

> A series of guides, products, and regulations created by GobLab UAI with the support of multiple national and international organisations to support the incorporation of responsibility standards throughout the entire lifecycle of a data science project: formulation, design, implementation, and evaluation (GobLab Universidad Adolfo Ibáñez, n. d.).

The project behind the Tools for Ethical Algorithms is the most recent addition to the Algoritmos Públicos website, and the team behind GobLab is currently planning to integrate the Repository and the Tools for Ethical Algorithms into one unified platform to contribute to algorithmic transparency.

3.1.2 Chile: Repository design

The repository aimed to register as many operating ADM systems as possible, opting for a wide arrangement of algorithms instead of relying exclusively on artificial intelligence systems. To achieve this, the project comprised four key tasks: creating a taxonomy for the selected cases, identifying and collecting relevant information, generating content for each case, and developing the webpage through which the information would be accessed. The criteria for including an algorithmic tool in the repository were: 1) the presence of an automated or semi-automated decision system, 2) currently implemented in a Chilean public institution, and 3) having at least one source of information about the tool from the institution itself.

The original objective of the repository was to generate a practical impact on public administration innovation and policies. A second goal that eventually became evident to the team was to contribute to algorithmic transparency in the public sector, as this information was already public, and gathering the

different cases in one virtual repository would contribute to more democratic and accountable public services.

Even though the repository was initially conceived for public sector operators and academics as the primary users, the information it contains is available to the general public on the website. The decision to make the repository public was made by the GobLab team to fulfil the second objective of building the repository — to contribute to algorithmic transparency.

The repository website stands out for its largely transparent and user-friendly design, miming the colourful overlay, readable text boxes, and accessible language of NYU's Open Data's Impact. The interaction required to access the repository's information is minimal, with all projects gathered on a single page of the site, which is also visually easy to digest due to the added graphic elements (images, colour, and text size). The categories for filtering projects focus on facilitating work with the public sector, even though the design reflects a clear interest in making the information accessible to any user. Each entry in the repository includes 17 variables regarding these algorithmic/AI tools. Additionally, the repository includes annual reports that provide both descriptive and analytical information about the projects included in the repository and the state of algorithmic visibility at a national and international level. In this regard, the repository is a solid contribution to algorithmic transparency in the public sector in Chile. The site also features a More Information section describing the team behind the repository. However, certain key information, such as the project's start year or its funding, is not disclosed at any point. The information and data provided in the repository are available in Spanish.

Figure 1: Algoritmos públicos homepage

Source: Algoritmos públicos repository

3.1.3 EU: History of the repository

Throughout the second half of the 2010s, the Joint Research Centre (JRC) of the European Commission developed a set of repositories on various topics, including a Citizen Science repository in 2016 and the Innovative Public Services Observatory in 2018. In 2019, the Directorate-General for Communications Networks, Content, and Technology (DG CNECT) and the JRC collaborated to establish the AI Watch Observatory, which provided information on multiple AI technology initiatives in the EU until 2021. The original AI Watch website introduces the reader to the context in which the project originated by explaining the Coordinated Plan on Artificial Intelligence published by the European Commission and the Member States on December 7, 2018, which aims to monitor AI uptake and development in the region by establishing standard benchmarks and parameters that allow for comparison between the nations' institutions. The site then indicates the role of AI Watch within the Coordinated Plan, stating that:

> [It] will monitor the implementation of the Coordinated Plan including strategies and investment. From these in-depth analyses, we will be able to better understand Europe's areas of strength and areas where investment is needed to boost AI in Europe. AI has a wide range of potential economic and social implications including new forms of economy and governance. AI Watch will provide an independent assessment of the impacts and benefits of AI on growth, jobs, education, and society. (European Commission, 2021)

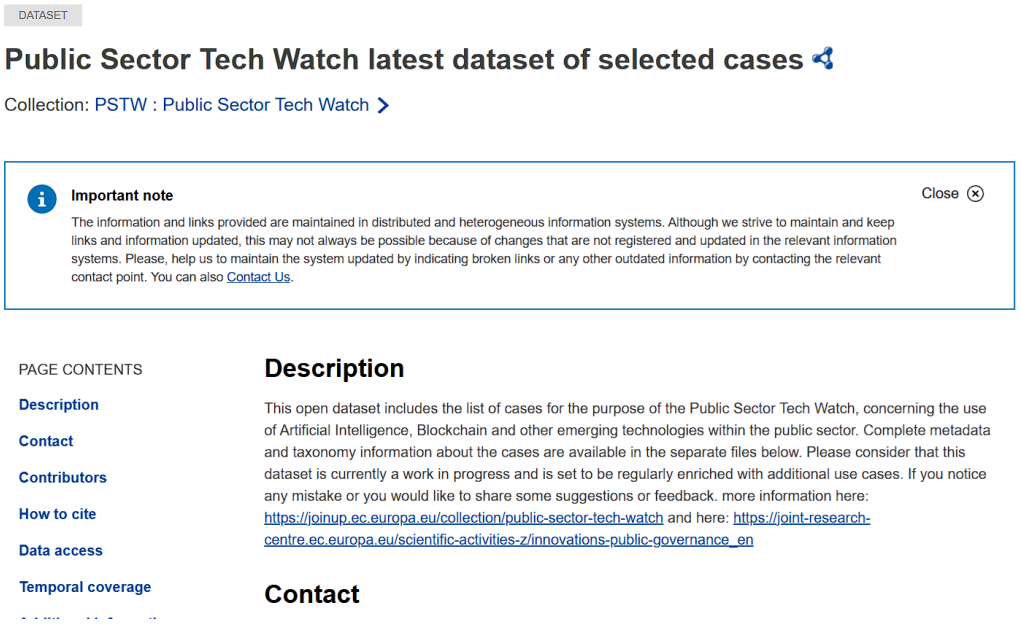
Based on these experiences, the Directorate-General for Digital Services (DIGIT), together with the JRC, decided to create Public Sector Tech Watch in 2022. This new observatory was set up to provide streamlined and central access to cases in which public sector institutions in the EU use emerging digital technologies. It concentrates on the public sector within the region and expands its scope beyond AI by also including Blockchain, virtual worlds, Augmented Reality, Virtual Reality, Quantum Computing, and other emerging technologies at the time of writing. As described on the Public Sector Tech Watch website, the project "aims to become the 'one stop' for all stakeholders – public sector officials, policy makers, private companies, academia – interested on the latest technological developments to improve public sector operations and service delivery" (European Commission Interoperable Europe, n. d.). Hence, the

web presence does not only provide the repository itself, but also supports community building, sharing of related materials, and dedicated events are organised both online and in person.

3.1.4 EU: Repository design

The European Union's Public Sector Tech Watch successfully contributes to algorithmic transparency by presenting relevant information on 1666 initiatives[8] across the Union's 27 member states and other European countries. The methodology implemented for case collection in the Public Sector Tech Watch observatory, including details about the included information for each case, is well documented (Tangi, Combetto, Martin Bosch, 2024)). The underlying data can be downloaded from the PSTW portal and directly from the JRC. The JRC website from which the repository's dataset can be accessed (Public Sector Tech Watch latest dataset of selected cases) is limited to presenting information through text using only blue as an emphasis colour, which allows clear reading but is less visually appealing or interactive compared to platforms such as Algoritmos Públicos UAI. It offers two different formats for the database (CSV and XLS) and provides the option to visualise them through the Cases Viewer and Statistics dashboard, a third-party interactive interface. The dashboard can be accessed through the Public Sector Tech Watch page at the Interoperable Europe website. The website is originally provided in English, but it can be automatically translated into the 24 official languages of the EU, as well as Russian and Ukrainian, which might boost its accessibility. The database itself is not equipped with a translation from English.

Figure 2: Public Sector Tech Watch Selected cases in the public sector webpage



Source: Tech Watch Repository

3.1.5 UK: History of the repository

[8] The number of cases corresponds to the date the website was revised (October, 31st 2024). Also, these initiatives are not only AI systems, but also including Blockchain, Virtual Worlds, Augmented Reality, Quantum Computing, and other emerging technologies by the time of writing.

The UK Government's Algorithmic Transparency Recording Standard (ATRS) originated from a budget assigned to the Department for Digital, Culture, Media and Sport (DCMS) to research new potential opportunities related to AI and data-driven technologies. In a national context where public concern about the implementation and ethics of algorithmic tools started to rise (for instance, the 2020 student protests caused by a controversial algorithm employed to calculate A-level grades), the DCMS proposed that a standard be designed and implemented to promote transparency. A standard was developed by the Centre for Data Ethics and Innovation (CDEI), which consequently became the ATRS. Administration of the ATRS project later shifted governance from the DCMS to the Responsible Technology Adoption Unit (RTA) of the Department for Science, Innovation and Technology.

In September 2020, the UK Government published a National Data Strategy (NDS)[9], which suggested, for the first time, that information on algorithms employed in the public sector should be compiled and made available to the public. As it states on the website, "The NDS is a framework for the action this government will take on data. It is not the final answer, but part of a conversation about the way we support the use of data in the UK" (UK National Data Strategy, 2020). Following this, in 2020 and 2021, the CDEI conducted a public engagement study in collaboration with research firm BritainThinks to identify the best algorithmic transparency strategies to improve public understanding of and trust in these technologies. The results of this research are summarised in a review in bias (BritainThinks, 2021), which provided insight into what type of transparency the public expects from algorithmic tools and became the basis on which the Algorithmic Transparency Recording Standard was constructed:

> One of the recommendations this report made was that there should be a transparency mechanism for public sector organisations to be open and transparent about algorithmic tools that they were using in processes that affect members of the public or sort of decision-making processes. And that recommendation essentially was taken up, and the ATRS, [which] essentially builds on or responds to commitments made in a few different government policy papers. So, for example, the National Data Strategy in 2020 would, I think, be one of the earlier places where a recommendation was made, and then much more recently in the AI white paper in 2023. (ATRS implementer)

Regarding the objectives of the ATRS repository, the ATRS Hub webpage states that "the Algorithmic Transparency Recording Standard helps public sector organisations provide clear information about the algorithmic tools they use and why they're using them". The goals of the ATRS explicitly comply with the 2020 National Data Strategy:

> The strategy has a commitment to explore an appropriate and effective way to deliver greater transparency on algorithm-assisted decision making in the public sector. The National AI Strategy reiterated this commitment, with an action to conduct research that will help develop a cross-government standard for algorithmic transparency. (UK Government, n. d. a)

Therefore, the main goal, in this case, seems to be inscribed in institutional mandates that originated from a commitment to algorithmic transparency and an ethical implementation of algorithmic tools.

