IN DEPTH ANALYSIS

Requested by the JURI committee

European Parliament

The way forward for better regulation in the EU

Policy Department for Citizens' Rights and Constitutional Affairs
Directorate-General for Internal Policies
PE 736.129 – August 2022

EN

The way forward for better regulation in the EU – better focus, synergies, data and technology

Abstract

This in-depth analysis, commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs at the request of the JURI Committee, looks at the use of data for the purpose of regulatory assessment/evaluation. The author finds that data is needed to support evidence-based regulation, that information technologies, and in particular AI, can enable a more extensive and beneficial use of data, and that the use of data in ex-post evaluations can improve the regulatory process. The in-depth analysis offers policy recommendations.

This document was requested by the European Parliament's Committee on Legal Affairs.

AUTHOR

Giovanni SARTOR, European University Institute of Florence; University of Bologna.

ADMINISTRATOR RESPONSIBLE

Mariusz MACIEJEWSKI

EDITORIAL ASSISTANT

Fabienne VAN DER ELST

LINGUISTIC VERSIONS

Original: EN

ABOUT THE EDITOR

Policy departments provide in-house and external expertise to support EP committees and other parliamentary bodies in shaping legislation and exercising democratic scrutiny over EU internal policies.

To contact the Policy Department or to subscribe for updates, please write to:
Policy Department for Citizens' Rights and Constitutional Affairs
European Parliament
B-1047 Brussels Email: poldep-citizens@europarl.europa.eu

Manuscript completed in July 2022

This document is available on the internet at:
http://www.europarl.europa.eu/supporting-analyses

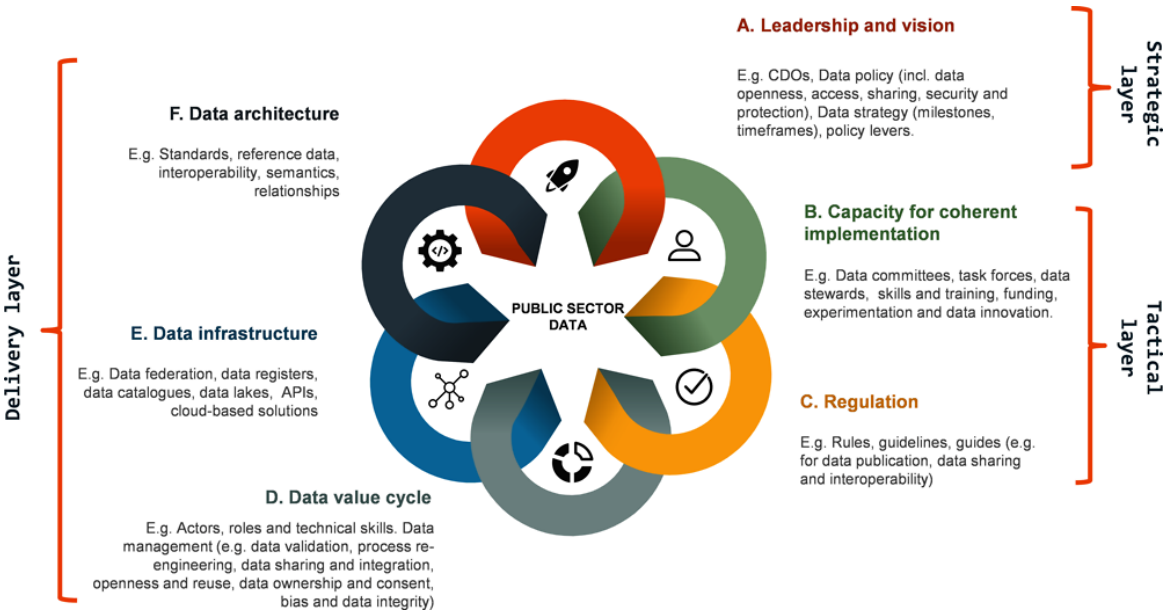
CONTENTS

LIST OF FIGURES

EXECUTIVE SUMMARY

The accelerated growth of data has been counterbalanced by the development of technologies for storing and processing data in such a way as to enable efficient use of vast resources. Moreover, thanks to artificial intelligence and data-analytics such vast datasets can be exploited to extract useful information. Governments, at all levels, must improve their ability to access data that are relevant to defining and implementing policies, and to process such data as needed.

The efficient and effective use of data in the public sector requires governments to engage in data governance, i.e., to design data policies, implement them, adopt appropriate regulations, involve actors with the requisite skills, create a data infrastructure, and define technical architectures. Data are essential to anticipatory governance, as they support forecasting, foresight, policy design and policy evaluation.



The collection/generation, curation and processing of data involves costs, as well as benefits. Costs may be reduced by reusing and repurposing the data. In particular, data collected for the purpose of implementing a policy can be reused for the purpose of evaluating the policy and possibly reforming it.

Computing techniques can be applied to (big) data to do descriptive, predictive and prescriptive analytics. Thanks to AI (machine learning), the very models used for this purpose can be automatically learnt (partially or totally) from vast datasets. Through prediction and simulation, the effects of regulations can be anticipated and assessed.

Technologies to support drafting of legislation can be woven into synergy with technologies to assess/evaluate outcomes. Different approaches may be adopted for regulatory assessment, which make different informational demands. More inclusive multi-criteria approaches for assessing regulatory impacts, such as the UN sustainable development goals, require additional data.

To make data-collection cost effective and sustainable, it is necessary to design data collection and creation policies under which data are automatically produced as a side effect of administrative processes, resulting from the actions by governments and citizens. Private companies have excelled at collecting data in providing services. The same should apply to governments. Data protection and ethical issues pertaining to data collection should be considered from the very start, so that the data processing respects data subjects' right and meet their factual and normative expectations. Risk reduction measures should be adopted, in accordance with the principles of data protection by design and by default (Art. 25 GDPR), with a particular emphasis on the anonymisation or pseudonymisation of data.

The uncertainty of ex-ante assessment, even when carried out by relying on the best methodologies and on adequate datasets, underscores the significance of both interim monitoring and ex-post evaluation. Interim monitoring and ex-post evaluation provide evidence of the outcomes, support democratic dialogue with evidence, contribute to government's accountability toward elected assemblies, such as the European Parliament.

The traditional policy cycle is characterized by evaluations happening at the very end of policymaking. It should now be possible to take advantage of instantaneous or near-instantaneous data processing, so that evaluation results become available the very moment data arrives. Rather than being a neatly defined step at the end of the policy cycle, the evaluation of policies could take place continuously, opening permanent possibilities of reiteration, reassessment, and consideration.

Interim monitoring and ex-post evaluation enable legislators to control the effectiveness and efficacy of the measures proposed, assessed, and implemented by executive branch. Thus, they contribute to ensuring the accountability of towards elective bodies. Parliamentary committees and units should place themselves at the apex of the accountability structure and make efforts to be widely known as the prime location and focus of ex post legislative evaluation, so that information, research and analysis is submitted to them as a matter of routine.

Given the importance of regulatory assessments and evaluations for the EU Parliament, the JURI Committee should consider setting up a permanent Working Group on Better Regulation, to ensure a more active and persistent critical involvement by the Parliament. The Working Group should systematically contribute to identifying shortcomings and proposing improvements, with a special focus on data practices and corresponding technologies, and with the support of academic and research institutions.

Policy recommendations:
1. The public sector should catch up with the private sectors in the capacity to collect and use data.
2. Better and larger datasets should support the policy cycle, for the purpose of forecasting, foresight, policy design, assessment, and evaluation.
3. The cost and benefits of data collection, curation and use should be considered, and addressed by adopting cost-effective solutions, also involving the reuse and repurposing of data. The statistical processing meant to provide aggregate information should be

distinguished from the processing of personal data meant to provide individualised outcomes.

4. The opportunities offered by AI for collecting and interpreting data, extracting aggregate information though analytics, and exploring scenarios and possible developments through simulation, should be exploited.

5. The need to support multicriteria evaluations with appropriate data, according to methods such as the UN sustainable development goals, should be considered. Quantitative metrics should be used whenever possible, consistently with the nature of the data and the goals pursued.

6. Compliance with law and ethics should be ensured, as to uphold the rule of law, respect citizens' rights, and foster trust.

7. Predictions should be matched against reality through monitoring and evaluation. An ex-post evaluation should accompany every ex-ante regulatory assessment, making it possible to adjust policies, in a process that contributes to democratic debate and makes for greater government accountability.

8. To expand and facilitate the role of the European Parliament in participating in the Better Regulation process and in contributing to its improvement and reform, the JURI Committee should consider setting up a Working Group on Better Regulation.

9. A platform involving academic and research institutions could be created, with the task to identify strength and weaknesses of the Better Regulation process, and propose methodologies and technologies to make it more effective.

