

From Facts to Decision Data: About the Factual Basis of Automated Individual Decisions

Dag Wiese Schartum

1	Introduction	380
2	Facts in Rules	381
3	Simplifying the Processing of Facts by Reuse and Common Use	384
4	Simplifying the Processing of Decision Data by Means of Self-Service	388
5	Directly Capturing Facts of the Physical World	389
6	Combining Sources of Decision Data	390
7	Regulation of Facts in Individual Decision-Making Processes and Digitalisation-Conscious Legislation	393
8	Modernising Public Administration Legislation	397
9	Concluding Remarks	399

1 Introduction

Digital or “electronic” government¹ has been one of the pillars of research at the Swedish Law and Informatics Research Institute (IRI), with Professor Cecilia Magnusson Sjöberg’s doctoral thesis (1992) as the foremost contribution.² Digitalisation of government administration was already debated in 1969, the year IRI Stockholm was established. The founder of IRI Stockholm, Professor Peter Seipel, summed up discussions from one of the very first seminars on law and EDP³ at Stockholm University.⁴ The overall issue of the seminar was “automatic data processing within the legal system”, and the subjects that were discussed included “design and application of algorithms for the purpose of legal application of rules ...”.⁵ Several of the more specific problems touched upon are still of current interest: the danger of a less dynamic application of legal norms; reduced, unfettered discretion; and the advantages of using decision tables for the formulation of legal norms.

My ambition in this article is to analyse some aspects of data describing facts as part of digital case processing of individual cases: How may digital government administration be organised in order to identify, interpret, formalise, collect, process and transfer data representing facts of the case? I will limit the discussion to the processing of decisions in individual cases concerning rights and obligations of citizens pursuant to legislation. This is a type of decision-making where digitalisation in the Nordic countries has come relatively far, with a rather high level of automation. My attention will be on automated collection of data constituting the substantive basis for individual decisions. In the following, I denote such data by the term “decision data”.

What I refer to here as “factual” and “facts” are physical and psychological characteristics of individuals, their social relations, as well as events, states of conditions in their environment that could be connected to them etc. In the article, I aim primarily to provide a basic, systematic description of how decision data may appear and may be processed. I will also touch upon effects. Decision data will largely amount to “personal data”.⁶ Thus, the processing of decision data raises a series of data protection issues. Although I will give some attention to the General Data Protection Regulation (GDPR), Article 22, regarding automatic processing, I will not discuss protection of personal data in particular. Rather, I will more generally maintain a rule of law perspective.

1 “Digital government” and “electronic government” are synonymous concepts.

2 See, Magnusson Sjöberg, Cecilia, *Rättsautomation. Särskilt om statsförvaltningens datorisering*, Norstedts Juridik, Stockholm 1992.

3 Electronic Data Processing.

4 The seminar was held 17 April 1969 by the Working Group for EDP and Law (in Swedish: ”Arbetsgruppen för ADB och Juridik”).

5 See, Seipel, Peter, *circular letter number 3*, 18 April 1969.

6 Cf. General Data Protection Regulation, article 4(1).

2 Facts in Rules

Government administration in Sweden, Norway and many other countries is largely regulated by Acts of Parliament and delegated legislation. Such statutory laws often contain detailed rules concerning the substantive contents of the various government arrangements and welfare schemes, encompassing pensions, benefits, public services, taxes and duties etc. Legislation, for example, establishes conditions for admission to universities, the right to receive unemployment benefits, the obligation to pay income tax, custom duties etc. Such laws are often designed and formulated in ways that enable – but do not necessarily facilitate – expressing rules by means of programming language; rendering it possible – wholly or partly – to automate corresponding decision-making in individual cases.

To decide in individual cases, the relevant government agency needs to have decision data regarding all legally relevant circumstances of each case. This is typically linked to two main types of facts: The first type of facts refers to circumstances that must be met if the party of the case should be seen as qualified for a certain right or as having obligation to comply with certain duties. The second type of facts are those needed to determine the detailed result of the decision. Some facts, for instance, are required to establish tax liability (cf. first type of facts), while other facts are necessary to establish the exact taxable income, assets etc. (cf. second type of facts).⁷ To put this differently, facts belong to both sides of the sequence:

IF fact, THEN legal effect.

Putting it this way, facts related to legal effects are not made visible, but become evident if we detail the right hand side of “THEN” in the sequence above:

IF fact₁ THEN add (fact₂ and fact₃) GIVING decision (containing fact₄ and fact₅).

Of course, real legislation contains much more complex rule structures. Here I will illustrate this by expanding the left hand side of the sentence one step further:

IF fact₁ AND (fact₂ or fact₃) OR (fact₄ and fact₅) THEN add (fact₆ and fact₇) GIVING decision (containing fact₈ and fact₉).

As already indicated, I use “fact” in a very broad sense, encompassing all circumstances that are qualified pursuant to relevant valid legal sources. Thus, facts could refer to people and objects and describe their physical appearance, opinions, qualities, whereabouts, position in time and space etc. Facts may also refer to assessments and expert assessments etc. For instance, if a government body has established something as “justified”, the results of that assessment

⁷ Sometimes, there is only one result possible and facts to establishing detailed results are unnecessary. No additional facts are given in a situation where people apply for Norwegian citizenship and meet the condition: There is one type of citizenship, and facts to determine details of the decision are thus irrelevant.

could be seen as a “fact”. Expressed in a different way, the word “fact” in this article denotes every legal relevant input to and output from the processing of individual cases.

Facts may also have an *output* from decision processes that results in input in new cases. The decision, for instance that a person is eligible to receive unemployment benefit at the sum of NOK 20.000 per month could be qualified by other laws as relevant *input* (basis of decision) for future decision processes, for instance regarding assessment of income taxes.⁸

Legally relevant facts are typically described by words and phrases in statutory law, and we therefore need to understand the semantics of these descriptions and assess whether or not a fact is of the type described by law. If it is confirmed, we need to establish a legally correct representation of that fact and submit this decision data to the data system. To perform as described, we need two types of knowledge: one about the law and one about the facts. In other words, we need i) to interpret the semantics of words and phrases of statutory law and other legal sources describing facts, ii) to assess potential corresponding legal relevant facts of individual cases, and iii) subsume by establishing a legally correct representation of the facts, see Figure 1.