3.1.6 UK: Repository design

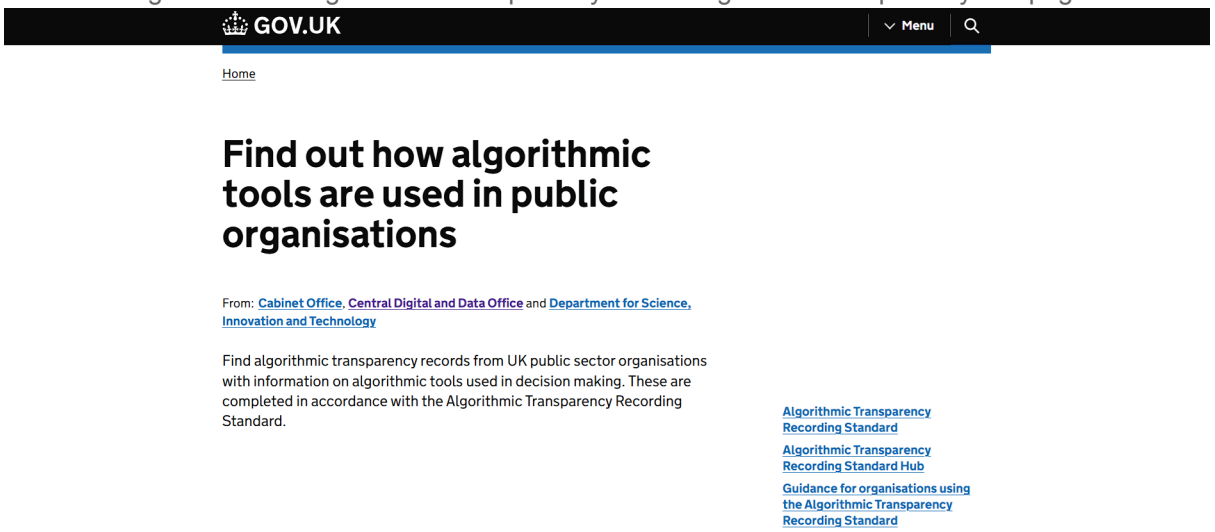
[9] Available at https://www.gov.uk/government/publications/uk-national-data-strategy

Compared to the other repositories, the Algorithmic Transparency Recording Standard offers a small selection of cases, but they are described in much greater depth on an extensive page devoted to each algorithmic tool. The provided information includes details about the organisations managing the tool, their objectives, development, decision-making processes, modelling and maintenance, used databases and their performance, and potential risks and impact assessments. In this sense, this repository focuses on the depth of the available data rather than on large quantities of cases:

> We have aimed for the depth of information, I suppose. So although I think the State, for instance, has logs or collections of a lot more algorithmic tools across its federal agencies, it offers quite minimal information about them, whereas our communications tool offers quite a lot of more detailed information, which for most people is probably not interesting, but if they wanted to understand, it allows them to look back on the process and understand who they've collaborated with to get to the point of use, and understand where the data has emerged from, and understand what the thinking around the use of the tool was. (ATRS implementer)

This is also evident because the site does not offer downloadable files or visualisations that compile systematised information about each tool, even though every record is built according to the same standard. Most of the tools offered on the page, which are very clear and thorough in their explanations, aim to help UK public bodies complete the ATRS form, suggesting that these public offices are directing their efforts toward including more algorithmic tools in the repository. The algorithmic transparency they seek appears to be driven by a practical interest in the public sector, as well as a democratic interest in line with the right to information. The website's data and information are provided in English.

Figure 3: UK's Algorithmic Transparency Recording Standard repository webpage



Source: ATPS Repository

3.2 Main Challenges faced by the repositories

The main challenges identified throughout the study relate to the processes of designing and managing repositories. Information on algorithms employed in the public sector remains a relatively opaque matter, and bureaucratic requirements for accessing information about public institutions were often stated as a significant obstacle to incorporating algorithmic tools into the repository. This, in turn, results in additional time and resources being allocated to the necessary procedures, as periodic meetings with contributing institutions need to be held to verify whether the information is being provided correctly. In direct relation to this issue, it was found that the areas most affected by the resource constraints faced by the repository teams are, on the one hand, the rigour with which algorithmic systems are classified and, on the other, the depth of the gathered information. Regarding challenges faced by users, the main issue identified is that, occasionally, information provided by repositories will need to be outdated and might represent a roadkill for scholars and journalists. (See figure 4).

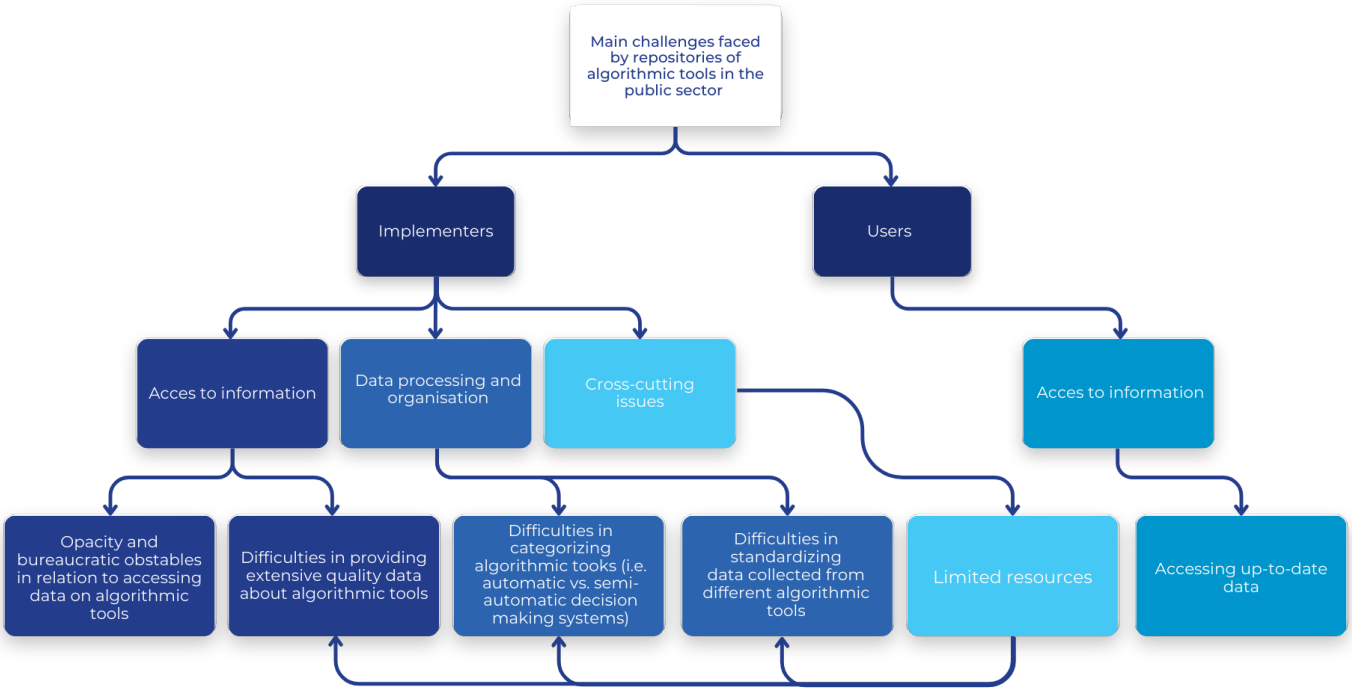


Figure 4: Main challenges associated to repositories of algorithmic tools in the public sector

3.2.1 Chile: Challenges faced by the repository

Regarding the construction of the repository, interviews showed that the main challenge faced by the team was a lack of resources and personnel for the project, which rendered specific tasks even more difficult, such as defining the categories the repository would include. This led designers to improvise within their capabilities to achieve the desired results:

> Taxonomy has been a big issue, and it continues to be in this project, like trying to understand what kind of taxonomy really matters, because we didn't want to create one from scratch. So, we tried to see which existing taxonomy we could align with. That was a major topic -deciding which fields we were going to collect data on for each project. We missed something there... Look, I would have always liked to have, maybe, a professor who was an expert in public administration or something like that available, but we had limited resources, and it was what it was. In the end, none of us [on the team] were experts in research methodologies, so we ended up structuring this in a way that was perhaps a bit more sui generis than what one would have preferred. (Algoritmos Públicos UAI implementer)

Another example of how limited resources impact the scope and depth of the information included is that the Editorial Committee[10] of the repository operates on a voluntary, *ad honorem* basis. As a result, its members can devote limited time and effort to identifying and selecting cases and creating their corresponding entries for the repository.

Interviewees noted that resource constraints are also evident when hiring rotating assistants as the primary information gatherers. This has led to inconsistent data collection, as the methods for storing and analysing the information differ from one assistant to another. Consequently, the team has been working to correct

[10] See 3.3.1. Governance: Universidad Adolfo Ibáñez's Repositorios de Algoritmos Públicos.

mislabeled information, such as algorithms wrongly classified as artificial intelligence, even though they do not meet the criteria for that category. From the users' perspective—such as researchers, academics, and journalists—the repository represents a valuable resource for accessing data from public bodies that use algorithms. It is particularly useful for those interested in mapping and exploring these cases, offering a critical tool for analysis and investigation.

Building a case from the various information sources available has been identified as a time-consuming challenge for the Algoritmos Públicos team. While the implementers of the repository may be aware of an algorithm's existence through press releases or public announcements from institutions, a case necessitates that robust evidence and data be publicly available before it can be included as an entry in the repository. In many instances, the publicly accessible information is limited or requires the submission of a formal information access request before such data can be disclosed to the team, complicating the process of constructing a case. The process lacks a defined protocol; decisions are brought to the repositories' committee, and if all members agree, they are included. According to the repository implementer: "There has never been a case where one person objects while the others approve."