1. DATA IN THE PUBLIC SECTOR

KEY FINDINGS

The accelerated growth of data has been counterbalanced by the development of technologies for storing and processing data in such a way as to enable an efficient use of vast resources. Moreover, thanks to artificial intelligence and data-analytics such vast datasets can be exploited to extract information. Governments, at all levels, need to improve their ability to access data that are relevant to defining and implementing policies, and to process such data as needed. The combination of AI and Big Data could improve government's performance across different dimensions such as detecting social issues, predicting the effect of policies through analysis and simulation, and supporting the tailored decision of individual cases. Much progress in the efficiency and effectiveness of public action can be achieved through a lawful and ethical use of data in the public sector.

It is well known that in the recent years we have witnessed veritably a "data deluge." The amount of data that is available, has increased at an accelerated pace, as human life has transferred online, and most economic, administrative and social processes are computer mediated. At the same time the cost of data storage and processing has shrunk significantly. More data have been stored in the last few years than in all humankind's pre-computer history.

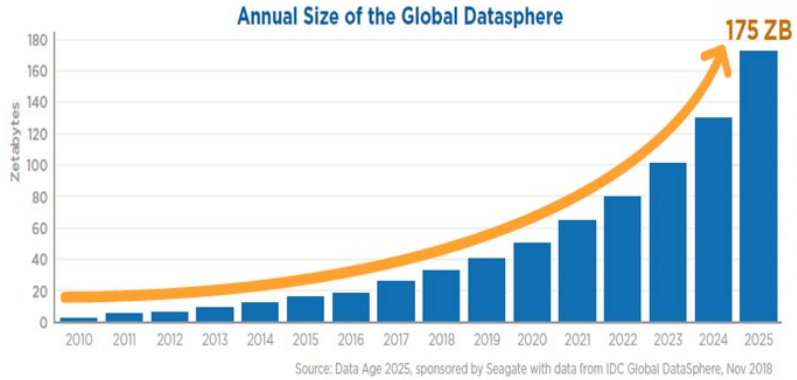


Figure 1. The growth of the datasphere [1]

The data deluge has indeed been countered by the development of technologies for the storage and processing of data that enable the efficient use of such vast resources. Moreover, thanks to artificial intelligence and data-analytics such vast datasets can be exploited to extract useful information. Artificial intelligence can indeed gain aggregate information from vast data sets, even when such data are characterized by the features usually associated with Big Data: high volume, variety in format and content and fast-paced change. Big Data sets can also be used to train AI systems, i.e., to build predictive models, which provide fresh inferences (predictions) when applied to specific cases, both in

[1] Source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018.

the private economy (e.g., for targeted advertising) and in the public sector (e.g., to detect unlawful behaviour).

Arguably, the public sector has lagged behind the private sector in adopting data-intensive technologies.[2] In particular, the private sector has excelled in collecting transaction data, registered when providing services. Online services are indeed characterised by a double flow of information: individuals receive information and services, and, at the same time providers automatically observe, verify, and analyse all transactions, using every character typed or link clicked.[3] The vast amounts of data collected by the private sector raises a number of concerns which cannot be considered here, in particular with regard to the violations of privacy and data protection, the manipulation of individual users/consumers, the impacts on the public sphere (e.g., the spread of fake news and extreme opinions), and the competitive advantages of controllers of Big Data resource. Together with these worries we need to consider the advantages that data resources offer to private companies, with regard to both the aggregate information they can extract from such data (e.g., as when looking to anticipate future demand and other economic trends) and the opportunity to personalise performance toward individual users/consumers. Public sector bodies are less ready not only to collect transactional data, but also to use the data they have for analytical and predictive purposes. It has been indeed observed that

> *In the classic Weberian model of bureaucracy, data are compressed within files, available for checking individual pieces of information, but generating no usable data for analytics. This characteristic of governments' information architecture persisted into the era of computerization, with a lack of usable data remaining a feature of the "legacy systems" of many governments.*[4]

In the era of Big Data and AI, governments, at all levels, must improve their ability to access data that are relevant to defining and implementing policies, and to process such data as needed. The combination of AI and Big Data should improve governments' performance across different dimensions such as detecting social issues, anticipating the effect of possible policies through analysis and simulation, and supporting decision-making in individual cases. [5].

The collection of data by the public sector raises important concerns: privacy, data protection and freedoms are at risk when data about individuals and groups are used for purposes of surveillance, control and manipulation Even when data are collected for valuable purposes, there is always the possibility that they are reused in ways inconsistent with a democratic society (so-called function creep). However, even though the prospect of a surveillance state raises well-justified worries, within a democratic society vast progress can be achieved through a lawful and ethical use of data in the public sector, by ensuring not only data protection, but also transparency and trust.

---

2 Van Oojen at al (2019).

3 Varian (2020).

4 Margetts (2022).

5 Allard et al (2018).

2. DATA IN THE POLICY CYCLE

KEY FINDINGS

An efficient and effective use of data in the public sector requires that governments design and implement data policies, adopt appropriate regulations, involve actors with the requisite skills, create a data infrastructure, and define technical architectures for it. There is a synergy to be had between the data produced when planning, delivering, and evaluating a policy, since such data can be reused at subsequent stages of the policy cycle. Data are essential to anticipatory governance, as they support forecasting and foresight as well as policy design and evaluation.

Data should be recognised as a key strategic asset for the public sector, with its value defined and its impact measured.[6] Consequently, active efforts are needed to remove barriers to managing, sharing, and re-using data, and data should be used to transform the design, delivery and monitoring of public policies and services. A data-driven public sector —in the sense that it relies heavily on data in predicting forecasting needs, shaping the delivery of services, and understanding and responding to change— is indeed considered a key dimension of the digital government strategy, as shown in Figure 2.

Figure 2. Data in digital government[7]

The efficient and effective use of data in the public sector should not be taken for granted. It requires, as shown in Figure 3, a governance structure that includes the ability to design a data policy, implement

[6] OECD (2019d, 8).
[7] Source: OECD (2019d, 13)

it, adopt appropriate regulations, involve actors with the requisite skills, create a data infrastructure, and define technical architectures for it.

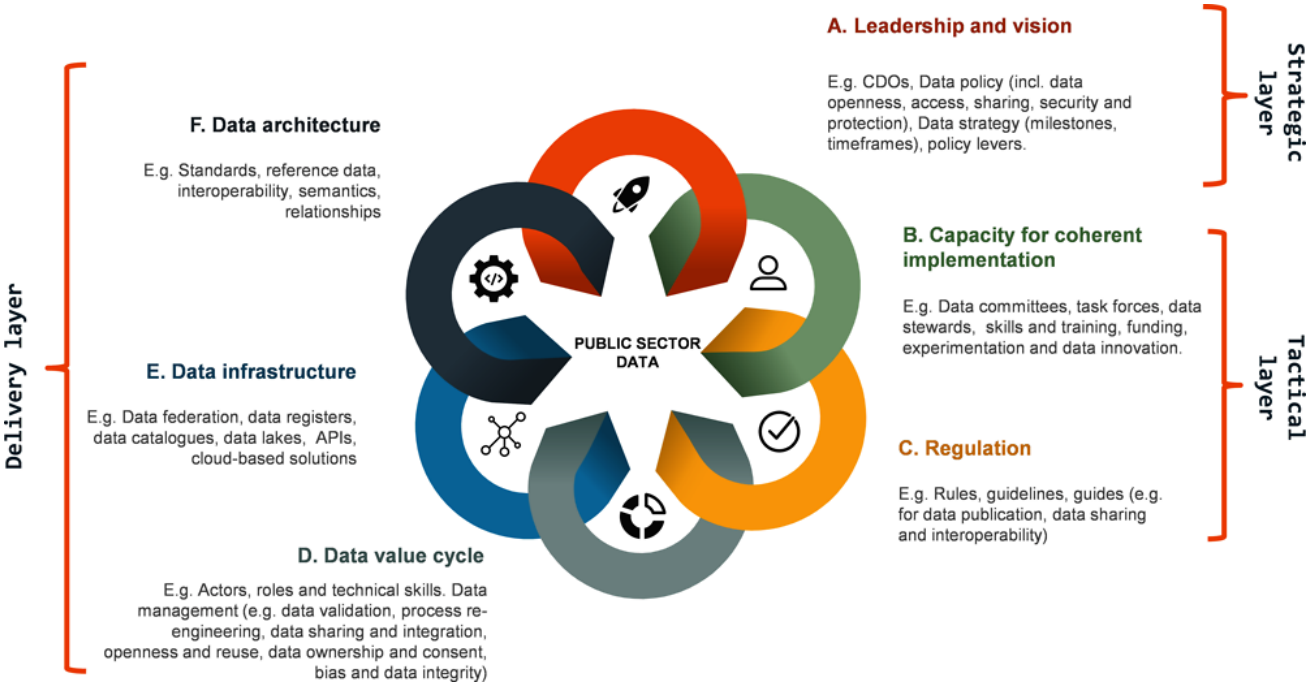


Figure 3. Data governance in the public sector[8]

In 2017 the Economist ran an article describing information as the new oil,[9] the idea being that data is a highly valuable asset which fundamentally powers the economy; this idea has since become a trope. However, information is relevantly different from oil (and from other natural resources): it is non-rival, and indeed its value tends to grow with use (by different users) and with increased links (with other items of information), while it tends to decrease over time (hence, the importance of acquiring new information and verifying and updating old one). Moreover, it is not depletable; on the contrary new valuable information can be generated by processing existing information. The features of information as an asset also tend to change as technologies evolve. It has been argued that additional data might have negative utility, because of information overload.[10] However, in the context of Big Data and AI, while it can still be argued that information often has a decreasing marginal utility (as the size of a data set increases, additional items contribute less to higher performance), technological solutions may enable the scalability of data resources, so that every new data item can profitably be integrated with existing ones, and contribute to better performance.