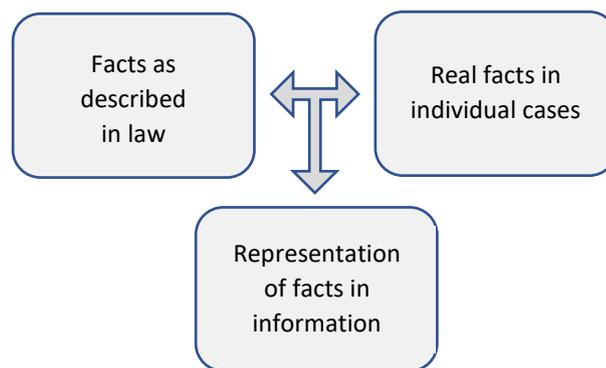


Figure 1: Assessment and representation of legally relevant facts.

It follows from the explanations above, that facts in law are often an integral part of legal conditions, for instance conditions that two people are married, that the average income for the previous calendar year must exceed NOK 200.000, or that the applicant is a Norwegian citizen. In other common situations, facts in law are elements of calculations, without being a condition. Income figures, for instance, are needed to calculate one’s total income in order to establish tax liability.

In order to establish facts of cases, we need first to understand the description of facts in relevant legal sources, with a word or phrase in a statutory text as

⁸ Cf. van Eck, Marliens, *Geautomatiseerde ketenbesluiten & rechtsbescherming* (“Automated Administrative Chain Decisions and Legal Protections”), Tilburg University, 2018. See English abstract, page 439–450.

starting point. What does the phrase “support a child” mean, for instance? (What does “support” mean, who may be supporter, what is a “child” in this context (cf. natural, adopted and foster child)? – and so on. Typically, statutory law will only give answers to a few of these questions. Many issues of interpretation will have to be resolved through more thorough legal methods, taking into account case law etc.

Based on a general understanding of facts as described in law, the task will be to consider if the facts of the case at hand may be subsumed under our understanding of the facts as described by law. In typical cases of “child support”, we have a natural child under the age of 18 living with one or two of the parents. If so, there is hardly room for doubt regarding the application of the law to the facts. In other cases, there may be factual elements creating serious doubts regarding what should be regarded a legally correct subsumption: Does a situation where a 17-year-old child from a previous relationship, living in the same household as the surviving husband imply that this child, pursuant to the law, is entitled to support by the stepfather? What if the child has his own income, and what if the child rarely visits home more than twice a week? Applicable legal sources may give us answers, but chances are good that conclusions will be uncertain at best.

In a totally manual processing of cases, i.e. where reasoning about the legal and factual questions of individual cases rests totally on cognitive processes, it may be that we allow ourselves to be in doubt concerning who should “support a child”. We wait and consider other aspects of the case, and in the final analysis make a holistic assessment in which we include public policy considerations. Automated legal decision systems in digital government, in contrast, would require that facts be described in certain formats and processed in a certain sequence. Expected input value, for instance, might be “yes” or “no”. “Do not know” will be impossible – the computer will always require that explicit choices be registered in the program.⁹ Representation of facts in data systems therefore requires us to make a formal, specific and clear representation of each legally relevant fact; they cannot be absorbed in a decision-maker’s unfettered discretionary assessment.

Formalisation of facts in data systems mirrors that the system may process only facts that are explicitly made a part of the system. One possible effect is that some types of facts cannot be processed as decision data in the system, even though they are claimed to be legally relevant. For instance, the legal criteria is that someone must “support a child under the age of 18 years”, and in the data system, the input has to be “yes” or “no”. If a child is 18 years old, as in this situation, and statements in the preparatory works of the relevant Act, as well as the statement of legal purpose, support the possibility that certain disabled children at 18 years of age might in rare instances be accepted in line with children younger than 18 years, there will be no way to make that fact part of what the system processes. If we accept the system as legally valid, this implies that the formalisation requirements and system design hinder certain legal interpretations that the legislator has deemed to be possible interpretations.

9 Provided that the condition about support of child is necessary to clarify.

It follows from this brief discussion of formal representation of facts that we must know whether or not the data system is legally valid, i.e. that it can accommodate legally correct representations of all valid facts of cases concerned. Any deviant situation would be unacceptable, because it would introduce uncertainty regarding the correctness of decisions produced by the system and would always provide grounds for claiming manual processing or lodging complaints. Requirements for how facts may be represented in a data system must thus be seen as binding decisions.

3 Simplifying the Processing of Facts by Reuse and Common Use

At least two conclusions may be drawn on the basis of the description in the previous section:

- i) It takes legal expertise to determine the meaning of facts described by law, and
- ii) it may be labour-intensive to subsume facts from real-life situations under legal facts.

In digital government, these two observations are often regarded as a major challenge. The basic reason is that digitalisation is about more efficiently carrying out tasks in government administration.¹⁰ Where the legal regime offers flexibility in describing facts by means of vague concepts and discretionary assessments, efficiency improvement and digitalisation pull in the direction of simplification and formalisation of the way facts are established. Such simplification and formalisation are prerequisites for a large amount of the automated decision-making. Data systems must be programmed to automatically collect decision data from digital sources, containing relevant data in the right technical format.

Here I will briefly examine two well-established main strategies for simpler and more effective processing of decision data. I will present the strategies as separate from one another, although in practice they might just as well be applied in combination; cf. the end of this section.