3.2.2 EU: Challenges faced by the repository

The main challenge for the EU repository is the difficulty of raising awareness of Public Sector Tech Watch among public administrators and expanding its reach to new audiences and interested stakeholders.

> I think it's like how we chose the cases. It's very difficult to… to have an overview, or, you know, a detailed description with everything, with interviews of each case. We try to focus on the cases that matter the most, or we think at least that matter the most, to invest there. The other is, of course, to reach out and disseminate what we are doing to public administration. (...) Another is, of course, to bring more people not only for the cases, but also to read our news or also to provide their news, to serve their news with us. (Tech Watch implementer)

Interviews with the repository users also evidenced occasional shortcomings regarding outdated information from specific algorithmic tools. This problem will be addressed in section 3.6.2. EU: Types of Users and Usage.

3.2.3 UK: Challenges faced by the repository

Reported challenges for the ATRS team were mostly related to regular procedures regarding the new governance processes that the Standard introduced. The task of periodically contacting organisations to request information that requires bureaucratic approval was mentioned as a particularly time-consuming job, especially on a wider scale of the entire UK's public institutions, as each one usually has its own procedures for disclosing data. The team has struggled with orchestrating the different components needed to build and publish each record in both the front and back end of the process. Operators are responsible for requesting information on each algorithmic tool while also being positioned at a lower level of the governance structure. This dual role requires them to maintain ongoing communication with institutions to prompt the necessary actions from their superiors, enabling the authorization and release of data:

> "When something like this is still in its relative infancy as a mandate, there are also challenges for organisations in getting that information cleared and signed off (...). We think of ATRS records as communications products. So we expect the sign-off procedure to be similar to if you were

> producing a press release or a blog post about that tool. But typically, something like that in government, you might go through a comms clearance, you might take it past a legal team to make sure they were happy, you'd probably take it past a policy team, and the appropriate senior people would sign it off as well." (ATRS implementer)

The team expects a more streamlined flow of information with the various institutions involved in the ATRS repository in the future:

> "Currently we are updating records if teams tell us that something has changed (e.g. the tool is retired or if content for the questions in the standard has changed). This is a manual process currently - we upload updated content onto the Gov.uk repository. In the future, we may be looking into automating the process through a platform with more functionalities, but this is currently not possible." (ATRS implementer)

The amount and type of information that the repository can collect have also been mentioned as a challenging aspect of the project. Though the team aspires to compile as much data as possible regarding each algorithmic tool, some of it might compromise national security or the integrity of public sector services:

> "The more sort of salient challenge for us at the moment… There are of course whole categories of tools, or whole categories of information within tools that we wouldn't expect organisations to be transparent about, for example, for national security reasons. And again, this is covered in our scope and exemptions policy [...]. Other concerns might be around, you know, ability to game systems, but we think that the ATRS has been designed and developed with those concerns in mind, and with the ability for organisations to mitigate those concerns proactively." (ATRS implementer)

3.3 Governance structures

Information on the repositories' governance structures for this study has mostly been accessible through interviews. The only website that provides some information on the governance structure of the repositories is UAI's Algoritmos Públicos, which details the members of both its editorial committee and the repository's team, even including a history of previous collaborators. The UK's ATRS and the EU's Tech Watch specify the institutions responsible for managing the repositories and allow website visitors to contact the teams via email. However, they do not provide details about the governance structure or how tasks are distributed within the managing institutions.

Interviews with members of the implementing organisations shed light on the different institutions managing the repositories. Government-managed projects like the ATRS or Tech Watch have a history of collective efforts from different public bodies, whereas the only repository developed by an academic institution, UAI's Repositorio de Algoritmos Públicos, appears to function more independently but still benefits from the interest of its external collaborators.

For all the cases studied, data collection and analysis processes are very simple (they collect and treat data manually) in order to build a case for the repository. The pro of this approach is that repositories' team members develop knowledge in order to trace data and compile it to share it publicly. The con of this

approach is the lack of automation of data collection, which translates into the impossibility of developing a constant workflow of updated cases online.

3.3.1 Universidad Adolfo Ibáñez's Repositorio de Algoritmos Públicos

At the time of writing, the Algoritmos Públicos team is managed primarily by GobLab members but relies nonetheless on an Editorial Committee that holds monthly meetings to decide whether new algorithmic projects should be included in the repository. Members of the Committee participate voluntarily after being invited by GobLab operators. The Committee comprises GobLab and external researchers, including members of public institutions such as the CPLT and the Digital Government Secretariat (Secretaría de Gobierno Digital). The Committee also holds a spot for a research assistant, who is typically a UAI postgraduate student. Research assistants usually last approximately 3 to 6 months in the role and leave the team once their studies have finished.

The Editorial Committee and research assistants are responsible for gathering primary public information on potential algorithms for inclusion in the repository. Subsequently, the research assistants further develop the case by seeking additional sources of information to construct a more robust algorithm profile. If sufficient data is available, the case is reviewed during the monthly meeting with the Editorial Committee, where it is decided whether the algorithm should be included in the repository.

One designated investigator is in charge of producing the annual report published by Algoritmos Públicos, which involves thorough research on the current state of algorithmic transparency in Chile and South America, as well as deep qualitative analysis of the cases that comprise the repository. However, the Editorial Committee plays a key role in its construction by proofreading, commenting on, and suggesting changes to the reports.

3.3.2. European Commission's Public Sector Tech Watch

The Public Sector Tech Watch repository is the responsibility of the Directorate-General for Digital Services (DIGIT) of the European Commission, whereas the scientific lead remains with the JRC. Although DIGIT administrators have the final decisive authority to include information in the Observatory, Public Sector Tech Watch operators describe its governance as a collaborative effort between DIGIT and its stakeholders, such as contractors and private companies, as they will often submit information and start discussions before they are posted. In this sense, Public Sector Tech Watch is both an active and receptive initiative insofar as it intentionally searches for pertinent cases for the repository and simultaneously allows third-party institutions to contribute to its database.

3.3.3. UK Government's Algorithmic Transparency Recording Standard

The UK's ATRS is managed by two teams: the Responsible Technology Adoption Unit (RTA) and the Central Digital and Data Office (CDDO), both of which are part of the Department for Science, Innovation and Technology (DSIT) within the UK Government. These teams collaborate so closely on the project that their members often do not find the distinction between them particularly significant. While the RTA primarily focuses on designing, collecting information for, and delivering ATRS-related products, the CDDO provides support on data ethics issues. Both units have been involved in the design and deployment of the ATRS since its inception and continue to work together closely.

3.4. Type of information included

While some basic information is present in all three repositories, such as institutions in charge, temporal and regional coverage, or a general description of the algorithms, the depth and detail provided about these tools vary greatly from one case to another. For instance, UAI's repository covers 19 general variables, while the UK's ATRS repository offers data on 72 variables in a wide spectrum of depth regarding the tool and model specifications. Regarding relevant meta-data that are unavailable in the studied repositories, none of them showcases the source code included in the algorithmic tools. As mentioned above, reported reasons for this include the public risk of sharing critical information that hackers could exploit to breach system security or disrupt its proper functioning. The repositories also lack variables that specify whether humans control the systems included and, if so, at which stages of implementation this control occurs. As stated in the state-of-the-art report (GPAI, 2024), the inclusion of variables that track "humans in the loop" is still limited mostly to repositories developed in the Netherlands and a few exceptions in the UK and Brazil.

Based on the classification of variable types established in the state-of-the-art report (GPAI, 2024), only four types of variables are consistently present across the three repositories studied: Information on the status of the system and its building process, the name of the system; the system's objectives, tasks, and outputs; and the unit in charge of deploying the system (see Annex B). Following this classification, which is structured around the five stages of the data life cycle defined by Denis et al. (2021), the Chilean and UK repositories include information on the first phase of Planning and Design (three variables each, see Table 3). Regarding information available about the data collection and processing stage, the three repositories seem to have a similar amount of information between one and two variables. The ATRS repository stands out for covering the most variables regarding the Model building and validation stage, with six variables compared to three in the Public Sector Tech Watch repository and only one in GobLab's Algoritmos Públicos. The Deployment and monitoring stage is where the three repositories share the most types of variables available: three in the Chilean and European case and five in the UK case. Finally, the Accountability stage is covered by one variable in the Chilean and European repository each, and two variables for the UK repository. With this in consideration, the stages that are covered most often are Planning and design, Model building and validation, and Deployment and monitoring.

Table 2: Number of Variables Covered in Each Repository by Stage of the System's Life Cycle[11]

Stage of the system's life cycle	Chile (Algoritmos Públicos UAI)	EU (Public Sector Tech Watch latest dataset of selected cases)	UK (Algorithmic Transparency Recording Standard)
I. Planning and design (6)	3	0	3
II. Data collection and processing (5)	1	2	2
III. Model building and validation (10)	1	3	6
IV. Deployment and monitoring (11)	3	3	5
V. Accountability (4)	1	1	2

[11] Cells highlighted in light blue indicate a high coverage of a stage in the corresponding repository.

3.4.1. Chile: Type of information included

The Chilean repository's website collects and publishes information on the following 17 aspects of each algorithm or AI: name of the project, Sustainable Development Goal (SDG) it aligns with, DIPRES classification, public institution, institution unit, executors, region, objective, whether it uses personal data or not, whether it is considered AI, strategic product/support, description, launch date, known status, funding, source, and source links. Additionally, the Excel document available on the website's download page provides two more variables (19 in total): the amount of monetary resources destined for the projects and the corresponding currency. From the three case studies, UAI's repository appears to provide the most general-level information on the included algorithms, categorised into a restricted number of variables. GobLab researchers have mentioned that the selection of such limited criteria is a consequence of the type of information that is typically available through public sources. However, this allows for a wider arrangement of available cases.