Figure 4 shows the process for collecting and using data in such a way as to maximise their value for the public sector. This process starts with collecting and generating data and proceeds by storing and securing such data, curating, and distributing them, and finally using them to extract information and define and implement policies, which may require further data to be collected or generated, so that the cycle may start over.

8 Source: OECD (2019b, 164).
9 Economist (2017).
10 Moody and Walsh (1999).

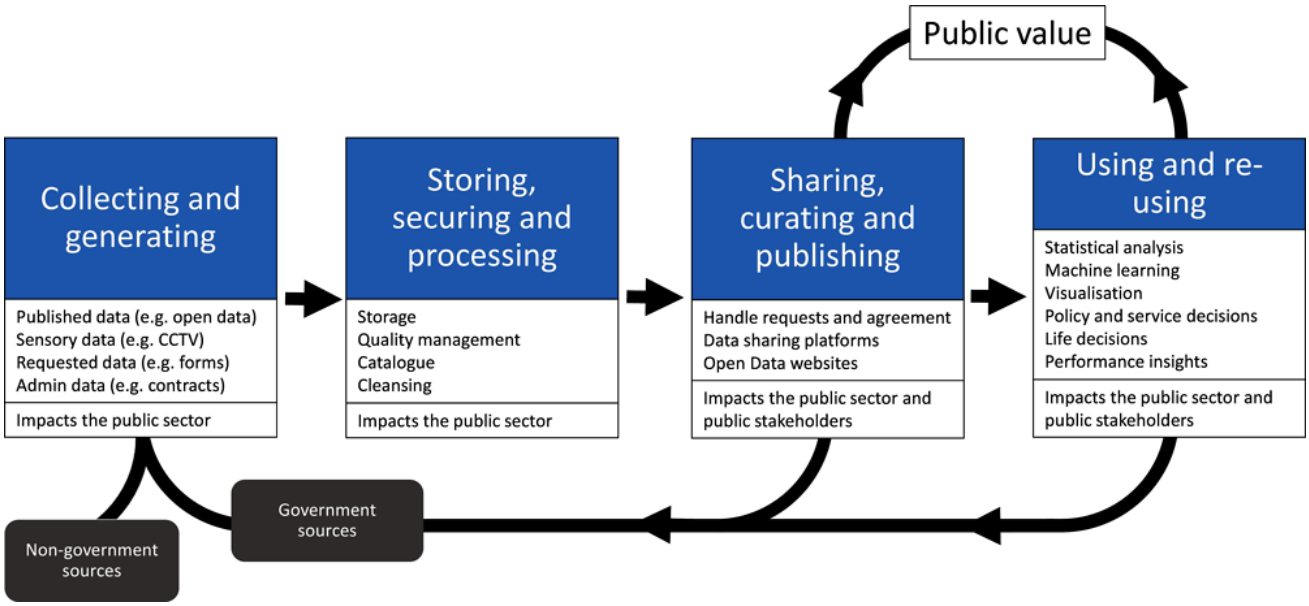


Figure 4. The government data cycle [11]

A broad picture of the use of data in the public sector is presented in Figure 5, which shows how data should be used for connected and mutually reinforcing purposes: anticipating changes, forecasting needs, and designing policies. Matching needs (anticipating and planning) requires adequate information, implementing policies (delivery), requires further data but also provides opportunities for data collection, assessing the policy on the basis of its implementation (evaluation and monitoring) provides information that can be used to improve both the implementation of the policy and the design of new or revised policies.

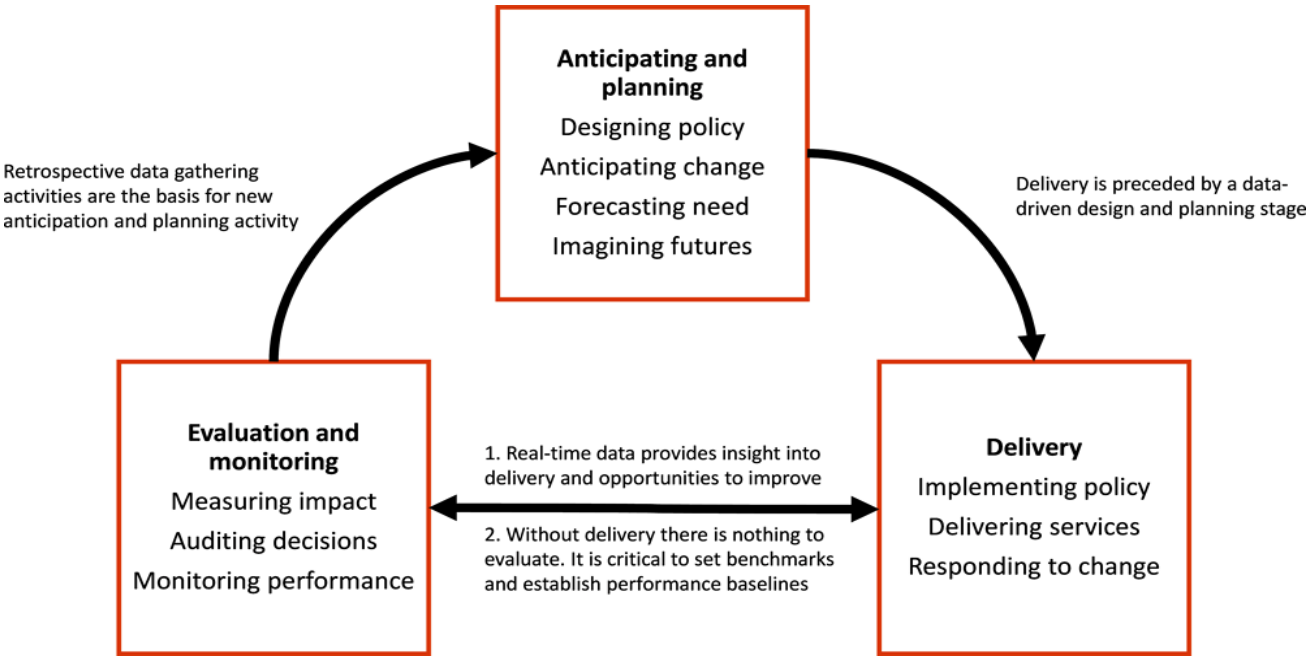


Figure 5. The use of data in the public section [12]

[11] van Ooijen et al. (2019, 11).
[12] OECD (2019d, 88).

The top box in Figure 5 is concerned with what may be called "anticipatory governance", in a broad sense, namely the

> *systematic efforts to consider the future in order to inform policy decisions today. In this context, governments respond proactively rather than reactively, based on knowledge and evidence rather than experience and protocol".*[13]

It is important to distinguish the different data driven activities listed in Figure 5:

- Forecasting (which includes anticipating change) is meant to predict the future, namely, to use existing data and trends to try to predict the most likely developments and outcomes, so as to anticipate "societal, economic or natural developments that are likely to occur in the future."[14]
- Imagining futures, also referred to as foresight is instead meant to identify scenarios, i.e., it "systematically explores multiple plausible versions of how the future could be different from expected, and then uses them to make policies more prepared and agile today," [15]
- Designing policies, rather than forecasting the independent evolution of society, requires. predicting what differences a policy would make, by identifying the expected causal effects of a regulatory intervention. It also requires comparing these effects with the effects of alternative options This aspect is the focus of the ex-ante assessment of regulatory impacts.
- Evaluating policies, requires engaging with actual causation,[16] i.e., determining the causes why certain state of affairs took place or failed to take place, and in particular assessing the extent to which a policy produced or failed to produce its intended outcomes or produced some side effects (perhaps unwanted).