The first strategy is based on the recognition of legal facts as something valuable and complex. This strategy is not directed towards simplifying the law, but instead trying to make as effective use as possible of the efforts that have already been invested in establishing facts. The overall keyword is “sharing” of

¹⁰ See, e.g. Ministry of Local Government and Modernisation, Norway (2016) *Digital agenda for Norway in brief – ICT for a simpler everyday life and increased productivity*, section 1, IV: “Common solutions shall be used to create effective, user-friendly digital services for the whole of the public sector.”, available from “www.regjeringen.no/en/dokumenter/digital-agenda-for-norway-in-brief/id2499897/”.

decision data and “once only”,¹¹ i.e. reuse of data derived from previous processes of interpretation and subsumption, or “chained use” of decision data from one agency in a decision-making process of another government agency. The premise for this strategy is that the first government agency has established decision data pursuant to their special legal sources, which also is in accordance with requirements for decision data pursuant to legislation governing decisions of the second government agency.

This first strategy could be divided into three substrategies:

- Reuse of decision data which have been the *basis* of previous individual decisions
- Use of decision data which represent *results* from individual decisions that have been made
- Use of data that have been established as a *common* government service.

Firstly, it is possible to reuse data that have served as the basis for existing individual decisions. The tax administration, for example, bases its decisions on an assessment of “income”, or “support a child” or “Norwegian citizen”. Each of these established decision data do not individually represent decisions, but constitute the basis for decisions regarding tax liability. Government bodies have an obligation pursuant to the principle of investigation. If fulfilled, this obligation results in facts that have sufficient legal quality (cf. correct, complete, precise, updated etc). If facts established pursuant to other branches of law are based on the same conceptual understanding of, for instance, “income”, “support a child” etc., then data representing these facts may be reused. To the extent these data exist in machine-readable formats, collection and further processing may be digitalised.

Secondly, there is a possibility to use the facts that express individual decisions. Decisions by the Norwegian Directorate of Immigration regarding conferral of citizenship could be seen as a type of decision data. Such decisions thus represent binding establishment of a certain legal fact. When valid decision is made, citizenship for the person in question has been established as a fact deriving from exercise of government power. The relevant decision data is thus more stable and reliable than facts that have been considered only as the *basis* of decisions.

Thirdly, even though no government body or court of justice has made a decision and thereby established a new legally valid representation of facts, official registration arrangements may constitute a collection of relatively stable data. If, for instance, persons make a reservation against digital communication with government agencies, this does not represent a formal government decision but is nevertheless established as a fixed and binding fact. Similarly, people may choose to register themselves as owners of a car in the vehicle registration system, or register themselves as residents at a certain address in the Population Register.

¹¹ See, e.g. *European Interoperability Framework Promoting seamless services and data flows for European public administrations*, European Commission (2017), page 16, available from “ec.europa.eu/isa2/eif_en”.

Legal obligation or the right to register, combined with an appurtenant government data system, may create the expectation that such facts must be observed. Obligations to register, moreover, may make it legitimate for legislators to establish by law that information in such registers has certain legal effects. Applications for housing benefits are decided on the basis of residence data as registered in the Population Register by the party in question, because according to law, everybody has an obligation to register any change of residence.¹²

The three types of sub strategies mentioned could be applied in relation to four types of actors: one's own government agency, other government agencies, special government agencies, private businesses. Here, I will briefly explain the possibilities.

	Same gov. agency	Other gov. agency	Gov. registration agency	Private businesses
Reuse of decision data being bases of previous individual decisions	x	x		x
Use of decision data in results of previous individual decisions	x	x		x
Use of officially registered data			x	

Table 1: Strategies for collection of decision data and possible actors.

Here, I will limit myself to some brief comments regarding each of the four groups of actors. "Same government agency" indicates the possibilities of accessing data pertaining to facts within the same agency that requires or collects this information. Several government agencies have comprehensive responsibilities covering many different administrative schemes and arrangements, each of them handled by means of separate legal decision systems. National welfare agencies and tax administrations are typical types of agencies in which one might expect to find many unexploited in-house possibilities for retrieving stored decision data from previous decisions. The situation is similar for "other government agencies". Central government agencies are probably those with highest potentials as sources for other government agencies.

"Government registration agencies" in Table 1 refers to government agencies established with the purpose of supplying government administration (and other parts of society) with certain types of data. Important examples are agencies with responsibility for the land registry and cadastre, population register and central coordinating register for legal entities. Identification of data from "other government agencies" with a view to potentially reusing the data is an overall

¹² See, Act of 9 December 2016-12-09 no. 88 regarding national population register, § 6-1.

assumption shared by all government agencies in Norway as part of digitalisation processes and efforts to increase the efficiency of government administration. Thus, while available data from government registration agencies are the result of planned establishment of services, the potential sharing of data between two or several agencies is typically something that is discovered and developed over time.

“Private businesses” is the last category of possible sources of data in Table 1, above. I use this term in a broad sense,¹³ the main point being that private businesses are not organised as government bodies.¹⁴ Private businesses may have data about people in their capacity as employees, customers, clients, members etc., and the data may have great relevance and potential as the basis of decision-making in government agencies (e.g. information about income, qualifications, health, type and contents of services, shopping, membership fees etc.). Private businesses will usually have data systems containing information of potential interest for collection and use by government agencies. Their data will furthermore often represent an economical potential for government agencies. If private data systems could supply decision data online in certain formats corresponding to government requirements, this could represent big savings on government budgets.¹⁵

Reuse of data held by different types of government and private enterprises representing established facts relevant to government decisions in individual cases could obviously be an efficient strategy to reduce needs for expertise and labour in decision-making bodies. To the extent that technical formats of data and online communication set-ups are standardised, this strategy contributes to possibilities of fully automated decision-making routines: Both collection of data and further processing of them could be automated. Applied to private businesses, however, this strategy requires that very many different actors hold data referring to the same semantical content, and that these data have sufficient quality to be used as a basis of individual governmental decisions. Thus, a minimum prerequisite related to private businesses will often be to have a clear legal basis laying down the obligation to provide certain data in a certain format, at a certain time etc.¹⁶ In terms of the government sector, access to data sources could sometimes be based on softer measures, for instance internal agreements or instructions from a superior authority, for instance a ministry.