No information is provided about whether the code that makes up these tools is open or not, or if it is accessible in any way. According to the Algoritmos Públicos 2024 Report:

> The Repository does not provide information about the code used for modelling, as it is not usually available. In many cases, moreover, there is no detailed background information on the specific technique used in the data analysis, which would be of great interest for this report and the Repository. (GobLab Universidad Adolfo Ibáñez, 2024)

3.4.2. EU: Type of information included

The variables on which the EU's database is built are name, website, description, scope (national, regional, local), country of coverage, NUTS 2021 category to which it belongs, responsible organisation, responsible organisation category, functions of government (COFOG levels I and II), implementation status, start and end years, AI classification, subdomains of AI classification, AI keywords, type of process, type of application, technology, interaction, data scope, improved public services, improved administration (improved administrative efficiency and open government capabilities), source, AI classification (I), AI classification subdomain (main, others I, others II, and others III), AI keyword, and date updated. In this regard, the database contains over 40 variables that are general and understandable to all audiences yet also specific to the assessment of public policies. The complete details, such as the taxonomy designed to categorise collected cases, are publically available on the public body's platform (Tangi, Combetto, Martin Bosch, 2024).

3.4.3. UK: Type of information included

The UK's Algorithmic Transparency Recording Standard is designed to collect deep qualitative data on over 70 different variables concerning, for instance, the following tiers of information (which vary depending on each tool): 1. General information, 2. Owner and responsibility, 3. Description and rationale, 4. Decision-making process, 5. Tool specification, 6. Technical specification and data, 7. Data specification, and 8. Risks, mitigation, and impact assessment. As mentioned above, most of these variables consist of at least one paragraph detailing information relevant to each aspect of the tool. These categories aim to provide information that is both easy to understand for the general public and specific enough for researchers, journalists, or other technology and public service developers.

3.5. Data lifecycle management

The collected information indicates that the primary strategies for accessing data on algorithmic tools involve, on the one hand, desk research by the repository teams, which includes reviewing press releases, social media, and official documents from public institutions, and, on the other, direct contact with the responsible organisations, whether initiated by the repository team or the system implementers. In all cases, the developers of the repositories have designed mechanisms for public institutions to submit data on their algorithmic tools independently to facilitate the collection process. Implementers of the repository will mostly focus on searching for new algorithmic tools, which makes updating information a secondary priority.

While in the Chilean and European repositories information is collected and mainly organised by the implementers, the data for the UK's ATRS is first filled out by the developers of the tools and later verified and officially published by the ATRS team. Repository operators usually carry out the process of selecting new data and updating old data on different algorithmic tools, such as information on their status, after discussions in recurring meetings. Regarding the final stages of data lifecycles, there seem to be no established protocols for data deletion in any of the three scenarios.

3.5.1. Chile: Data lifecycle management

As mentioned above, the algorithms included in UAI's Algoritmos Públicos are mostly identified through press articles, news, and media from the institutions that implement them. Assistants of the repository then write preliminary forms with basic information about newly identified algorithms in the public sector, which will be discussed later by the editorial committee, which decides whether the project is admissible for the repository or not. This information is stored in an Excel sheet, where the data on the algorithmic tools is categorised as "accepted", "rejected", and "under observation". The last category is reserved for algorithms that have not yet been implemented but have been officially announced by a public institution.

3.5.2. EU: Data lifecycle management

The repository operates by collecting information on cases from press releases and official institutional websites, allowing public administrators to provide information on the tools themselves and further elaborating on selected cases -depending on particular areas of interest, such as interoperability or generative AI- stories and the rationale behind them through interviews with the public administrators or developers of the tools (Tangi, Combetto, Martin Bosch, 2024). The team behind Public Sector Tech Watch is constantly incorporating new emerging technologies into the repository and updates the website every month. Moreover, the repository provides a survey through its webpage for tech developers to submit information about their tools, which is later evaluated and included in the repository.

> So, one part is a collection of information about this case. Okay, there are three actually: one is this [press releases]; the other is that people can submit their cases, like public administrations or even the private sector through the public administrations that implemented the solution. This is the second, and the third is that we further elaborate on specific cases. So we do extra research and we provide some stories explaining and actually describing and presenting 'what is this case about?' (Public Sector Tech Watch implementer)

3.5.3. UK: Data lifecycle management

The information collected by the ATRS for each record is acquired through direct contact with the institutions in charge of the algorithmic tools, which is achieved after the team actively searches for and finds technologies that fit the scope or after institutions find the repository through the ATRS' multiple efforts to promote the project via official announcements, talks, or government blogs.

The process of building a dataset involves the participation of public operators from the institution responsible for the algorithmic tool, who fill out a form designed to facilitate the process of systematically organising the data. Then, the ATRS team verifies the quality of the received information and publishes the record:

> There's a drafting process where they go away and they fill out the information. This is generally like a multi-person effort because of the quite broad scope of the information we asked for in the Standard (...). They may need to lean on a lot of different people to access the information required by the Standard. So, they'll do that, it usually takes a few weeks, possibly months… they'll come back to us with a record, at which point we quality assure it and make sure that it's in a state that people can understand it and there's not… any sort of informational gaps or things that could be misconstrued. (ATRS implementer)

At the moment, since the project is still at an early phase of its implementation, there appears to be no official process for deleting the information contained in the records.

Table 3: Construction and Operation of the Repositories

	Chile (Repositorio de Algoritmos Públicos)	EU (Public Sector Tech Watch latest dataset of selected cases)	UK (Algorithmic Transparency Recording Standard)
Data Lifecycle Management: *Collection*	• Press releases • Social media • Websites • Public bidding lists	• Press releases • Submissions by developers and public sector operators • Interviews	• Direct email contact with public institutions • Quarterly algorithmic tools mapping report from each UK ministerial department
Data Lifecycle Management: *Updating*	New data found on previously covered algorithms is discussed and ultimately approved in the monthly meetings of the Editorial Committee.	The Public Sector Tech Watch team will get updates through web search as well as the website's submit survey on information that needs to be updated.	Records are updated manually based on team notifications (e.g., tool retirements or content changes in the standard). Updates are uploaded to the Gov.uk repository in markdown. Future automation may be explored, but it is not currently feasible.
Data Lifecycle Management: *Case building*	Research assistant collects additional information and connects data to build a case for each repository. Each	For some particular cases (selected for topics of particular interest), members of the Public Sector Tech Watch team	ATRS implementers guide public sector operators to fill out the Standard form. Operators will collect the data from different

	case is then evaluated by the Editorial Committee to be published in the repository.	carry out interviews with public administrators or developers and operators of algorithmic tools to provide deeper information beyond the data provided by the initial collection process.	sources and send it back for a final verification from the ATRS team before publication.

3.6. Types of users and usage

In general terms, the repositories have been used mostly by researchers and public sector operators to find detailed information about specific algorithmic tools or find an arrangement of projects related to their area of interest. Users indicate the repositories as a starting point so they can then contact the developers of the projects personally to further expand their knowledge about their objectives, modelling, or implementation. Some users have pointed out that the entries can occasionally need to be outdated, which might lead the user to trace currently existing projects and stumble upon dead ends.

Four types of users and four different uses have been identified for the repositories, which vary in frequency: Public operators, who visit the repositories to find information that can inspire new public policies; academics, who employ the repositories as part of their research processes and to promote algorithmic transparency; journalists, who search for information about algorithmic tools for press articles; and university professors, who utilise the repositories as a pedagogical tool to teach the potential of data science in the public sector (see figure 5).