All the activities just mentioned require appropriate data, and on this basis, they produce new aggregate information for policy makers, which information can be used for further activities.[17] In particular the ex-post evaluation is a key aspect of anticipatory governance, since its outputs can direct forecasting, foresight and policy design. In particular, with regard to legislation, an ex-post evaluation of the real impacts of legislative acts is strongly needed for the purpose of improving anticipatory governance, and specifically, ex ante assessments.

[13] OECD (2019d, 90).
[14] OECD (2019d, 90).
[15] Ubaldi et al (2019, 18).
[16] Halpern (2016).
[17] Hochtl et al (2016).

3. THE COSTS AND BENEFITS OF DATA

KEY FINDINGS

The collection/generation, curation and processing of data involves costs, as well as benefits. The costs can be avoided or reduced by reusing and repurposing the data. In particular, data collected for the purpose of implementing a policy can be reused (at no collection cost) for the purpose of evaluating the policy and possibly reforming it. The use of data must comply with law and ethics, in such a way as to respect the rule of law, align with the collective and individual good, and contribute to fostering citizens' trust, which in turn facilitates the collection and reuse of data.

When considering the use of data by government it is important to stress that the collection/generation, curation, and processing of data involves costs, along with benefits (see Figure 6). A cost-benefit analysis also applies to the data, since the cost of data-related activities is an aspect of the larger cost of designing and implementing a policy. However, in considering the cost of collecting data for the purpose of the assessment of a legislative policy, it is important to stress the possibility of reusing and repurposing the data. In particular, data collected for the purpose of implementing a policy can be reused (at no collection cost) for the purpose of evaluating the policy and possibly reforming it. In particular, data collected while implementing a regulation are a key asset in evaluating that regulation.

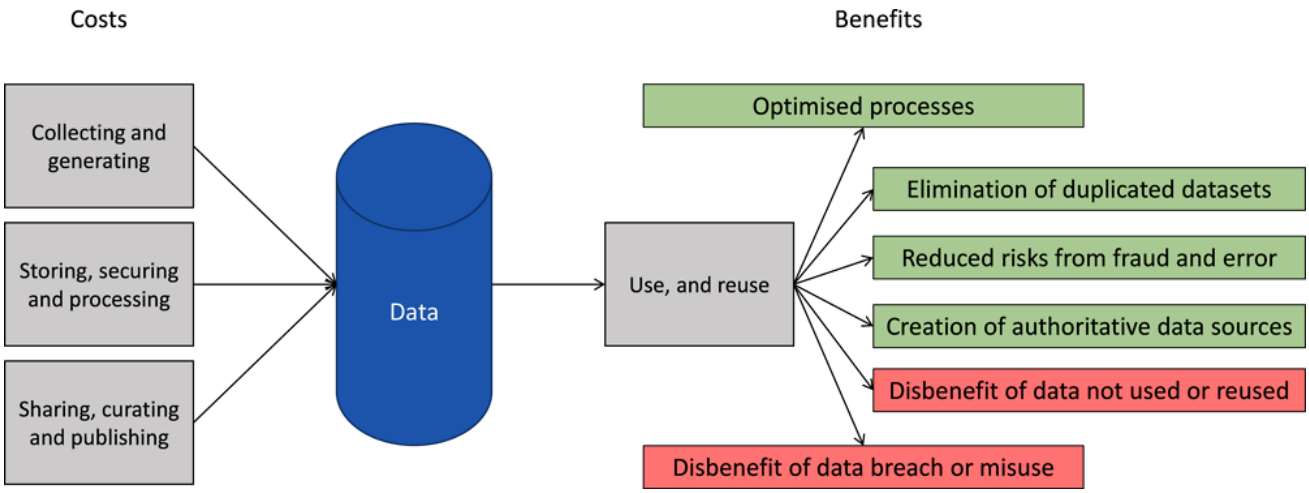


Figure 6. The data-value cycle [18]

As a final, but most important, consideration, it must be stressed that the use of data must comply with law and ethics. Among the legal requirements, a key role is played by data protection law (including the GDPR, the ePrivacy Regulation, the Directive on data protection and the law enforcement), as well as the new emerging data laws (including by the proposed Data governance act and Data act, which

[18] OECD (2019d, 77).

also address the reuse of publicly collected data by private companies, and access to privately collected data by public administrations).[19]

Going beyond legal requirements, we need to consider that data ethics, which indicates moral values to be respected and pursued when engaging with data and corresponding morally adequate conducts, being characterised as the branch of ethics that:

> *studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g., right conducts or right values)".*[20]

Lawfulness and morality in the management of public data are valuable in upholding the rule of law and achieving a fit between governmental action and social good. Moreover, they help to engender citizens' trust and their support of policies, which in turn facilitate the collection and reuse of data.[21]

[19] On access to data, see recently Mayer-Schoenberger (2022).

[20] Floridi and Taddeo (2016).

[21] OCDE (2019d, Ch. 4).

4. NEW PROSPECTS FOR THE USES OF DATA THROUGH BIG DATA AND AI

KEY FINDINGS

Computing techniques can be applied to (big) data for the purpose of descriptive, predictive and prescriptive analytics. Thanks to AI (machine learning), the models used for analytics can be automatically learnt from vast datasets. AI is already helping government detect issues, predict phenomena, and simulate social dynamics. Consequently, it can contribute to improving policy-making and service delivery. AI-based prediction and simulation can increase our ability to anticipate the effects of new regulations.

Today we have a range of computing techniques that can be applied to (big) data for analytics, i.e., for extracting insights. Indeed, analytics has been defined as "the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions", and the following aspects of it are distinguished:

- Descriptive analytics (also called business intelligence or performance reporting) provides access to historical and current data. It delivers the ability to alert, explore, and report using both internal and external data from a variety of sources.
- Predictive analytics uses quantitative techniques (e.g., propensity, segmentation, network analysis and econometric forecasting) and technologies (such as models and rule-based systems) to predict the future on the basis of past data.
- Prescriptive analytics uses a variety of quantitative techniques (such as optimization) and technologies (e.g., models, machine learning and recommendation engines) to specify optimal behaviours and actions.[22]

In recent times AI has indeed become a key instrument for analytics. Through machine learning descriptive, predictive, or prescriptive models can be (partially or totally) learnt automatically from vast datasets. It has been claimed that leading organisation "are rapidly making a strategic shift toward cognitive technologies in general, and machine learning in particular", this being the only feasible option if they are "to handle the amount of data they have at their disposal and to create the personalized, rapidly-adapting models they need"[23] It is worth pointing out that the terminology pertaining to analytics and prediction is not used consistently. In some cases, all activities engaged in inferring information from data, are called predictions.[24] Here the term "prediction" is used to refer only to inferences made about the future.

Analytics, and in particular AI-based analytics is mostly deployed by companies in the private sector, and in particular by tech companies, but analytics can also be used by the public sector, for valuable

[22] Davenport and Harris (2017, 30).
[23] Davenport and Harris (2017, 18).
[24] Agrawal et al (2018).

purposes. [25]I In recent years there has indeed been much interest in government's use of data science and AI.[26] A 2020 study on AI in the US federal government found that nearly half of the agencies studied had experimented with AI and related machine learning tools.[27] Among the AI application listed, are the following: the extraction of information from textual reports on adverse drugs events; the analysis of consumer complaints, the processing of worker injury narratives.

It has been claimed that, more generally, AI can contribute to improve policymaking and service delivery, by helping government to perform three key tasks, namely, detection, prediction, and simulation.[28]

The detection task is concerned with understanding societal and economic behaviors, trends, and patterns as a precondition for gauging public policy accordingly. An important concern is with detecting unlawful or anyway unwanted behaviour, where AI is already supporting the task of identifying tax frauds, money laundering, and terrorism threats. The use of Big Data and AI for detection purposes can play a useful role, but it should be applied fairly across the population, with appropriate caution and careful human assessment of potentially unlawful cases identified by the machine. A most significant instance of abusive deployment of AI has recently emerged in the Netherlands, where a large-scale project aimed at automatically detecting welfare frauds has led to a vast number of people being wrongfully deprived of benefits and subject to fines, and certain groups being subjected to unfair treatment.[29]

Prediction is concerned with anticipating individual events or aggregate phenomena. Examples exist of government using AI to predict aggregate demands, in settings such as schools, prisons, or childcare facilities, or to anticipate the spread of disease, or again to categorise and aggregate criminal complaints. Worries have been raised about applying prediction to individuals, a practice which is widespread in the private sector —e.g., in health care, the insurance industry, credit scoring, and job recruitment— but which has some controversial applications also in the public sector, e.g., in assessing recidivism, in predictive policing, or in identifying families at risk of violence or neglect of children. Less problematic is the use of prediction to anticipate aggregate phenomena. Such aggregate predictions may concern future conditions (e.g., economic, and social trends), or the outcomes of policies. In the latter case, the expected outcomes of alternative interventions have to be determined, which calls for a causal analysis. Predictions of both social conditions and outcomes of policies may be needed for the ex-ante impact assessment of regulations, the first ones providing the context for the second.