13 Even state enterprises and NGOs should be included in this category.

14 But could be owned by the government. Individual persons could also be seen as possible sources of information, but in practice only legal persons are eligible.

15 While at the same time introducing new costs to private businesses. Thus, the obligation for private businesses to deliver data to government agencies must have statutory basis.

16 Cf. the principle of legality.

4 Simplifying the Processing of Decision Data by Means of Self-Service

To a certain extent, the Internet has made it possible to establish self-service decision-making arrangements within government administration, where citizens are asked to contribute to questions regarding legal matters and, possibly, to implement actions on their own initiative or pursuant to a legal obligation or right they have. In line with such possibilities, a strategy for collection of decision data that are alternative to those described in section 3 above could be to allow individual parties– or their representatives – to assess case-relevant facts and register decision data in their case. Citizens may thereby become “their own officer in charge” and make required interpretations of applicable law, subsumption of relevant facts, as well as representation and registration of these facts in the relevant government data system.

Self-service as described is associated with some obvious limitations and problems. Firstly, there is the limitation that a fraction of the population is not “digital” and cannot be reached by means of online routines. Even when people are “digitally available”, some will not be motivated or willing to process their own case. In addition to these limitations, it may obviously be a problem to safeguard a sufficient quality of citizens’ contribution as officers in their own case. Even though the majority of actions and inputs that are expected from citizens are “easy”, a small percentage – but possibly still a large number of cases – will represent hard cases and require legal expertise that cannot be expected from the average citizen. The example above in Section 2 about determination of a “supporter” may serve to illustrate this point.

The above-mentioned limitations could exist in certain situations on a manageable and acceptable scale, for instance because parties to cases are limited and homogeneous: The rate of university students with access to the Internet and ability and willingness to actively participate in online case processing, could be high, for instance. If the target group is broader and more heterogeneous, the number of people that could not be expected to fill their role as contributors to case processing, could easily be higher than what is legally acceptable without supplementary measures to capture fall-outs etc. Even if limitations are manageable and legally acceptable, problems regarding the legal quality of data registered by lay parties remain. Among measures necessary will be the preparation of adapted and sufficient information regarding required legal problem-solving, combined with routines in the data system giving automatic feedback in case of erroneous and inconsistent inputs etc.

Data quality problems related to self-service as described are, not least, associated with the risk that parties of individual cases are tempted to manipulate the decision of their case by giving distorted or incorrect inputs. Self-service routines require extensive and intensive information to the citizens in order to give them correct and adequate understandings of the legal questions they need to address and solve. In complex legal matters, even advanced data systems to support self-service activities will be unable to guarantee sufficiently correct results. Thus, this type of contribution from citizens may create the need for successive control, and necessary controls may, in some situations, be too

expensive compared to what the government agency gains from the self-service arrangement.

It is hardly possible to make general statements regarding acceptable risk of self-service routines as described. However, limitations and problems mentioned will easily imply that self-service seldom will be the preferred way of collecting decision data, when seen from the perspective of government administration. To the extent that this strategy is selected, it will most likely be a supplementary way of collecting data as input to government decision-making routines.

5 Directly Capturing Facts of the Physical World

So far, I have discussed four strategies for collection of data representing relevant facts for decisions in individual cases. The common feature of these strategies is that a person perceives the fact, assesses it and makes a representation of it in the data system of the relevant government agency. All activities are carried out based on some person's understanding of the applicable law. Facts, in other words, are transformed into decision data, by the sensory apparatus and mind of people. A *fifth* strategy of collecting data for decision-making purposes, is to avoid dependencies of cognitive contributions from people, and instead collect data describing legal relevant facts *directly*, by means of sensors, biometrics and similar apparatus with the capability to "read reality".

I use "sensor and similar apparatus" (in short, "sensors") to denote electronic devices designed to sense the state of a condition or events or change thereof; for instance regarding physical particularities of people, movement of objects and people, light, temperature, sound, radiation, chemicals etc. Sensors make it possible to measure and represent a state and change of affairs, for instance that an object is moving, how fast it is moving, in which direction, at what acceleration etc.; or changes of frequency of a person's heartbeat, heartbeat pattern (e.g. indicating heart disease) etc. The point here is that sensors can automatically collect streams of data concerning people, objects and other parts of the environment. Provided sensors can produce digital data, output data may be transferred to and processed by decision-making systems, and if desirable, be subjected to advanced analyses. Each stream of data may be processed separately and analysed (for instance regarding how a person travels), but may have a much bigger potential if combined with other sources; for instance data about topography (maps), body temperature, weather data, traffic data etc. In other words, sensors may enable government agencies to measure and analyse several aspects of people, objects and environments.

Data from sensors represents facts. These could be very simple facts based on one type of sensor data or be composite facts based on two or several sets of sensor data (movement, topography, traffic situation). Sensor facts may also be parts of "multifaced facts" where sensor data are combined with other types of data, for instance from statistics and data analyses of comprehensive and heterogeneous materials, cf. machine learning analyses of big data.

Sensor data may represent *legally relevant* facts, i.e. the legislator may establish that a certain bulk of compiled sensor data, or certain results from analyses of these data, should qualify as a certain legal fact, for instance as a

condition for a legal right. In Norwegian legislation, for example, one of several conditions for citizenship is that applicants have resided in Norway for at least seven years, and not been abroad for more than two months during any one of the calendar years.¹⁷ This type of condition could partly be translated into GPS parameters, and the question of residence could thereby be transformed into something that is not answered by people alone, but by machines. Another radical possibility that illustrates some of the potentials of sensor technology is to construct a “sickness” concept by means of data patterns from sensor technology combined with statistics etc. Individuals who have been diagnosed as “sick” pursuant to the sickness allowance scheme could automatically be declared fit to return to work again based on data stemming from wearable biosensors, GPS data logging of persons’ movements, stored medical data and statistical data related to the diagnosis in question etc.