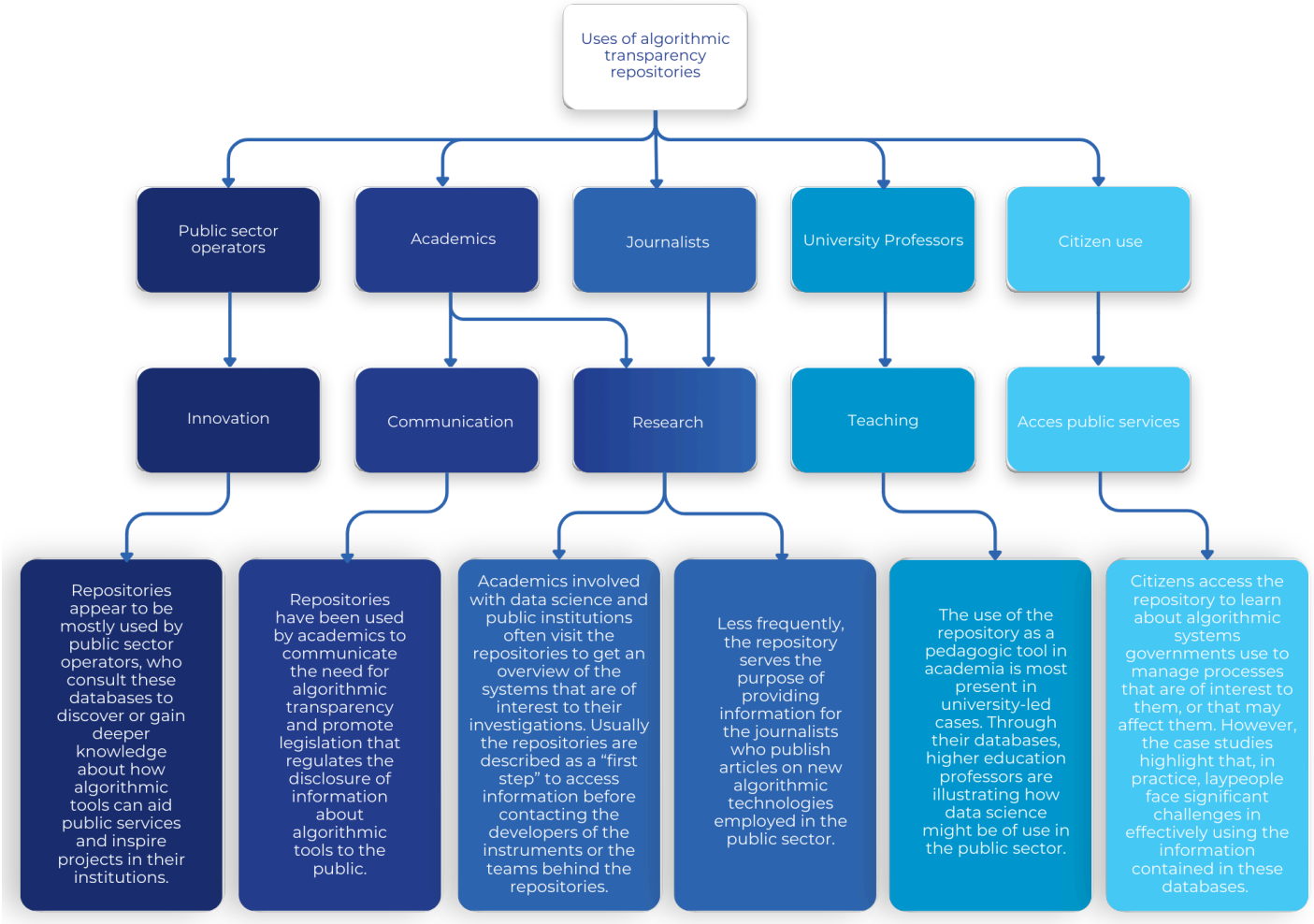


Figure 5: Uses of algorithmic transparency repositories

3.6.1. Chile: Types of users and usage

The main type of use identified for the UAI repository is in relation to academic research. Since the repository acts as a centralised and accessible website that provides an overview of different public sector policies and projects incorporating AI into their implementation, academics have acknowledged that Algoritmos Públicos is the only consolidated repository where information on currently implemented artificial intelligence tools in Chile can be found. The repository satisfies the researchers' need to collect information about multiple and varied AI-related cases in the public sector from one source. The repository was mentioned as a starting point for deeper research:

> Personally, I do a lot of user research, and for that, I need specific cases. A system is required where people have to make particular decisions, and so, I often used it to define what my case studies would be depending on the hypothesis I have. I frequently navigate by domain or by the Sustainable Development Goal they are aiming for in order to find cases that are interesting to continue analysing. These are generally my entry points into a system. (Algoritmos Públicos UAI user)

Implementers also indicated that the existence of the repository has proven useful when discussing the need for transparency in the public system with public institutions since it serves as detailed evidence of how algorithmic systems are being employed in the Chilean public sector and the fundamental rights they might put at risk without proper regulation.

The public sector also benefits from UAI's Algoritmos Públicos, as public administrators and service developers access its database to discover information on related technologies and contact the implementers of these systems. In this sense, the Algoritmos Públicos team has often taken on the role of mediator between public sector operators:

> We've also had cases where they look for ideas there, so they've written to us on YouTube asking 'Do you know the person who created such-and-such system?' and we're like 'No, we don't know them,' or sometimes we try to see if we know someone in that institution to try to connect them. For example, we had a case where a woman from the Melipilla hospital got in touch, and we said 'Maybe, if you're interested, we could help you get in touch.' Essentially, someone said 'I want to do this' and ended up meeting with the team of the University that had helped another hospital create that system. (Algoritmos Públicos UAI implementer)

Another reported use of the repository is by journalists who have utilised it to investigate how automated decisions are made in the public sector. Additionally, the repository has found a pedagogical application in the context of its academic emergence, as professors at UAI have used its cases as examples in their courses on digital democracy. It has also had an international impact, serving as a model and inspiration for other public repositories in Mexico (Riquelme, 2023) and Colombia (Gutiérrez and Muñoz-Cadena, 2023).

3.6.2. EU: Types of users and usage

Implementers of the Public Sector Tech Watch repository have stated that the most common use associated with it is for public sector operators to find emergent technology solutions that can inspire their projects or suggest alternative answers to public issues related to their areas. At the same time, the repository provides a rich knowledge base for the JRC scientists to carry out more dedicated research on topics of particular interest.

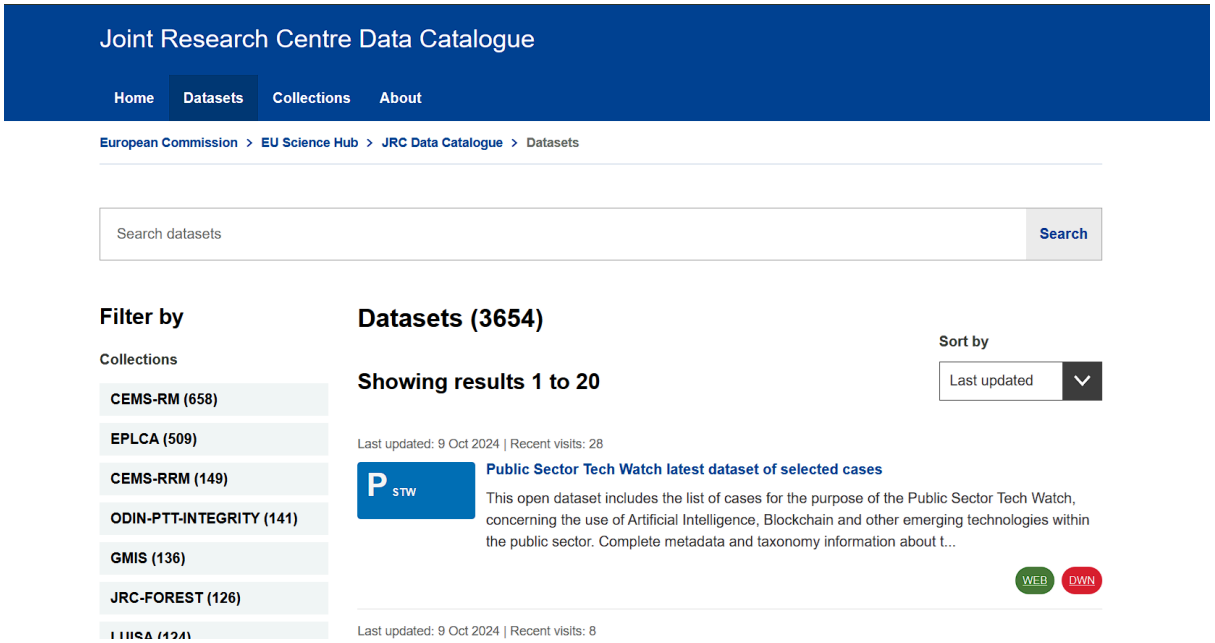
Interviews with users have evidenced a concern regarding accessibility issues that the site might have for ordinary citizens. It was mentioned that the information is too complicated to be easily understood by inexperienced users and that the only way it can interact with civil society is through third-party organisations.

> I'm not sure how easy it is for an average citizen, an ordinary citizen, to go to the EU Commission website, look at the repository, and understand what is written on the repository… Even though right now it's definitely more user-friendly, I still think that probably ordinary citizens are cut out of this repository, unless there are civil society organisations in between. (Public Sector Tech Watch academic user)

Although the repository's page at the JRC's data catalogue site contains a large amount of useful information, a relevant issue that has been identified seems to be how hidden it is for inexperienced users. To access it, a user must visit the JRC's data catalogue site, from where it is necessary to navigate to the Datasets section and manually search for Tech Watch in the search bar or, alternatively, browse through all

3654 collections available at the centre. From this point, the database file must be downloaded. Currently, however, the Public Sector Tech Watch dataset sits at the top of the search list since the search system first displays the most recently updated datasets, making it easier to find. This might vary over time, nonetheless, depending on which datasets are being updated at the time of visiting the website.

Figure 6: Joint Research Centre Data Catalogue website



Source: JRCD webpage

This repository is the only one from the three cases that have presented inconveniences to users regarding its use. An interviewee indicated difficulties in tracking down a project they were researching when they discovered the project had only reached piloting status and was never developed beyond that point:

> For me, I mean, I don't know. I just have this sense of trust in institutions. And so for me, it was quite reliable, but once I conducted fieldwork and I realised that one role of the database was, as I mentioned, not even ever implemented by the Italian Anti-corruption Authority, I started, like, taking it but with, you know, being aware of the… of some gaps it might have. (Public Sector Tech Watch academic user)

3.6.3. UK: Types of users and usage

The Guidance for Public Sector Bodies available at the ATRS Hub explicitly states the intention to facilitate information for the "general public" and "interested stakeholders" regarding the algorithmic tools used in the United Kingdom's public sector: "These records will be uploaded to our GOV.UK repository, where they will be accessible to the general public and interested stakeholders."

Due to the way it must be accessed, it appears to be exclusively aimed at users with a personal interest in obtaining this information, primarily professionals in the public sector. It is not advertised on any other public pages of the GOV.UK website or social media.

Considering that the ATRS was established with the intention of helping UK public bodies comply with the algorithmic tools disclosure mandate, there is little evidence of who reads the data that the repository collects and for what purposes. However, the way in which the Standard has been designed, how the

website is accessed, and the information for each record is displayed suggests that the type of user the ATRS team had in mind when developing the standard is more likely to be an experienced technology developer or a public sector professional than to a layperson. However, the developers have mentioned that they aimed to provide different levels of depth for various audiences. The target user, in this sense, is multiple:

> We want it to be accessible to the general public, but the more detailed tiers of information within the ATRS are aimed at more sort of specialist audiences, that could be civil society, that could be researchers, journalists, industry, other algorithmic tool owners… (ATRS implementer)

3.7. Contribution to algorithmic transparency

Interviews revealed two main perceived contributions of these repositories by their implementers, collaborators, and users. One contribution is directly linked to algorithmic transparency, while the other relates to a long-term contribution to constructing more ethical algorithmic tools. Two additional contributions that may not have appeared in interviews but are worth considering: 1) Contributing to better use of public resources, as the repository allows governments to know which ADMs they have (instead of duplicating costs by acquiring the same tool multiple times through different public agencies). This also means that public bodies can learn what their peers are doing. Additionally, 2) informing the market of ADM system developers about the needs of government bodies.

The algorithmic transparency promoted by the repositories through the systematic collection and sharing of public sector information is widely seen as a first step toward the long-term goals of the institutions implementing them, such as democratising public information or fostering technological innovation. Academics who have accessed these repositories to search for information state that the available data enables them to see an overview of the systems they are interested in, as opposed to detailed insights about modelling or code.