Automated prediction can rely on different models, e.g., on econometrics or statistics, possibly supplemented by machine learning. The current excitement about AI should not make us forget that all models are fallible, and that more traditional statistical-optimisation models can often outperform AI approaches.

A distinct development —made possible by the availability of vast computer resources and appropriate computational techniques— is computer simulation.

> *Governments need ways of testing out interventions before they are implemented to understand their likely effects, especially those of costly new initiatives, major shifts in resource allocation, or cost-cutting regimes aimed at saving public resources. In the past, the only option for trying out initiatives was by running field experiments: randomized trials in which the intervention is applied*

[25] An application to corruption and fraud risk assessments is presented in OCDE (2019a).

[26] Margetts and Dorobantu (2019)

[27] Engstrom et al (2020).

[28] Margetts (2022).

[29] Heikkila (2022).

> *to a "treatment group" and the results are compared with a "control group." But such trials are expensive and take a long time, challenge notions of public equity, and sometimes are just not possible due to attrition or ethical constraints. In contrast, the availability of large-scale transactional data, and innovative combinations of agent computing and machine learning, allow the simulation of interventions so unintended consequences can be explored without causing harm.*[30]

Social simulation relies an agent-based modelling, rather than on analytical formalisation through mathematical equations.[31] Agent-based models are in principle experimental: the model specifies the features and behaviour of individual agents, so that the aggregate dynamic of the system emerges through the interactions of such agents. For instance, to anticipate patterns of road traffic, the behaviour of each vehicle (given its kind, ownership, etc.) may be algorithmically specified, along with the features of the environment (places, roads, etc.) in which the vehicles are operating. The traffic (with congestions, accidents, etc.) will result by the interactions of the individual vehicles. Similarly, by creating digital agents that will behave in certain ways when trading with one another, it is possible to investigate the dynamics of markets.

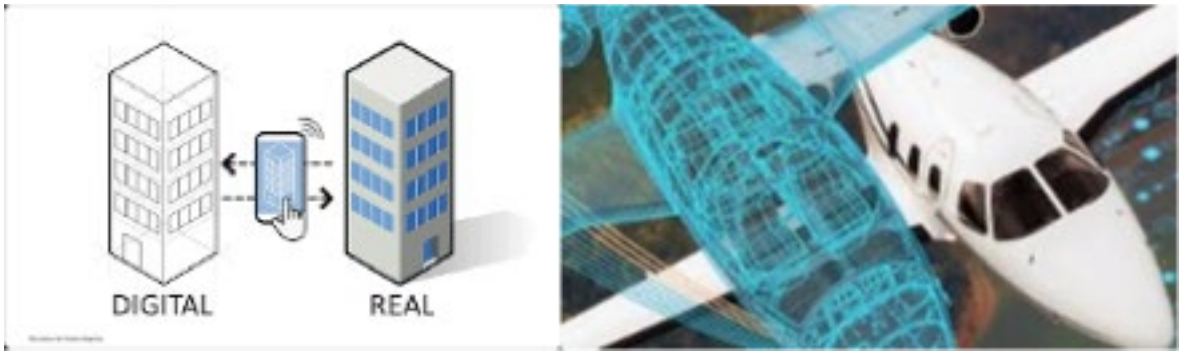


Figure 7. Digital twins

An interesting idea, related to the concept of simulation, is that of "digital twin", i.e., a virtual representation of a physical or social entity, which is constantly linked to that entity through dataflows. [32] As defined by IBM:

> *[a] digital twin is a virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning and reasoning to help decision-making*[33]

The idea of a digital twin originates in engineering (initially in space-engineering) but has since been expanded to cover not only technological devices, but also buildings, factories, cities, and other systems. The digital twin is used to anticipate problems of the corresponding real system, and to test potential solutions before implementing them in the real world. It has been argued that AI and Big Data

[30] Margetts (2022, 364).
[31] Hamill and Gilbert (2016).
[32] Fuller et al (2020).
[33] https://www.ibm.com/topics/what-is-a-digital-twin.

technology can enable us to create digital twins not only of physical systems, but also of social and socio-technical systems, i.e., societal twins.[34] Such models could be used to proactively determine how social systems may respond to future contingencies, identify future issues, and evaluate possible interventions, such as the enactment of new regulations.

[34] Birks et al (2020).

5. DATA AND ASSESSMENT / EVALUATION METHODOLOGIES

KEY FINDINGS

Technologies to support drafting of legislation can be woven into synergy with technologies to assess/evaluate outcomes. Different approaches may be adopted for regulatory assessment which make different informational demands. More inclusive multi-criteria approaches for assessing regulatory impacts, such as the UN sustainable development goals, may require additional data.

In this section and in the next one, we will focus on the deployment of data in ex-ante assessments and ex-post evaluations of regulatory initiatives.

As a preliminary observation, we need to consider that information technologies can be used in law-making to manage legal sources, supporting their drafting and retrieval.[35] This includes:

- the digital management of legal sources to support drafting, amendments, consolidation, referencing, the annotation of documents with metadata, and the production of electronic versions;
- the insertion of semantic information within legal sources, to support retrieval and support consistency in drafting;
- the computable modelling of legal documents to test for logical consistency and completeness, as well as to test their application in real/hypothetical cases, check alignment between different texts (e.g., EU and national laws), and detect transposition issues;
- the use of machine learning to analyse regulatory documents, assess their terminological consistence, evaluate their language, and identify related documents, including administrative and judicial decisions.

Here we will not consider these application domains, but rather only focus on assessment of policies. However, synergies between drafting support and impact assessment/evaluation should be developed in the future and exploited to the benefit of both.

It must be considered that regulatory assessment may rely on different approaches, which make different informational demands.

A popular model is *cost-benefit analysis*, which measures the potential benefits (advantages) of a measure under consideration against its potential losses (disadvantages). Benefits and losses are quantified by numbers expresses in the same unit, usually corresponding to a common currency (e.g., euros or dollars). These quantities—the measure of the overall benefit or loss at stake—are often determined by referring to the individuals' willingness to pay in order to gain a benefit or their willingness to accept payment in exchange for a loss, while also taking into account the probability that such a benefit or loss should take place. By summing up all expected benefits and subtracting all expected losses, for all individuals, we come up with a single number that indicates the overall merit of

[35] On legal analytics, see Ashley (2017). On computer support to legislative drafting, see Palmirani et al (2022).

the measure being considered, which can be compared with a corresponding number assigned to each of the alternatives under consideration.[36]

Cost-benefit analysis is a powerful approach, which has the advantage of providing comparable numerical outcomes for all options under consideration. However, in many domains it encounters difficulties, pertaining to the possibility of operationalizing it, as well as to some normative issues. In particular, cost-benefit analysis makes informational demands —collecting data on willingness to pay or accept, for all the individuals concerned— that cannot be directly met, making it necessary to rely on proxies. Further issues pertain to the fact that cost-benefit analysis tends to disregard the distribution of benefits and costs between different individuals and groups (it just consider their sum total), and tends to under- or mis-appreciate collective goods as well as the goods that cannot be assigned a meaningful monetary value (such as human life, or respect of fundamental rights).[37] Thus, it is generally understood that cost-benefit analysis can contribute to the appreciation of certain policies, with regard to some of their impacts, but in many cases cannot provide an overall evaluation.

Alternative/complementary, less demanding criteria for assessing legislative measures consist in least cost analysis and cost effectiveness analysis. *Least cost analysis* looks only at costs, in order to select the alternative option that entails the lowest cost; thus, it does not adequately address those cases in which a measure having higher costs would deliver greater benefit (the greater benefit fully justifying the higher costs). *Cost-effectiveness analysis* consider the relation between units of benefit and units of expense (e.g., dividing the number of lives saved by the euro amount needed for healthcare measures); thus it does not adequately address cases in which a measure having a broader scope would provide a greater overall net benefit (benefit-cost) , even though a more restricted measure might have greater cost effectiveness (compare vaccinating all individuals, rather than only those at greater risk, during a pandemic).[38] I

In multicriteria decision making the achievement of different objectives is separately considered, such as the contribution a policy can make to GDP, to the environment, or to individual rights, possibly using different methods and scales to measure such impacts. In such cases, in order to determine what measure is preferrable, all things considered, a further evaluation is required, whenever, as it often happens, one measure is more favourable under some criteria and less favourable under some others in comparison with other measures (e.g., a regulatory option is more protective of certain individual rights, such as privacy, and more costly for companies). Thus, in such cases it needs to be determined whether the advantages of one measure in certain regards (e.g., privacy rights) are more or less important than the advantages of the alternative measure in other regards (e.g., cost reduction for companies). Thus, a human assessment may be needed to establish what measure (what package of benefits and losses) is preferable overall, or a calculation has to be defined that transforms each benefit or loss into a common currency (units of utility, or "utils", dollars, euros, etc.), or that otherwise makes it possible to compare alternative measures.[39]

Multicriteria decision making appears to have been adopted in EU regulatory assessments, at least for the most significant regulations. For instance, the impact assessment for the Digital Services Act[40] separately considers economic impacts, social impacts, impacts on fundamental rights, and

[36] Zerbe (2006), Boardman et al (2018)

[37] Hansson (2010)

[38] For a discussion of different approaches to regulatory assessment, see Renda (2015). For some critical considerations, see Micklitz (2022).