Sensor data, alone or in combination with other data, could not only constitute conditions for legal obligations and rights, they could also be connected to legal *effects* and thus enforcement. Government decisions may, for instance, entail a prohibition against staying in certain areas or a ban on drinking alcohol or taking other intoxicating substances. Since sensors may register relevant data, sensor data may be used to detect any breach of previous decisions made and may thus make enforcement automatic and more efficient.

6 Combining Sources of Decision Data

In Figure 2 (below), I have combined and supplemented the above-described strategies for collection and processing of decision data in digital government. It shows two basic choices for the collection of data; human and machine based. Human assessment is carried out either by a case officer or by the party in question through a self-service procedure. Employment of case officers is an expensive strategy as seen from the machinery of government. Manual case processing is labour-intensive and is funded over government budgets. This question often receives much attention when government digital decision-making routines are redesigned in the course of digitalisation. Figure 2 shows that there are both alternative and supplementary ways to collect data.

17 Cf. Citizenship Act § 7 (1)(e) and Citizenship Regulation § 3-4.

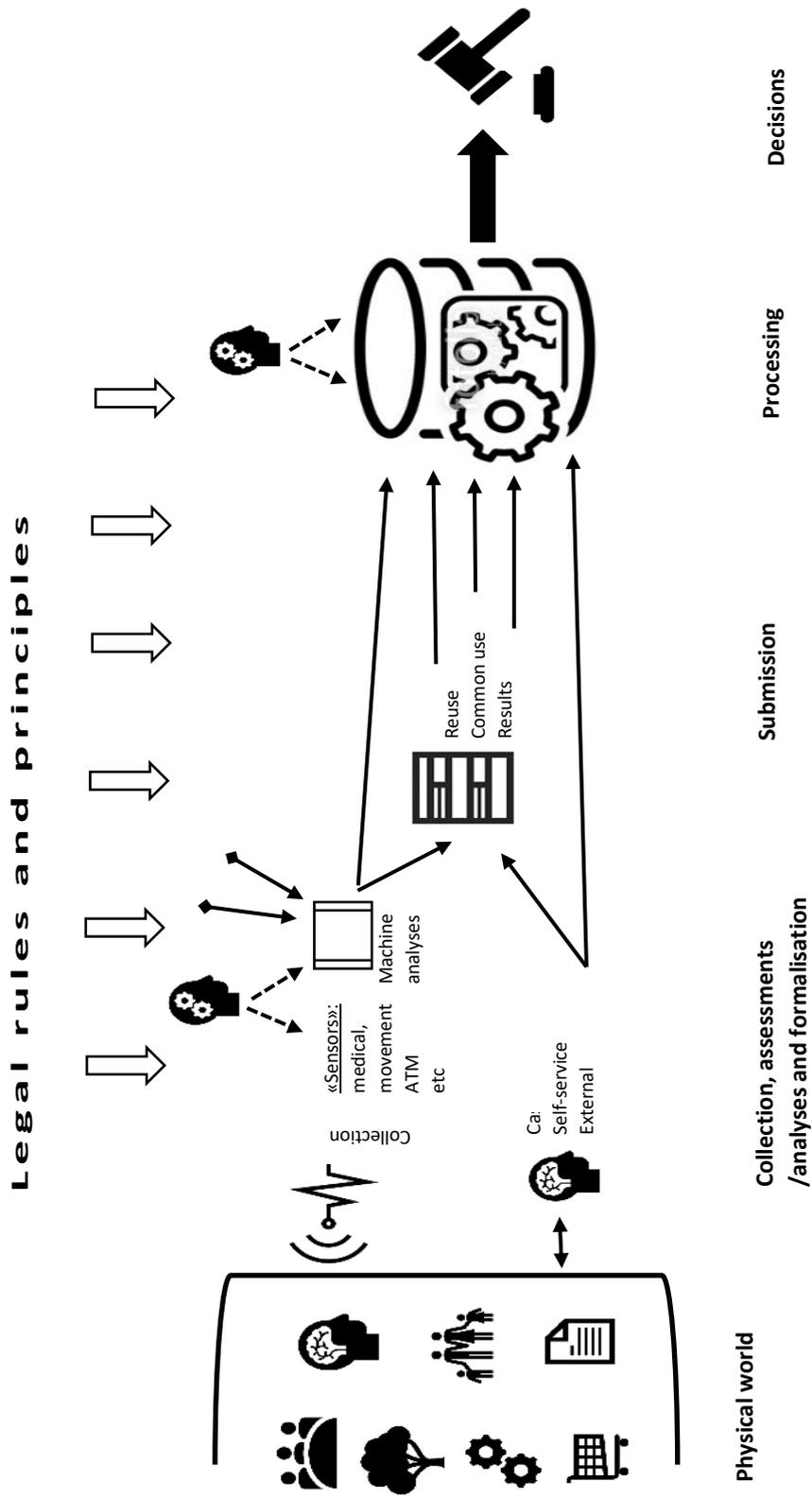


Figure 2: Survey of strategies for generating decision data, from relevant facts to bases of decisions.

The first alternative is to let someone else outside government administration (and off-budget!) perform manual assessment. In Figure 2, this is exemplified inter alia by “self-service”, that is, that parties assess and collect data pertaining to their own case. A related line of action is to let another government agency or a private business do the assessment and submit the results into the decision-routine (cf. “external” in Figure 2). Moving assessment and submission from one government agency to another in this way, does not in itself yield any cost reduction or efficiency improvement for the government sector. However, shifting over to private business may save the government money.¹⁸

One of the major strategies for improving efficiency and cutting costs is to make better use of manual assessments that are already made, by reusing data that have been the basis of decisions, or by utilising data expressing decisions made as decision data in chain-like decision-making processes. This type of solution could be seen as *reactive*, because the objective is to find out where in the existing decision-making processes a government agency can find suitable decision data to be used in *future* decision processes. In contrast, the *proactive* approach would be to design data systems with the aim of making available data that are intended to be the source for many decision processes (cf. “common use”).