Implementers and developers have noted that a secondary consequence of the emergence of repositories is the promotion of critical thinking about algorithmic tools. Developers often discover ethical shortcomings in their products when questioned about their purpose and construction, such as the reproduction of biases. In this sense, the implementation of these repositories does not merely contribute to algorithmic transparency but also to a wider ethical use of data technology.

3.7.1. Contributions of Universidad Adolfo Ibáñez's Algoritmos Públicos

The implementers of the Algoritmos Públicos UAI see the repository as only a starting point for algorithmic transparency and government accountability. According to their definitions, gathering and publishing information about algorithmic tools is not enough to evaluate/assess whether or not these algorithms protect civil rights. Therefore, additional efforts are needed to properly guarantee that an algorithmic tool is harmless to civilians. The Algoritmos Públicos team believes that the new Algorithmic Transparency Sheet, part of their Tools for Ethical Algorithms, is a step towards a more thorough evaluation of Chilean institutions' compliance with civil rights:

> Knowing that an algorithm exists, one can, using Law 20.285,[12] request information about a system, and basically, yes, obtain much more information to make a decision. Or one could, through Law 19.88... I don't remember what it's called, basically file a complaint or take action with a public service regarding a specific decision. So it's a starting point, so to speak, for accountability; transparency is ultimately a means for algorithmic accountability, it requires the existence of sufficient information to evaluate the system and determine whether I think it's right or wrong, whether it affects me or not. So, the repository on its own, in its current state, doesn't achieve that, but it helps. (Algoritmos Públicos UAI implementer)

As noted earlier, the repository has proven valuable in UAI's classrooms as a learning tool for better understanding how algorithms are utilised within the Chilean public system. In this regard, it is a solid contribution to education on algorithmic transparency. Similarly, GobLab has highlighted its use within public institutions, offering examples of ADM systems in the public sector and demonstrating how data science can enhance public services. However, interviewees have said the information provided is for specialists, instead of the general public.

3.7.2. Contributions of the European Commission's Public Sector Tech Watch

Notably, the Public Sector Tech Watch team has declared in interviews that algorithmic transparency is not considered one of the goals of the repository since their efforts are aimed at promoting innovation and communication between public sector operators and private technology developers:

> I don't think that there is something there for us, like transparency of the algorithms. It's more like we report the case. It's not that we have access to the source code. It's not that… we ask them to describe if it's compliant, for example, with the AI Act, which is another Act that actually tests the criticality of cases. But yeah, this is not the purpose of the Observatory. The Observatory is more to promote and make these things visible online, but also the mission is working through the AI Office and the AI Act on developing guidelines for the use of AI, first internally and then it will go all externally, when it comes, for example, to the transparency of the algorithms, as you say, but we don't do this under the website that we have. (Tech Watch implementer)

Even if the information has not been designed with the primary intent of being read and used by the general public, our analysis of the repository's walkthrough shows that the information it contains on more than 700 cases across the EU is still a significant contribution to algorithmic transparency, considering the extensive data collected and the complementary qualitative information acquired through in-depth interviews for specific cases of interest to the Public Sector Tech Watch team. This approach provides a broad range and detailed insights into the various cases included in the repository.

3.7.3. Contributions of the UK Government's Algorithmic Transparency Recording Standard

Members of the ATRS team were adamant about the idea that algorithmic transparency is not an end but a means toward more relevant values regarding algorithmic tools employed by governments, such as fairness and accountability. In this regard, the ATRS repository represents only a preliminary step toward making information accessible to public sector operators and civilians. However, it does not ensure that the compiled algorithms are free of risks to the beneficiaries they intend to serve. The main contribution

[12] Law on Transparency of the Civil Service and Access to Information of the State Administration. Available at https://www.bcn.cl/leychile/navegar?idNorma=276363

acknowledged by the team is the availability of information about these algorithmic tools, regardless of whether civilians read it or not:

> Most people probably won't read all these records in great detail. What has kind of been a reflection from that public engagement study we initially did was that members of the public like to know that the information is out there even if they don't want to read it, which is kind of contradictory, but not because they think that just it being out there already provides kind of a level of assurance, it provides them with kind of the confidence that experts or like people who actually know their algorithms, like, you know, academics or other like experts, can look at this information, interpret it for them, and would flag if there's anything that doesn't look right. While as a normal member of the public, you wouldn't necessarily read all this information. So the value in the information isn't necessarily even in everybody reading it, but just it being there for scrutiny purposes rather than for anyone to read. (ATRS implementer)

3.8. Key performance indicators

Overall, the three repositories examined in this study seldom undergo performance evaluations. This is particularly true for the most recent projects, UAI's Algoritmos Públicos and the UK's ATRS, as the teams behind them are currently focused on expanding these initiatives, which have only recently been made available to the public. Operators might have a superficial notion of how the repositories are used and by whom, but this information is not actively collected or analysed. The tools closest to performance indicators are feedback systems, which primarily seek to gather suggestions from public sector workers on incorrect information, potential improvements to the repository categories, or their respective websites.

3.8.1. KPIs for Universidad Adolfo Ibáñez's Algoritmos Públicos

The Algoritmos Públicos team states that the website's visitor numbers and types are checked approximately every six months through Google Analytics. Still, no active efforts have yet been made to analyse the site's user base. Operators have identified that most visits to the site occur after press or media appearances, with most coming from Chile and some from other Latin American countries.

The website provides visitors with a "Comment here" section, which redirects to a Google form for evaluating the user experience in the repository. The form collects information about the user's motivation to visit the repository and whether the feedback pertains to project categories, information fields for each project, the website, or all of these aspects. The form also offers a final box for general recommendations on the repository. In 2022, an analysis was conducted of all feedback received through the form, leading to the following improvements: enabling data downloads from the repository in open data format, enhancing site display on mobile devices (responsiveness), and adding a "unit" field within the institution to include information about the teams that developed the algorithms, as requested. Additionally, the content of algorithm descriptions was standardised that year based on a suggestion from the form. Currently, the suggestion to add a "beneficiaries" field and publish the methodology manual for a more detailed explanation of each field is being implemented.

Figure 7: First screen of Algoritmos Públicos "Comente aquí" (Comment here) form

Source: Algoritmos públicos repository

3.8.2. KPIs for the European Commission's Public Sector Tech Watch

The EU's Public Sector Tech Watch does not currently have an evaluation system for the repository. In contrast to UAI's Algoritmos Públicos, Public Sector Tech Watch does not provide its own form for comments or suggestions regarding the categories or information included. Instead, the repository's website indicates where the visitor might find more information regarding the project and provides an email address for contact: "Please consider that this dataset is currently a work in progress and is set to be regularly enriched with additional use cases. If you notice any mistake or you would like to share some suggestions or feedback. More information here: ..." (European Commission Joint Research Centre, 2023).

3.8.3. KPIs for the UK Government's Algorithmic Transparency Recording Standard

As mentioned above, the ATRS team does not yet have access to user analytics, which might allow interpretation of how the repository's data is used. The ATRS repository's website has one feedback system exclusively that belongs to the entirety of the UK.GOV website. At the bottom of the repository's page, a box asks the user, "Is this page useful?" to which the visitor can reply with "Yes", "No", and "Report a problem with this page" by clicking on the corresponding button.

3.9 Comparison with other algorithmic transparency instruments

As stated above, the state-of-the-art report (GPAI, 2024) proposes a new classification of algorithmic transparency instruments: supply-driven mechanisms and demand-driven mechanisms. Considering their proactive approach to disclosing information, repositories fall under this first category, and as such, might not be the best fit to satisfy transparency demands for specific cases that affect a defined group of people.

Algorithmic transparency instruments to be implemented should be chosen based on criteria such as the impact of the algorithms being disclosed, the institutional resources available, and the target audience.

Public algorithm repositories are a proactive algorithmic transparency instrument with high potential to provide a broad and diverse audience with access to information about algorithms that directly impact people's lives through a single, centralised system. This contrasts with other supply-driven instruments, such as publications of audits or evaluations of ADM systems, which are oriented toward providing in-depth analysis of specific algorithms. This positions repositories of public algorithms as a useful instrument for identifying how, when, and where systems operate, but it rarely allows for their evaluation or impact assessment. The ATRS implemented by the UK Government is an exception, as it provides meta-data on audits and impact assessments conducted regarding the systems it includes, at the expense of having a narrower scope.

When comparing repositories of public algorithms to demand-driven algorithmic transparency instruments, the first and most evident advantage that emerges is that repositories allow for faster and more effective access to information. As mentioned in the state-of-the-art report (GPAI, 2024), requests for information and judge-ordered disclosures have a low success rate, the latter even more so. Repositories can be an appropriate option for accessing information on systems with extensive reach. However, if a detailed and tailored response is required to address specific legal or judicial concerns regarding these systems, repositories are not designed to fulfil that need.

When the context for implementing a repository is properly considered, this instrument can generate social and cultural value: the City of Amsterdam Algorithm Register and the Artificial Intelligence Systems of Helsinki have proven to be valuable resources, enhancing social value by helping manage "increasingly complex urban environments" (Floridi, 2020) and fostering cultural value in the public sector through transparency and openness. Similarly, interviews conducted for this case study reveal that the three repositories in this report generate value in the public sector by promoting transparency and encouraging ethical responsibility among developers and implementers of algorithmic tools. In this way, repositories of public algorithms can consistently promote transparency as a core cultural value in the public sector.

It should be noted, however, that, as standardised instruments that compile large amounts of information, repositories of public algorithms can also posit the danger of reproducing biases that privilege some groups to the detriment of others if governments approach them from a "just do it" perspective (Cath & Jansen, 2021). Moreover, the impact of public registers and repositories is often overrated since they only trace the algorithms employed in the public sector and obfuscate the impact of privately-owned systems that manage and use public information (Cath & Jansen, 2021). Considering these risks, public institutions should thoughtfully consider whether repositories are the best instrument to achieve their algorithmic transparency objectives.