[39] A vast set of approaches to multicriteria decision-making exist. For a seminal contribution see Keeney and Raiffa (1993); for a recent account Ishizaka and Nemery (2013).

[40] Brussels, 15.12.2020 SWD (2020) 348 final.

environmental impacts, further detailing such impacts, attempting to quantify some of them (in particular, economic and environmental impacts).

In Figure 8, a summary assessment of three options is presented: (1) baseline, limited measures against illegal activities, (2) full harmonization, and (3) asymmetric measures and EU governance. The chosen one is the third option.

Clearly, making a comprehensive assessment of a policy having such a vast set of different potential impacts by monetising all inputs on individuals according to a cost-benefit analysis seems unfeasible, aside from raising the previously mentioned normative issues.

Recently an increasing interest has been taken in the UN's Sustainable Development Goals, which consist of seventeen goals—(1) no poverty; (2) zero hunger; (3) good health and well-being; (4) quality education; (5) gender equality; (6) clean water and sanitation; (7) affordable and clean energy; (8) decent work and economic growth; (9) industry, innovation, and infrastructure; (10) reduced inequalities; (11) sustainable cities and communities; (12) responsible consumption and production; (13) climate action; (14) life below water; (15) life on land; (16) peace, justice, and strong institutions; and (17) partnerships for these goals. The Sustainable Development Goals framework also provides targets for each goal and indicators for the achievement of the goals. While methods have been proposed for computing the merit of policies relative to this framework, this is a challenging task, given the diversity of the goals and the multiple indicators on which basis they are quantified.[41]

In a recent Report on Better Regulation,[42] the European Parliament supports a broad and inclusive approach to impact assessments and evaluations:[43]

> *[The Parliament] welcomes the Commission's intention to improve the analysis and reporting of proposals' impacts, for example on competitiveness and SMEs, territoriality, sustainability, equality, subsidiarity and proportionality, which could also help identify gaps, needs and opportunities, as well as help discover existing risks and trends, and therefore contribute to defining policy priorities and devising strategic planning with a long-term perspective, especially in the least developed countries and with regard to achieving the sustainable development goals (SDGs);*

[41] For a proposal, see Guerrero and Castaneda (2020).

[42] European Parliament (2022).

[43] For a critical discussion on methods for regulatory assessment in the EU, see Renda (2022), for an analysis of some shortcomings and delays, see Sion et al (2022).

	Effectiveness	Efficiency		Coherence		
		Costs	*Benefits*	*a*	*b*	*c*
Baseline	~	~	~	~	~	~
Option 1	+	>	+	+	+	+
Option 2	++	>>	++	++	+	+
Option 3: Sub-option 3.A	+++	>>>	+++	+++	+	+
Option 3: Sub-option 3.B	+++	>>>>	++++	+++	+	+

Figure 8. The top-level outcome of the ex-ante assessment of the Digital Services Act.[44]

Impacts assessed	***Baseline***	***Option 1***	***Option 2***	***Option 3***
Economic impacts				
Functioning of the Internal Market and competition	~	+	++	+++
Costs and administrative burdens on digital services	~	>	>>	>>[137] / >>>[138]
Competitiveness, innovation, and investment	~	+	++	+++
Costs for public authorities	~	>	>>	>>>
Trade, third countries and international relations	~	+	+	+
Social impacts				
Online safety	~	+	++	+++
Enforcement and supervision by authorities	~	+	++	+++
Fundamental and rights (as laid down in the EU Charter)				
Freedom of expression (Art 11)	~	+	++	+++
Non-discrimination, equality, dignity (Art 21, 23,1)	~	+	++	+++
Private life and privacy of communications (Art 7)	~	+	+	++
Personal data protection (Article 8)	~	~	~	~
Rights of the child (Art 24)	~	+	++	+++
Right to property (Art 17)	~	+	+	+
Freedom to conduct a business (Art 16)	~	+	+	+
User redress	~	+	++	++
Overall	~	+	++	+++

Figure 9. The ex-ante assessment of the impacts of the Digital Services Act[45]

[44] From European Commission (2020, 67).

[45] From European Commission (2020, 67-8).

6. THE COLLECTION AND REUSE OF DATA

KEY FINDINGS

To make data-collection cost-effective and sustainable, data collection and creation policies need to be designed in such a way that new data are automatically produced as side effect of administrative processes. The use of data in regulatory assessment and evaluation is in principle concerned with aggregate data, so that input data may in most cases be anonymous or at least pseudonymous, and the output data consists in aggregate, non-personal information. Thus, it seems that this processing belongs with the concept of statistical processing under the GDPR.

As noted, in the context of AI and Big Data new opportunities are available for ex-ante assessments and ex post evaluations of the merits of regulations. To this end, however, large datasets have to be made available, which requires expanding the data collection process. AI itself can contribute to this process, for instance by analysing and interpreting data inputs, e.g., by extracting data out of natural language reports (e.g., on accidents, or complaints). AI can then be used to mine for information, build models, test hypotheses, and develop what-if analyses and simulations.

To make data-collection cost effective and sustainable it is necessary to design data collection and creation policies, under which data are automatically produced as a side effect of administrative processes, resulting from the actions by public administrations and citizens. As noted, private companies have excelled at collecting data in providing services. The same should apply to governments. Data protection and ethical issues pertaining to data collection should be considered from the very start, so that the data processing respects data subjects' right and meets their factual and normative expectations. Risk reduction measures should be adopted, in accordance with the principles of data protection by design and by default (Art. 25 GDPR), with particular emphasis on the anonymisation or pseudonymisation of data.

The use of data for the purpose of regulatory assessment and evaluation is in principle concerned with aggregate data, so that the input data may in most cases be anonymous or at least pseudonymous, and the output should in any event consists in aggregate, non-personal information. Thus, it seems this processing should fall in the concept of statistical processing which, according to Recital 162 of GDPR, requires that the result of the processing "is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person". Note that personal data collected for other purposes can, in principle, be reused for statistical processing (Art. 5 (1) (b) GDPR).

The reuse of data for regulatory assessment/evaluation should indeed be encouraged, within the framework provided by the Data Governance Act and the Data Act. Relevant data can also be obtained through tools originally designed for use by individuals.

An example would be a diverted use of Claudette, an AI tool meant to enable individuals and consumer associations to assess the legality and fairness of online terms of service and data protection policies.[46] The systems has been applied to a large set of data protection policies collected by crawling websites

[46] Lippi et al (2019).

before and after the enactment of GDPR. By automatically comparing such policies it has been possible to determine to what extent GDPR stimulated changes, and to assess whether such changes went in the direction of enhancing data protection (fewer unlawful clauses).

The European Parliament, in its Report on Better Regulation, has stressed the importance of making impact assessments for all legislative proposals and the need of providing adequate resources.

> *[The Parliament] calls on the Commission to perform impact assessments on all legislative proposals [...]; calls for a sufficient amount of time and resources to be allocated to impact assessments in order to ensure their quality […] calls for impact assessments to be published immediately upon their completion, and not only when the policy proposal is presented, thus ensuring greater transparency on how EU decisions are taken; acknowledges that the effective implementation of better regulation and, in particular, of the ex-ante impact assessments will require an appropriate level of resources; urges the Commission to allocate the appropriate means in this regard.* [47]

[47] European Parliament (2022).

7. COMPLEXITY AND UNCERTAINTY IN ASSESSING AND EVALUATING IMPACTS

KEY FINDINGS

Even when impact assessments rely on advanced statistical or AI methods, they remain highly conjectural, since they involve predicting behavioural changes and further direct and indirect effects of such changes. The uncertainty of ex-ante assessments, even when based on the best methodologies and on adequate datasets, underscores the need for interim monitoring and ex-post evaluation. Both are needed to check the extent to which ex-ante predictions are confirmed or rather contradicted by subsequent facts, as well as the extent to which unexpected side-effects emerge. Ex-post monitoring and evaluation provide evidence of real outcomes, supports democratic dialogue with evidence, and contributes to making governments accountable toward elected assemblies, such as the European Parliament.