In Figure 2 (upper left part), the use of sensors is marked as a main strategy for collection of data.¹⁹ In the figure, I have differentiated between collection and analyses. Typically, sensors, GPS and similar devices collect streams of data from the physical world. To be used in decision-making processes, these data must be analysed and processed in order to be transformed into aggregated data. In the next phase, these aggregated data will be submitted into the decision-making routine. For example, a stream of detailed GPS-data is collected but will only be submitted to the decision routine if the aggregated data shows the person concerned has stayed in areas of the city where he, according to decision from a public authority, is not allowed to be. Similarly, submission of heart beat sensors may only have effect on decisions if their values are higher or lower than certain defined limits. Machine analyses of such data may range from simple to advanced. Advanced analyses could imply application of data from other sources, for instance historical and statistical data etc,²⁰ possibly by means of big data analyses applying machine learning technology or the like.

The head icons with gearwheels in Figure 2 represent human decision processes that determine the content of automatic analyses and processing of data. The icons signify that even automatic processes are the result of human decision-making.²¹ With a human interface towards the world, case officers and others subsume facts of each individual case under legal rules. People creating

18 But would require a sufficient legal basis.

19 The word sensors is put between inverted commas to remind the reader that it is used in an specialised meaning, cf. section 5.

20 Cf. arrows with rhomb shaped start.

21 This is the case even when machine learning is applied, because necessary programming and manual training represents human decisions (despite the fact that humans do not survey possible effects of their contributions).

machine interfaces for sensors will, in contrast, decide in advance the general characteristics of the physical environment that should count as legally relevant. The core automatic decision-making process (“processing” in the figure) is in similar ways programmed, implying that legal rules determining the decisions based on collected facts are represented in the program, and are established in advance by people in the course of a system development process.

The last element of Figure 2 that I will comment on is “legal rules and principles”, at the top of the figure, with broad arrows pointing to every other part of the figure. Basically, this element represents the general assumption that in a government administration under rule of law, there will be legal rules and principles safeguarding fundamental rights and freedoms of citizens. This presumption will be elaborated on in next section.

7 Regulation of Facts in Individual Decision-making Processes and Digitalisation-Conscious Legislation

Facts derived from individual decisions in government administration – as previously mentioned – are to a large extent personal data, and thus the GDPR apply. Personal data are data that could be linked with an identifiable person. For instance, data from a GPS device only represents personal data to the extent that it is possible to connect these data to one or several physical persons.²² Physical persons are *in themselves* not “personal data”. Thus, GDPR apply from the point where personal data are deduced on the basis of human sensing or machine-based registration and analyses, and through their existence, until effective deletion or irreversible anonymization.

The GDPR leaves rather broad leeway for national legislation regarding, for example, the processing necessary for the exercise of official authority vested in the controller, cf. art. 6(1)(e). Firstly, member states may maintain or introduce more specific provisions to adapt the application of the rules of the regulation with regard to processing associated with the exercise of public authority as mentioned. Secondly, the legal basis for such processing should be laid down in national legislation, and this legislation may contain specific provisions to adapt the application of rules of the regulation concerning, for instance, types of data that are subject to processing, the purpose of processing, the disclosure of data etc. In other words, the regulation encourages the integration of provisions to safeguard data protection, within the framework of the regulation and into national special administrative legislation. The regulation does not restrict the national legislator’s authority to decide which type of personal data are to be applied as statutory conditions in order to receive benefits, be obliged to pay taxes and duties, be in the custody of child welfare authorities etc. Thus, the regulation has no direct effect on how each member state may design their

22 Cf. GDPR Article 4(1).

legislation in order to “modernise” their laws and prepare the ground for further digitalisation.²³

Already in the mid-1970s, scholars discussed “automation-friendly” and “digitalisation-conscious” legislation.²⁴ Automation-friendliness implies design of legislation so that digital representation of legal rules, particularly automatic application of law, becomes easy. Digitalisation-conscious is not specifically directed towards automation,²⁵ but could be said to represent a wide perspective encompassing many types and degrees of digitalisation. *Consciousness* in this regard points to awareness of possibilities for digitalisation, rather than the point of view that digitalisation should always be preferred. Here, I will not revive previous discussions, but limit myself to some reflections regarding assessment of facts and determination of decision data as a basis for individual decisions in digital government.

There are many possible degrees of automated decision-making. Fully automated decision-making routines presuppose a high degree of automated collection of decision data. There is no authoritative definition of what totally automated individual decisions means. The Article 22 of the General Data Protection Regulation regulates a “decision based solely on automated processing”, something which seems to presuppose total automation, but there is as yet no certain interpretation of where the lower limit of total automation should be drawn.²⁶ A reasonable delimitation from both a technical and legal point of view would, in my opinion, be to presuppose that decisions are based solely on automated processing if they i) are based solely on digital data sources, and that only these digital data are processed automatically, without persons having a possibility to influence the substantive output. My aim here is not primarily to argue in favour of a certain interpretation of GDPR Art. 22. Thus, I will not substantiate and pursue this line of argumentation further, but limit myself to applying this as a first basic definition of what totally automated individual decisions may mean. The implication is that individual decisions are completely automated if data sources and processing of these data are digital and automatic. In Norway, the processing of housing benefit cases and assessment of taxes for individual taxpayers are examples of types of processing where a very high percentage of cases are processed without interference from case officers or other decision-makers: Human decision-making is removed entirely from the layer of individual case-processing, and instead is present at the layer where general rules are formulated and imbedded in the programming code,

23 But the GDPR implies a number of legal requirements regarding how processing of data must be carried out.

24 See e.g. Bing, Jon, “Automatiseringsvennlig lovgivning”, in *Tidsskrift for rettsvitenskap* 1977, page 195–229.

25 But covers that too.

26 See discussion in Bygrave, Lee and Mendoza, Isak Esteban Sveinhaus, *The Right not to be Subject to Automated Decisions based on Profiling*, In Tatiana-Eleni Synodinou; Philippe Jougoux; Christiana Markou & Thalia Prastitou (ed.), *EU Internet Law: Regulation and Enforcement*. Springer, 2017, pages 77–98.

something that normally covers (almost) all necessary operations needed to process individual cases – from initiation of cases to final individual decisions.