4. Conclusion and recommendations

This report presents a case study analysis of public algorithm repositories in Chile, the European Union, and the United Kingdom, focusing on their effectiveness in promoting algorithmic transparency in the public sector. Through case studies, interviews with key stakeholders, and a review of regulatory frameworks, the report evaluates each repository's strengths, weaknesses, and challenges. Despite sharing the common goal of enhancing transparency, the cases differ in their approaches and regulatory contexts, highlighting the need for flexible, context-specific strategies when designing and managing public algorithm repositories.

To develop a conclusion based on this study's findings, we will address the two overarching questions of the investigation: “How are repositories built and operated?” and “Do these instruments contribute to algorithmic transparency, and if so, what type of value is created and for whom?” In the next section, we summarise the key findings and offer recommendations to highlight the importance of public algorithm repositories in promoting transparency within society.

4.1 Government (state) involvement in the success of repositories

Government or State involvement is crucial for enhancing the creation and implementation of repositories and facilitating a systematic analysis of algorithms used by public services. For instance, the EU’s Public Sector Tech Watch is an institutional effort to promote innovation by enabling information sharing and interaction between private and public organisations. This, in turn, shapes both the design and operation of the repository, focusing on internal information exchange among these entities to the point where implementers seldom regard the repository as motivated by transparency. Hence, the contribution of Public Sector Tech Watch to algorithmic transparency can be considered a valuable by-product - as compared to a primary goal. In contrast, the UK’s ATRS appears to focus on fulfilling the primarily ethical goal of providing proactive algorithmic transparency, directing its efforts toward publishing extensive quality data about algorithmic tools in the territory to guarantee a democratic use of these instruments. In this sense, the algorithmic transparency sought by these repositories differs in their ultimate intentions, which supports the conceptualization of algorithmic transparency as a means to achieve fundamental rights (GPAI, 2024) and not exclusively as an end in itself. On the contrary, initiatives like the Chilean Algoritmos Públicos repository started as a project committed to enhancing the visibility of automated decision-making and support systems in the Chilean public sector, intending to motivate public institutions to leverage data as a means of fostering innovation while also describing how state institutions utilise computational systems and algorithms (GobLab, 2022).

One of this repository's main limitations is its information-gathering process, which relies on public sources. Due to the lack of a national transparency standard regarding the use of these systems, in some cases, it is impossible to obtain all the necessary information.

4.2 Operational responsibilities and data collection approaches in government and university-led repositories

Operational responsibilities vary across the cases depending on their organisational structures and the presence of mandatory data collection processes imposed by governments or regulatory bodies. Repository management teams typically consist of moderately sized groups, ranging from four to six operators. Algorithms in these repositories are subject to regular review and selection of new cases, often through monthly meetings involving the implementing teams alone or in collaboration with external entities such as editorial committees or key stakeholders. Operational responsibilities can vary significantly across different repositories. For example, government-led repositories like Public Sector Tech Watch or the ATRS often require ongoing interaction with other public institutions or stakeholders as their main strategy to collect information. In contrast, university-driven repositories such as UAI's Algoritmos Públicos only use this approach as a supplementary method if the publicly disclosed data or press releases lack sufficient information.

As we observed across the cases, the type of information collected and analysed varies depending on the objectives of each repository. For instance, Algoritmos Públicos aims to support the state's role in providing consistent reports on AMD systems used in various public services across the country. However, this is an independent effort, using its methodologies for data collection to build cases. As a result, its methodology does not rely on state bureaucracy, such as submitting information requests to individual public services. This repository promotes algorithmic transparency by mapping public innovations and uses of algorithms in public services.

4.3 Institutional interactions for data collection and transparency

The primary challenges encountered by repository management teams in two of the three cases (Public Sector Tech Watch and the ATRS) have centred on maintaining a consistent flow of information with the entities responsible for the algorithms they incorporate. Operators of the repositories dedicate substantial time and effort to these information exchanges, time that could otherwise be directed towards identifying new cases and enhancing records to ensure the repositories function more efficiently. This suggests that the demand for transparency regarding specific details of these algorithmic tools remains a relatively new phenomenon within the public sector, and it requires further development before it becomes standard practice. This value emerges primarily as a principle and an obligation within institutional interactions to foster transparency.

4.4 Audience analysis and other KPIs to enhance the impact of repositories

In the three cases analysed, there was no mention of using KPIs to measure the impact of each repository on its audiences. The teams generally struggled to collect and analyse data about the types of users (e.g., journalists, researchers, and the general public) and how these audiences used the repository's data (e.g., for academic research, general information, or to oversee government and public services that use algorithms). If repository teams had the time and resources to analyse this data, they could improve their communication processes, enhancing transparency within public bodies and empowering citizens to oversee the use of algorithms in the public sector.

The examined repositories have demonstrated their effectiveness in providing high-quality data on algorithmic tools, even when the scope and depth of this data vary significantly. As previously noted, users primarily consist of public sector officials and academics who value having a centralised virtual platform to access information about helpful tools employed by public institutions. But what if they focused on bringing this information closer to citizens? In all three cases, there remains a lack of accessible information for the general public on the uses and impacts of AI systems in citizens' everyday interactions with public institutions. A key challenge for each of the initiatives analysed is to effectively communicate their role to the general public in an easily understandable way. Each repository has successfully gathered and developed cases on AI applications within different institutions at this stage. However, these efforts appear to be directed primarily toward close stakeholders without placing the general public at the centre of their communication strategies. The teams managing the repositories, in partnership with their stakeholders, could take additional steps to raise awareness of these algorithmic tools among a wider audience. This could include actively promoting them through institutional social media channels and official websites.

The interviews reference democratising public management, but this appears to be a secondary goal. The three repositories selected in this work are mainly designed for professionals in data science or the public sector. In the UK and EU cases, responsible institutions provide minimal visibility or promotion of the

repositories on their official websites or social media, suggesting a focus on operational functions rather than public engagement. In contrast, Algoritmos Públicos—an independent, university-based repository—offers clear information and facilitates public access, though journalists and academics are still the primary users of the data.

Additionally, while these types of KPIs may help measure outputs—such as audience reach—they may not be sufficient to evaluate all outcomes or the full impact of each repository. For instance, they do not account for internal data collection procedures or the influence on public bodies' organisational culture. Therefore, there is a need for audits and impact assessments that extend beyond audience analysis. For instance, it would be valuable to introduce a KPI that measures the impact of the data provided to citizens, specifically in terms of improvements in public or private services, while prioritising privacy and placing citizens and consumers at the centre of technological design and innovations.

4.5 Financial alternatives

Based on the analysis of the case studies, national and supranational repositories funded by the state or public bodies can focus specifically on achieving their goals and building cases of uses of algorithms in the public sector, even though these teams have some limitations in terms of the number of people who can build and analyse a vast number of cases. On the contrary, repositories whose main funding source comes from other sources (e.g., teaching courses, research projects, consultancy work) are exposed to financial scarcity and uncertainty in achieving their goals. For the Algoritmos Públicos repository, every year, they must find new ways to secure financial resources to continue collecting data with their team to build new cases. Thus, the repository's effectiveness is compromised due to limited resources. A key issue is the shortage of experts involved in its development, along with the inability of those responsible to unify criteria effectively, which impacts the repository's quality. This highlights the need for innovative and collaborative funding mechanisms to ensure repositories remain functional and up to date.

A recommendation to make these efforts more efficient is to provide public resources for building teams capable of collecting data and compiling cases from the vast number of organisations using algorithms as part of their services. If repository teams lack the public resources needed to collect data and track all the public and private institutions incorporating AI technologies, it will be impossible to learn about and understand the impact of these implementations on citizens and consumers.

4.6 Value associated with the repositories

These repositories offer significant functional utility by fostering innovation within the public sector through the exchange of ideas to enhance public administration. Moreover, they create cultural value by generating demand for transparency and accountability from public institutions regarding the impact of their algorithmic tools on citizens. This, in turn, promotes greater awareness of data ethics among algorithmic tool developers, mobilising cultural change within the public administration sector. Additionally, these repositories facilitate research using public data collected by expert groups for the general public. This not only enhances transparency but also helps improve services through evidence professionally analysed by experts.

We identified that these repositories promote "meaningful transparency" as essential tools for communicating accountability to independent stakeholders (Suzor et al., 2019). However, the information collected is not always accessible to the general public, such as ordinary citizens. Instead, the case studies identify journalists and academic researchers as the primary users of this data.

Lastly, economic value creation is also identifiable in this study, as some repositories of public algorithms might work in tandem with private sector companies to promote technological innovation. These interactions simultaneously benefit the private and public sectors, providing valuable assets for profit-oriented organisations while offering new strategies to improve public administration.

4.7 Alignment or divergence with GPAI's data governance principles

The analysis of the case studies in this report highlights significant alignments with the mandates of GPAI's two working groups — Responsible AI and Data Governance. The Working Group on Responsible AI focuses on developing human-centred, fair, and inclusive AI systems that respect human rights and democracy. Its mandate supports the responsible governance of AI following the UN Sustainable Development Goals. The three cases analysed — Algoritmos Públicos, Public Sector Tech Watch, and ATRS — demonstrate how data on the use of algorithms in public services can empower citizens by placing them at the centre of the process. Whether through repository directors, implementers, or users, efforts are consistently made to provide clear and accessible information about algorithms in the public sector.

The Working Group on Data Governance promotes responsible data management for AI, ensuring alignment with human rights, inclusion, and innovation. Its mandate fosters sustainable development by advancing research and projects on ethical data use and governance. The Algoritmos Públicos repository exemplifies these principles, promoting responsible data management by gathering and analysing data on AI and algorithms used in public services and offering improvement recommendations. Similarly, Algoritmos Públicos collaborates with researchers and developers on projects that explore ethical data use in the public sector. Both the Public Sector Tech Watch repository and the UK's ATRS, which are mandated to systematise data and build case studies on AI usage in the public sector, rely on specialised teams of data scientists to analyse data, identify public innovations, and ensure responsible data governance.