The assessment of impacts —even when relying on advanced statistical, computational, and AI methods— remains highly conjectural. This is due to the combination of different issues, such as the availability of data, the complexity of the regulated social contexts, the difficulty of causal analyses. As shown in the previous sections, the optimal assessment of regulations requires a vast amount of data, which may not always be available. Even when sufficient data are available, pervasive uncertainties may persist, since the assessment of regulatory impacts requires us to specifically determine what differences a legislative measure will make. Hence the need of identifying causal connections, rather than merely detecting correlations and projecting them into the future.

Figure 10 shows the complexity of the evaluations involved in ex ante regulatory assessment. Uncertainty is due to a need to estimate to what extent a regulation will prompt behavioural change and what further outcomes this would trigger in society, interacting with multiple further influences, within evolving and complex social contexts.

These uncertainties should not be underestimated, and consequently, we should keep in mind that ex-ante assessments are only conjectural. In some cases, uncertainty can be treated mathematically, since we can assign probabilities to the occurrence of future events, but there are also cases in which uncertainty also covers these very probabilities, i.e., we do not know what likelihood there is that a future event will happen. This is often the case where the impact of new factors has to be assessed (e.g., technological innovations), or when complexities are involved (as in many ecological and social contexts).[48] In such cases quantities should only be assigned in full awareness of their uncertainty (and of the fact that the uncertainty will inevitably spread to the implications of such quantities).

[48] Hansson (2016).

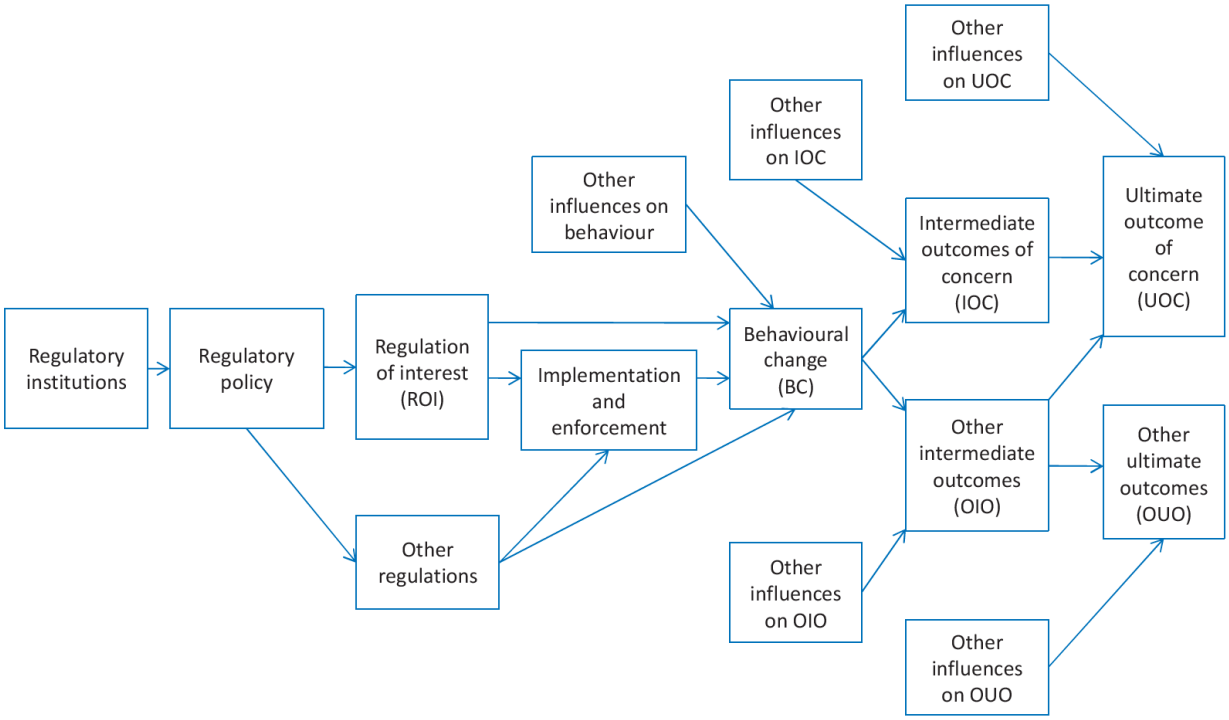


Figure 10. The object of ex-ante regulatory assessment[49]

The uncertainty of ex-ante assessment, even when made by relying on the best methodologies and on adequate datasets, make it necessary to emphasise the significance of both interim monitoring and ex-post evaluation. Both are meant to check the extent to which the ex-ante predictions are verified or contradicted by subsequent facts, as well as the extent to which unexpected side-effects have emerged. Thus, monitoring and evaluation can provide evidence for corresponding adjustments and revisions of policies. More generally, through monitoring and evaluation, policymakers can be held to account for the real (rather than the expected) outcome of their policies, and the very process of *ex ante* assessment (the data collected, the methodology used, etc.) can be subject to critical examination.

It has been argued that, while the traditional policy cycle is characterized by evaluations happening at the very end of policymaking, it should now be possible to "take advantage of instantaneous or near-instantaneous data processing", so that "evaluation results become available the very moment data arrives."[50] Rather than being a neatly defined step at the end of the policy cycle, the evaluation of policies could take place continuously, opening permanent possibilities of reiteration, reassessment, and consideration.

On the other hand, it has also been claimed that policy evaluation should not be confused with monitoring. Monitoring consists in continuously checking implementation in relation to an agreed schedule. It involves the systematic collection of data on specified indicators to give management and the main stakeholders a sense of the progress and achievement of the objectives and of the delivery of outputs and outcomes. On the contrary, evaluation involves a deliberate and responsible "looping back" into the regulatory cycle.[51]

[49] Coglianese (2012, 11).
[50] Hochtl et al (2016, 162).
[51] Allio (2015, 193).

It seems to us that that both monitoring and evaluation are valuable ways to retrospectively assess policies and provide useful feedback: they can be optimally integrated by using the information extracted through monitoring as input in evaluation, while using evaluation to assess how well the monitoring is working and whether it ought to be expanded.

The ex-post monitoring/evaluation, while being complementary to the ex-ante assessment, is grounded on specific reasons.

In the first place, interim monitoring and ex-post evaluation provide evidence of the real outcomes of regulative measures. Given the uncertainty inherent in social predictions, these outcomes may differ, in size and kind, from those predicted through ex-ante assessment, even when the latter has adopted appropriate methodologies. Thus, it may be necessary to fine-tune, tailor and complement such measures, in order to better achieve their intended goals and prevent or mitigate unwanted side-effects. It has also been observed that the systematic use of ex-post evaluations can engender in decision-makers an aptitude for openness and learning:

> *carrying out retrospective evaluation and analysis is helpful in keeping an open mind as it encourages an ongoing learning from experience and stimulates efforts to adapt future policy as a result. Putting in place mechanisms to gather, and apply, new insights set an expectation that lessons will be learnt, and new insights gained.*[52]

Secondly, ex post monitoring and evaluation play an important role in the democratic debate. They enable civil society to scrutinise the real outcomes of the policy process and check whether government goals have been achieved and whether public resources have been effectively managed. The availability of data on the impact of policies can improve the democratic dialogue, making it less ideological and more evidence-based.

Thirdly, when ex-ante assessments of legislative proposals are entrusted to governments (in the EU, to the Commission), interim monitoring and ex-post evaluation enable Parliaments to control the effectiveness and efficacy of the measures proposed, assessed, and implemented by governments. Thus, they contribute to ensuring the accountability of governments towards elective bodies. Is has indeed been affirmed that Parliament should play a key role in ex-post evaluation:

> *Parliamentary committees and units should place themselves at the apex of the accountability structure and make efforts to be widely known as the prime location and focus of ex post legislative evaluation so that information, research and analysis is submitted to them as a matter of routine. In some countries, like Australia or Canada, one central motivation of ex post evaluation by the legislature is to make a judgment on the effectiveness of the RIA [regulatory impact assessment] and seek improvement from the executive when this is shown to be required.*[53]

The effectiveness of ex-post evaluations is highly dependent on the quantity and quality of the available data:

> *Increasing the amount of data associated with the outcome of a given policy allows for agile policy adjustments in the short term, but more importantly will generate better insights into the policy process in terms of accountability and learning in the mid- to long term. Those responsible for a given policy can establish whether their policies have had the desired effect or not and, if*

[52] OECD (2019d, Section 3).

[53] Allio (2015, 198).

> *those data are published as OGD [open government data], so can other stakeholders. As a result, policy evaluation can turn into an open, inclusive and ongoing process rather than an internal, snapshot moment.*[54]

The EU Parliament —through its Research Service's Directorate for Impact Assessment and European Added Value as well as through the DG IPOL's Policy Departments— is already playing a significant role in *ex post* evaluations, but our view is that this role should be expanded and facilitated.