Even though it is possible to regard the above-referred definition of totally automated individual decisions as fruitful and valid, it is only tenable if we delimit our analyses to the core of each decision-making process. Even if decisions are made “totally automatically” in line with the referred understanding of Article 22, such routines only apply to *secondary sources*. The primary sources will almost always – still – be based on manual analyses, assessments and formalisation by a case officer, by the individual party to a case as part of a self-service arrangement, or by another government agency or private business, or the like. Thus, in this perspective, individual decisions are not *really* totally automated unless automatic procedures cover everything from sensing all legally relevant aspects of the real world, to valid decision. Put in another way, real totally automatic individual decisions require the use of sensor technology or similar, i.e. technology that reads reality and thereby makes people redundant in the individual processing of cases.

Returning to the question of automation-friendly and digitalisation-conscious legislation, what I have pointed out above regarding primary and secondary sources of data demonstrates that the legislative strategies mentioned may be said to address two layers: one regarding regulation of *digitalisation of facts*, and a second regarding provisions regulating *decision-making* based on existing decision data from digital sources. This second layer represents challenges that have been addressed for many years. It encompasses the question of avoiding discretionary rules and favouring comparably stable and “fixed”, computable rules instead; the use of standardised concepts to denote data, preferably collected from machine-readable sources, the absence of “unnecessary” professional secrecy and other restrictions for access to required data, etc. The first layer is different because it raises topical questions of how digital equipment should be applied to digitise legally relevant aspects of reality in the process of individual decision-making. Below, I will briefly try to explore digitalisation-conscious regulation applied to this interface between the physical world and the digital representation of it.

The first group of possible types of provisions regulating digitalisation of the physical world that I will mention are those clarifying the legal status of digital representation. An “anti-digitalisation choice” would be to forbid digital representation and require that documents and other objects are authentic and are to be assessed by people. Even though I cannot exclude that this may be desirable and possible in very rare and limited areas of law, I will not comment further on this possibility. A much more frequently applied technique is that the law explicitly accepts digitalised representation. For instance, Section 2 of the Norwegian Administrative Procedure Act explicitly includes digital representation in the definitions of “document” and “in writing”. A tacit presumption, however, is that the digital representation already exists, either because it is originally created in digital form or because it is digitised. The Act does not regulate the actual digitalisation processes, and does not deal with questions of authenticity and data quality.

In general, I am not at all certain that legal regulations should encompass provisions regarding technical equipment for digitalisation, but the question

certainly needs to be considered. A comparable and relevant piece of legislation is that which regulates measures and weight etc., including digital measuring.²⁷ This legislation determines which quantitative measures are to be applied in the Norwegian society and allows particular measuring methods and tools.²⁸ Weights, electricity metres, water-supply metres, and taxi metres, are a few examples of what is covered by legal requirements. There are, however, no specific, general requirements for the use of several types of equipment that may be hot candidates for direct delivery of data to individual decision-making processes (e.g. GPS, biosensors, scanners). Legislation exists within a few specialised areas containing provisions with requirements regarding measures and measuring tools. Provisions concerning tachographs in professional vehicles, measuring of time, speed and distance, and speed cameras are two of several examples of such special regulation. These legal requirements would have direct effects if data stemming from the regulated measuring devices were used as the basis for individual decisions in public administration. For instance, whether or not the speed of vehicles is measured as exceeding speed limits depends on the safety margin programmed into the speed measuring device.

This is not the place to enter into discussions of overall legislative architecture and where to eventually systematically place provisions regarding requirements pertaining to digital equipment producing decision data. Both rule of law and data protection considerations, however, are arguments in favour of some type of regulation covering decision data in government administration. A basic requirement could for instance be that such equipment should always be in accordance with technical standards from national or European standardisation organisations.

A radical way of regulating decision data from sensors etc. would be to establish that only data derived from particular digital equipment that meets specific technical requirements would be accepted as decision data. The effect, for instance, could be that manually assessed, subjective data regarding the same circumstances is not relevant as a basis for individual decisions. Of course, this would not be acceptable in a criminal case, but in administrative decision procedures with a limited effect on individuals, such an approach is possible to imagine. A parallel can be found in the Norwegian Housing Benefit Act, where the income of people registered at an address in the Population Register is included in the household income, even though the people who have earned the money have not actually lived at that address.²⁹ *Registered* facts, in other words, take priority before real facts. Provisions that give priority to facts as they have been registered by certain digital equipment according to a certain method, before other observations and alternative methods of observation, could by comparison be seen as less controversial. A possible motive for such a type of provision, is that it eliminates a great proportion of potential disagreement

27 For Norway, see the Act of 26 January 2007 no. 4 on measurement units, measurements and standard time.

28 Cf. Measures and Measuring Act, § 2.

29 People who actually live at the address without being registered at the address in the Population Register are included too.

concerning correctness and adequacy of the digitalised collected decision data. Legal regulation establishing that data stemming from standardised and certified equipment is the sole source of decision data would imply that these data will be apprehended as close to objective. To contest the correctness of such data, the private party will have to prove malfunctions of the digital equipment, which is likely to happen very infrequently.

A third group of possible regulative measures to support digital collection of decision data is to pass rules that promote trust in the processing, in the relation between citizens and businesses on the one hand, and government authorities on the other. Requirements for documentation of how measuring equipment and logging routines have been operated, access to logged data and scrutiny, as well as supervision and control of equipment and the use of them, are examples of types of provisions that could contribute to building necessary trust.

8 Modernising Public Administration Legislation

Our legal system is largely built on the active participation in legal processes of private parties; for instance, those individuals that are directly affected by legal decisions made by government agencies. An important aspect of this role as party to individual cases is associated with assessment of facts relevant to the case: Parties should be able to make their own assessments of the facts in their cases, how facts should or could be considered from a legal perspective. They must also have the opportunity to contradict government agencies' assessment of facts and legal considerations. Digitalisation of decision-making in individual cases is changing this role, but how it is changing and how it will change in the future, is not self-evident.