References

BritainThinks (2021). *Complete transparency, complete simplicity. How can the public sector be meaningfully transparent about algorithmic decision making?* Available at https://assets.publishing.service.gov.uk/media/60ccae1c8fa8f57cef61fcc7/Complete_transparency_complete_simplicity_-_Accessible.pdf

Cath, C., & Jansen, F. (2021). Dutch Comfort: The limits of AI governance through municipal registers. arXiv preprint arXiv:2109.02944. https://doi.org/10.48550/arXiv.2109.02944

European Commission. (2021). Coordinated plan on artificial intelligence 2021 review. Digital Strategy. Retrieved September 5, 2024 from https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review

European Commission Interoperable Europe. (n. d.). Public Sector Tech Watch. Retrieved September 5, 2024 from https://joinup.ec.europa.eu/collection/public-sector-tech-watch

European Commission Joint Research Centre (JRC) (2023): Public Sector Tech Watch latest dataset of selected cases. European Commission, Joint Research Centre (JRC) [Dataset]. Available at http://data.europa.eu/89h/e8e7bddd-8510-4936-9fa6-7e1b399cbd92

Floridi, L. (2020). Artificial Intelligence as a Public Service: Learning from Amsterdam and Helsinki. Philosophy & Technology, 33, 541-546. https://doi.org/10.1007/s13347-020-00434-3

GobLab Universidad Adolfo Ibáñez. (2022). Repositorio de Algoritmos Públicos de Chile: Primer informe de estado de uso de algoritmos en el sector público. Universidad Adolfo Ibáñez.

GobLab Universidad Adolfo Ibáñez. (2024). “Repositorio de Algoritmos Públicos. Informe Anual 2024”. Universidad Adolfo Ibáñez.

GobLab Universidad Adolfo Ibáñez. (n. d.). *Herramientas Algoritmos Éticos*. https://algoritmospublicos.cl/herramientas

GPAI (2024). “Algorithmic Transparency in the Public Sector: A state-of-the-art report of algorithmic transparency instruments”. Report, May 2024, Global Partnership on Artificial Intelligence.

Gutiérrez, J. D., & Muñoz-Cadena, S. (2023). Building a repository of public algorithms: Case study of the dataset on automated decision-making systems in the Colombian public sector. In L. Beli & W. B. Gaspar (Eds.), *The Quest for AI Sovereignty, Transparency and Accountability*. IGF. Retrieved from: https://repositorio.fgv.br/server/api/core/bitstreams/210b5b35-df1f-4eed-b086-bd4631adf8b9/content

Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New Media & Society*, 20(3), 881-900. https://doi.org/10.1177/1461444816675438

Ragin, C. and Becker, H. (Eds.) (1992) *What Is a Case: Exploring the Foundations of Social Inquiry*. Cambridge University Press.

Riquelme, R. (2023, December 17). Académicos del CIDE crean un registro de algoritmos del sector público de México. *El Economista.* https://www.eleconomista.com.mx/tecnologia/Academicos-del-CIDE-crean-un-registro-de-algoritmos-del-sector-publico-de-Mexico-20231215-0101.html

Stake RE (1995). *The art of case study research*. Sage Publications.

Suzor, N., West, S., Quodling, A., & York, J. (2019). What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. *International Journal Of Communication*, 13, 18. https://ijoc.org/index.php/ijoc/article/view/9736

Tangi, L., Combetto, M. and Martin Bosch, J. (2024). Methodology for the Public Sector Tech Watch Use Case Collection, Publications Office of the European Union, Luxembourg. Retrieved from: doi:10.2760/078522, JRC137409.

UK Government. (n.d. a). Algorithmic transparency recording standard hub. GOV.UK. Retrieved September 5, 2024 from https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub

UK Government. (n.d. b). Algorithmic transparency recording standard: Guidance for public sector bodies. GOV.UK. Retrieved September 5, 2024 from https://www.gov.uk/government/publications/guidance-for-organisations-using-the-algorithmic-transparency-recording-standard/algorithmic-transparency-recording-standard-guidance-for-public-sector-bodies

UK Government. (n.d. c). UK National Data Strategy. GOV.UK. Retrieved September 5, 2024 from: https://www.gov.uk/government/publications/uk-national-data-strategy

Valderrama, M., Hermosilla, M. P. and Garrido, R. (2023). State of the Evidence: Algorithmic Transparency. Open Government Partnership. https://www.opengovpartnership.org/documents/state-of-the-evidence-algorithmic-transparency/

Yin RK (2018). *Case study research and applications: Design and methods* (Sixth edition). SAGE.

Annex A: Summary of conducted interviews

Table 4: Summary of Conducted Interviews

Number	Repository	Type [13]	Organisation	Date
1	Chile (Algoritmos Públicos UAI)	Director	Universidad Adolfo Ibáñez (UAI)	5-7-2024
2	Chile (Algoritmos Públicos UAI)	Stakeholder User Researcher	University	25-7-2024
3	Chile (Algoritmos Públicos UAI)	Stakeholder User Provider	Public entity	25-7-2024
4	EU (Tech Watch)	Director	European Commission	31-7-2024
5	EU (Tech Watch)	Stakeholder User Researcher	University	23-7-2024
6	EU (Tech Watch)	Stakeholder User Provider	European Commission	24-7-2024
7	UK (ATRS)	Director	Department for Science, Innovation and Technology (DSIT)	29-8-2024
8	UK (ATRS)	Team Member	Department for Science, Innovation and Technology (DSIT)	29-8-2024
9	UK (ATRS)	Stakeholder User Provider	Stakeholder	4-5-2024
10	UK (ATRS)	Stakeholder User Provider	Public entity	5-9-2024

[13] The categories used here for the type of interviewees are indicated in section 2. Research design and Methodology.

Annex B: Stage lifecycle and types of information in each case

The following table is an adaptation from Table 8 in the state-of-the-art report (GPAI, 2024), which indicates the type of information disclosed in each of the three repositories studied. An X indicates the presence of a specific type of information in the repositories listed under each column. The table allows for a comparison of the information available across the three cases. The data on which the table has been constructed was consulted on October 31, 2024.

Table 5: Stage Lifecycle and Types of Information in Each Case

Stage of the system's lifecycle	Type of information	Chile (Algoritmos Públicos UAI)	EU (Public Sector Tech Watch latest dataset of selected cases)	UK (Algorithmic Transparency Recording Standard)
I. Planning and design	1. Policy problem to be addressed by the system.			X
	2. System's objectives. [Only includes repositories in which the system's objectives are expressly disclosed].	X		
	3. Metrics of the system's objectives.			
	4. Justification for the adoption of the system, and/or assessment of the pros and cons of the system's deployment and/or justification on why the system is needed.			X
	5. Information about the system's pre-feasibility studies and/or risk and impact assessments.			
	6. Target population or intended beneficiaries.			
	7. Processes for building or acquiring the system.			
	8. Sources that financed the development or acquisition of the system.	X		
	9. Cost of building or acquiring the system.	X		
	10. Unit of the public body that led the building process.			X
	11. Information about the organisation and/or team that built the system.			
II. Data collection and processing	1. Data sources.			X
	2. Types or categories of data used to train, pilot or test, and deploy the system.		X	
	3. Whether personal data was used.	X		X

	4. Justification on how data protection laws were complied with.			
	5. Publication of Dataset Nutrition Labels or analogous means for disclosing basic information about the system.			
	6. Information on data quality (e.g., accuracy, quantity, representativeness, timeliness, and limitations)?			
	7. How often the database is updated and who is responsible for updating it.			
	8. Glossaries with definitions of terms and variables included in the repository.		X	
III. Model building and validation	1. Information on the status of the system and its building process (e.g., conceptual prototype, piloting, production, discontinued, completed)?	X	X	X
	2. Whether diverse techniques and methods were considered to build the system. Justification for the techniques or methods that were prioritised.			X
	3. Information about the source code.			
	4. Architecture of the system and techniques and methods used to build the system.		X	
	5. Information on variables, heuristics, limitations, and assumptions of the model.			
	6. In the case of Machine Learning systems, information on how the system was trained and the data used for training.			
	7. Whether the system was developed in-house or with external developers.			X
	8. Identification of external developers (when applicable).			X
	9. Information on public procurement procedures required for building, acquiring or operating the system (when applicable).			
	10. Hardware required to deploy and operate the system.			
	11. Use of cloud services.			
	12. Who can access and/or use the system.		X	X

	13. Metrics defined to measure the performance of the model.			X
	14. Information on whether the deployment of the systems could entail discrimination or unequal treatment. Mechanisms adopted to prevent discriminating vulnerable populations.			
IV. Deployment and monitoring	1. Name of the system.	X	X	X
	2. System's objectives, tasks, and outputs (Description of what the system does and/or how it does it).	X	X	X
	3. Processes that are supported by the system or decisions that are informed or made with the system.			
	4. Unit in charge of deploying the system (or responsible unit).	X	X	X
	5. Description of human oversight mechanisms and/or at what stage(s) are humans involved in the deployment and use of the system.			
	6. How often the system is used.			
	7. Types of system users.			
	8. Types of beneficiaries or recipients of the system.			
	9. Whether the beneficiaries or recipients of the system are informed about the use of the system and/or whether they receive information explaining how the system influenced the process or decision.			
	10. Precision of system.			
	11. Risk management practices.			X
	12. Information on the procedures for challenging or reviewing decisions taken based on or with the system.			
	13. Contact information of the unit and/or public official in charge of system's deployment.			X
	14. Information on who can access the system's data.			

	15. Does the responsible unit have access to the system's code and/or whether the unit monitors and/or audits the system's performance.			
	16. Whether officials revoke or annul a decision made based or made with the system.			
V. Accountability	1. Whether there is a unit or group accountable for the system.	X		
	2. Whether internal or independent algorithmic audits have been performed and/or what were their findings.			X
	3. Whether impact assessments with a human rights approach have been conducted and/or what were their findings.			
	4. Whether Data Protection Impact Assessments (DPIA) and/or Privacy Impact Assessments (PIA) have been performed and/or what were their findings.		X	X
	5. Whether Ethics Impact Assessments (EIA) have been conducted and/or what were their findings.			
	6. Information about the evaluation of the system's performance.			
	7. Level of achievement of the system's objectives.			
	8. Information about the positive or negative effects generated by the system generated and/or about who has benefitted and/or who has been harmed.			