Given the importance of regulatory assessments and evaluations for the EU Parliament, the JURI Committee should consider setting up a permanent Working Group on Better Regulation, to ensure a more active and persistent critical involvement by the Parliament. The Working Group should systematically contribute to identifying shortcomings and proposing improvements, with a special focus on data practices and corresponding technologies.

As noted above, the approaches to regulatory assessments and evaluations are today rapidly evolving, in connection to aspects such as accelerated dynamics of the regulated domains, the need for more comprehensive multi-criteria appraisals, the availability new digital technologies. In this context the involvement of academic and research institutions in reviewing current practices, identifying their strength and weaknesses, and proposing innovative methodologies, and technologies could be considered.

[54] OECD (2019d, 94).

8. CONCLUSIONS AND POLICY SUGGESTIONS

This report has discussed the use of data in the public sector, focusing on ex-ante regulatory assessments and ex-post evaluations. The following indications emerge from the present inquiry:

1. The public sector should catch up with the private sectors in the capacity to collect and use data.
2. Better and larger datasets should support the policy cycle, for the purpose of forecasting, foresight, policy design, assessment, and evaluation.
3. The cost and benefits of data collection, curation, and use should be considered, and addressed by adopting cost-effective solutions, also involving the reuse and repurposing of data. The statistical processing meant to provide aggregate information should be distinguished from the processing of personal data meant to provide individualised outcomes.
4. The opportunities offered by AI for collecting and interpreting data, extracting aggregate information through analytics, and exploring scenarios and possible developments through simulation, should be exploited.
5. The need to support multicriteria evaluations with appropriate data, according to methods such as the UN Sustainable Development Goals, should be considered. Quantitative metrics should be used whenever possible, consistently with the nature of the data and the goals pursued.
6. Compliance with law and ethics should be ensured, as to uphold the rule of law, respect citizens' rights, and foster trust.
7. Predictions should be matched against reality through monitoring and evaluation. An ex-post evaluation should accompany every ex-ante regulatory assessment, making it possible to adjust policies, in a process that contributes to democratic debate and makes for greater government accountability.
8. The JURI Committee should consider setting up a Working Group on Better Regulation, in order to expand and facilitate the role of the European Parliament in participating in the Better Regulation process and in contributing to its improvement and reform.
9. A platform involving academic and research institutions could be created, with the task to critically examine the Better Regulation process, identify strength and weaknesses of it, and propose methodologies, and technologies to make it more effective,

REFERENCES

- Agrawal, A., J. Gans, and A. Goldfarb (2018). *Prediction machines*. Harvard Business Review Press.
- Allard, S. W., W. E. R., C. Schlecht, A. R. Datta, R. M. Goerge, and E. Weigensberg (2018). State agencies' use of administrative data for improved practice: Needs, challenges, and opportunities. *Public Administration Review*, 78: 240–50.
- Allio, L. (2015). Ex post evaluation of regulation: An overview of the notion and of international practices. In *Regulatory Policy in Perspective: A Reader's Companion to the OECD Regulatory Policy Outlook*, pp. 191–237. OECD.
- Ashley, K. D. (2017). *Artificial Intelligence and Legal Analytics*. Cambridge University Press.
- Birks, D., A. Heppenstall, and N. Malleson (2020). Towards the development of societal twins. In *24th European Conference on Artificial Intelligence - ECAI 2020*.
- Boardman, A. E., D. H. Greenberg, A. R. Vining, and D. L. Weimer (2018). *Cost–Benefit Analysis*. Cambridge University Press.
- Coglianese, C. (2012). *Measuring Regulatory Performance: Evaluating the Impact of Regulation and* Regulatory Policy. OECD.
- European Commission (2019). *Commission staff working document: Taking stock of the Commission's Better Regulation Agenda accompanying the document Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Better regulation: taking stock and sustaining our commitment*. COM (2019) 178 final, SWD (2019) 156 final.
- European Commission (2020). *Impact Assessment Accompanying the Document Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, SWD (2020) 348 final.
- Davenport, T. H. and J. Harris (2017). *Competing on Analytics*. Harvard Business Review Press.
- Davidson, P., C. Kauffmann, and M.-G. de Liedekerke (2021). *How do laws and regulations affect competitiveness: The role for regulatory impact assessment*. OECD.
- Engstrom, D. F., Ho, D. E., Sharkey, C. M., and Cuellar, M.-F. (2020). *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*. Administrative Conference of the United States.
- European Parliament. Committee on Legal Affairs (2022). *Report - A9-0167/2022. Better regulation: Joining forces to make better laws*.
- Floridi, L. and M. Taddeo (2016). What is data ethics? *Philosophical Transactions of the Royal Society A*, 374.
- Fuller, A., Z. Fan, C. Day, and C. Barlow (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE Access* 8, 108952–108971.
- Guerrero, O. A., and G. Castañeda (2020). Policy priority inference: A computational framework to analyze the allocation of resources for the sustainable development goals. *Data and Policy* 2, e17–31.
- Halpern, J. Y. (2016). *Actual Causality*. MIT.
- Hamill, L. and N. Gilbert (2016). *Agent-Based Modelling in Economics*. Wiley.

- Hansson, S. O. (2010). Cost–benefit analysis: Philosophical issues. In *The New Palgrave Dictionary of Economics*, pp. 1-4. Palgrave Macmillan.
- Hansson, S. O. (2016). Evaluating the uncertainties. In *The Argumentative Turn in Policy Analysis Reasoning about Uncertainty*, pp. 79–104. Springer.
- Heikkila, M. (March 29, 2022). Dutch scandal serves as a warning for Europe over risks of using algorithms. *Politico*.
- Höchtl, J., P. Parycek, and R. Schöllhammer (2016). Big Data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce,* 26: 147–169.
- Ishizaka, A. and P. Nemery (2013). *Multi-Criteria Decision Analysis*. Wiley.
- Keeney, R. and H. Raiffa (1993). *Decisions with Multiple Objectives: Preferences and Trade Offs*. Cambridge University Press.
- Lippi, M., P. Palka, G. Contissa, F. Lagioia, H.-W. Micklitz, G. Sartor, and P. Torroni (2019). Claudette: an automated detector of potentially unfair clauses in online terms of service, 27: 117–39.
- Margetts, H. (2022). Rethinking AI for good governance. *Dedalus,* 151: 360–71.
- Margetts, H. and C. Dorobantu (2019). Rethink government with AI. *Nature, 568*:163–5.
- Mayer-Schoenberger, V. (2022). *Access Rules: Freeing Data from Big Tech for a Better Future*. University of California Press.
- Micklitz, H.-W. (2022). The measuring of the law through EU politics. In *The politics of European legal research*, pp. 223–238. Elgar.
- OECD (2019a). *Analytics for Integrity Data-Driven Approaches for Enhancing Corruption and Fraud Risk Assessments*. OECD.
- OECD (2019b). *Better Regulation Practices across the European Union*. OECD.
- OECD (2019c). *Digital Government Review of Argentina Accelerating the Digitalisation of the Public Sector*. OECD.
- OECD (2019d). *The Path to Becoming a Data-Driven Public Sector*. OECD.
- Palmirani, M., Vitali, F., Van Puymbroeck, W., and Nubla Durango, F. (2022). *Legal Drafting in the Era of Artificial Intelligence and Digitisation.* European Commission.
- Renda, A. (2015). Regulatory impact assessment and regulatory policy. In *Regulatory Policy in Perspective: A Reader's Companion to the OECD Regulatory Policy Outlook*, pp. 35–114. OECD.
- Renda, A. (2022). *Assessment of current initiatives of the European Commission on better regulation*. European Parliament.
- Siôn, J., G. Dohler, and L. Pate (2022). *Better regulation in the EU: Improving quality and reducing delays*. European Parliament.
- van Ooijen, C., B. Ubaldi, and B. Welby (2019). *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance*. OECD.
- Varian, H. R. (2010). Computer mediated transactions. *American Economic Review: 100*, 1–10.
Zerbe, R. O. J. (2006). *A Primer for Benefit-Cost Analysis*. Elgar

This in-depth analysis, commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs at the request of the JURI Committee, looks at the use of data for the purpose of regulatory assessment/evaluation. The author finds that data is needed to support evidence-based regulation, that information technologies, and in particular AI, can enable a more extensive and beneficial use of data, and that the use of data in ex-post evaluations can improve the regulatory process. The in-depth analysis offers policy recommendations.

PE 736.129

IP/C/JURI/IC/2022-020

Print ISBN 978-92-846-9704-5 | doi: 10.2861/079 | QA-07-22-778-EN-C
PDF ISBN 978-92-846-9703-8 | doi: 10.2861/183840 | QA-07-22-778-EN-N