Self-service case processing is the model closest to the ideal of active individual parties. Here, the degree of contribution from parties relies on how the system is designed, and if the defined leeway for the contributing party is made broad or narrow.³⁰ Thus, the parties' possibilities for influencing decisions could range from considerable to very modest. Other digitalisation strategies make parties more or less redundant. Decisions may be made, for instance, only on the basis of assessments made by others rather than by the individual party and with previous assessments and decisions as sources. Employers' decision to grant a certain salary; police decisions to issue speeding fines, and registration of transfers of real estate in the land register are examples of how reliable information in digital form may be accessible as a basis for decisions in new individual cases of government administration – without having to involve the party.

30 With a «narrow» design, for instance, individual parties may only accept or reject pre-stored information in his or her case, while within a broad design, individual parties may register information in the system based on their own observations and assessments. In any case, information must be represented within the formal framework defined in the system, regarding the accepted types of information, accepted values, defined mandatory pieces of information etc.

An alternative set of strategies is to base decisions on automatically collected information, i.e. information based on “electronic traces” in the form of different types of logs registering various types of deeds and events connected to individuals in question (payments, information searches, travels etc.). This differs from the previous strategies mentioned in that it is based on sources independent of direct cognitive assessments. Here, both the individual party himself or herself (cf. first strategy), and government agencies, businesses etc. (cf. second strategy), are set on the sideline with no direct influence over the information on which decisions are based.

In current digital government administration, we will find an application of different mixes of strategies, often in combination with traditional case processing where an officer is partly in control and interacts with the data system. Even though self-service case processing could be an option within some types of decisions, it is not likely, in my view, that this will be preferred within areas of administrative law with a certain complexity: It will be too difficult to attain an acceptable low error rate because too many people will lack the necessary motivation and knowledge etc. It is much more likely that the second category of strategies mentioned above will be increasingly important: If so, decisions will be based on information already established and stored in computer systems of other government agencies and private businesses, stemming from human assessments. This will allow reuse of information and use of data systems developed to serve many purposes, and will thus be more effective than using information from scattered non-standardised sources. The third category of strategies mentioned will probably also be important within selected sources, in which information is automatically generated from technical devices, analysed, and processed to produce input to individual decisions. Here, technological developments, in particular regarding sensor technology etc., in combination with advanced analysis methods (machine learning etc.) where data from many sources could be combined,³¹ point in the direction of a much more important role of technologies “reading reality” without taking the roundabout way via the minds of people.

In other words, there are vast potentials and possibilities in developing government administration in the direction of a high degree of automated collection of information as the basis for decision-making in individual cases. However, this creates challenges for legal ideals and principles, and for the rights, autonomy and protection of individuals. A minimum requirement should be that parties should still be able to question the legality of decisions by contradicting their factual basis. In Norway, this would require a much higher degree of clarity and transparency regarding how decision data are established and processed, and regarding legal requirement of establishment and processing of such data. Here, I will not enter into detailed discussion about required legal measures, but only mention a few elements related to legislation.

First and foremost, we should consider passing laws that clearly establish the legal basis for collection and reuse of decision data from machine-readable sources, cf. second strategy presented above. Such regulations should not be

³¹ Including non-personal data.

fixed to particular sources, but instead establish requirements regarding data definition, data quality, integrity and availability that would narrow the choice of possible sources. Not least, the existence of publicly available information describing in detail how processing of data is carried out should be established by law,³² thereby making available information that could reveal possible weak points and sources of errors that private parties may scrutinise and deal with. Considering the likely emergence of more extensive use of “reason-less” technology, i.e. technology operating independently from cognitive human processes (cf. medical sensors, GPS etc.), there is also a need for legislation addressing technical requirements for devices applied to collect and further process data for decision-making purposes. A possible point of view could be that only equipment certified pursuant to certain technical standards should be allowed as a source of decision data in government administration. Such a regulation should also impose preparation of technical documentation of applied equipment; again, preparing the ground for scrutiny and lodging of complaints from private parties.

Because the technology connected to both strategies mentioned of automated collection of decision data implies rather technologically advanced solutions that – for most people – are beyond their intellectual capacity, legal regulation defining technological requirements and documentation would in itself be insufficient. Thus, in addition, there will probably be a need for a government body to advise and provide services to citizens who wish to scrutinise problems and understand the technology and its possible effects in particular cases.

9 Concluding Remarks

Fifty years ago when Swedish Law and Informatics Research Institute (IRI) was established, digitalisation of public administration and automation of decision-making in mass-administrative systems was in its infancy. Facts of individual cases were usually entered into by case officers on the basis of manual assessment of relevant facts and relevant legal sources. Only processing of these manually registered facts was automated, not the initial collection of them. Five decades later, important parts of government administration, at least in the Nordic and other technologically advanced societies, are approaching a situation where several types of decision-making are fully automated, in the sense that decision-making is based solely on machine-readable sources and automatic processing of collected data. Moreover, development and application of data from GPS, biometrics, various types of sensor technology, and other technologies producing “reading the reality”, point in the direction of a government administration that will increasingly be reliant on “reason-less” sources, where technical equipment is employed to capture live processes as they occur and, on that basis, generate decision data.

When looking 50 years back in time, stating the marked difference between current processing of data describing facts for decision-making purposes with

³² Typically, in delegated legislation.

the situation towards the end of the 1960s, it seems probable that current technological possibilities will lead to further changes in the direction of a higher degree of automated decision-making. I have argued that such developments should be regulated by new legislation as a response. It should particularly be directed towards the level of *data systems*, regulating systems directly, rather than the traditional way of regulating decision-processes with a case-by-case perspective. Legal requirements for automatic systems developed to collect decision data should be key in such amended statutory responses to developments of digital government administration.