Retrieving the Sources of Legal Decision-Making – Technical Possibilities and Related Legal Issues

Erik Helling

1 Introduction

The present article discusses information retrieval pertaining to the "sources of legal decision-making". This latter term not only includes formal legal *norms* (laws and regulations) but also other forms of *legal sources* (e.g. precedents) as well as legally relevant *facts*.

Retrieval of aforementioned sources can raise various technical and legal issues. Many of these issues are discussed below, with particular emphasis being placed on matters of information security, intellectual property rights and personal data protection.

It should be observed that the following text makes fairly extensive use of stipulative definitions. These definitions are not utilised under the pretence that they represent the only correct way to define particular terms and concepts. Rather, their function is simply enabling terminological consistency for the purposes of this article.

1.1 Aims

Specifically, the present article has the following aims:

Defining legal decision-making and its sources (chapter 2)

Defining information retrieval and information retrieval systems (chapter 3)

Analysing technical and legal issues related to retrieval of "relevant sources" – i.e. the sources of legal decision-making (chapter 4).

1.2 Delimitations

This article deals only with retrieval of sources that are in some way related to the process of legal decision-making (as defined in 2.1). This does not mean that only retrieval *by* decision-making bodies is considered. The delimitation concerns the *nature* of the sources discussed, rather than the specific purpose for which these sources are de facto retrieved. For this reason, the article also touches upon issues pertaining to source retrieval performed by private individuals (see e.g. certain sections of 4.1.3 and 4.2.2).

One should also note that the article focuses on *retrieval* of information, rather than the subsequent *processing* of this information.[1] In other words, the utilisation of relevant sources for the purpose of reaching legal decisions is not discussed in depth.

It is furthermore important to observe that we are mainly dealing with *IT-supported* information retrieval, as opposed to information retrieval performed

1 As will be seen below, this narrow definition of "processing", as a stage taking place after initial information retrieval, does not correspond to the wide definition of the term that is utilised in EC legislation (compare 3.1 with 4.2.2).

in a manual fashion.[2] Nonetheless, it is necessary to analyse information retrieval in general before it becomes possible to focus on the particularities of related IT-support. Thus, some discussions concerning information retrieval are technologically neutral (see 3.1).

Another important delimitation of the present article concerns its focus on European legislation. Specifically, EU (European Union) directives constitute a great part of the substantive law discussed. References to American legislation have not been included, partly due to their limited relevance to the topics discussed, partly due to considerations regarding article scope. Thus, the article takes its starting point in a norm-bound, civil law tradition, rather than in a case law setting.

2 Legal Decision-Making and its Sources

2.1 Defining Legal Decision-Making

The present article focuses on "legal decision-making". This brings to question what decision-making is to be considered legal. Here, we will apply a rather wide definition, linking the categorisation of certain decision-making tasks as "legal" to whether they have a basis in the *application of legal norms*.[3] As this definition takes its starting point in the concept of "legal norms", it is obviously necessary to specify the meaning of this latter term. Furthermore, it is necessary to analyse other forms of sources that can be relevant to the process of legal decision-making (by supplementing legal norms). An analysis of the various sources of legal decision-making will thus be carried out later in this text (2.2).

It should be noted that a decision taken with a starting point in a legal norm, but not entailing an *application* of this norm, does not fall under our definition of legal decision-making. Consider for instance a lawyer, making a decision regarding suitable tactics for a pending trial. Though his decision is likely to be based on legal norms to a greater or lesser extent, it nonetheless falls outside our definition. Similarly, all manners of private decisions taken with a starting point in legal norms (e.g. a person building a fence, as the relevant legal norms allow for such a procedure) are obviously not to be seen as legal. A comparable situation can e.g. be seen in the case of a company planning its strategies according to legal requirements. Even though relevant legal norms are taken into consideration in all these scenarios, the norms are not de facto *applied* in a manner *directly* resulting in legal repercussions of any sort.

---

2 IT stands for "information technology" and refers to the utilisation of computer systems for the purpose of fulfilling various tasks. The expression "computer systems" is utilised to denote any manner of combination between hardware (i.e. physical components) and software (*e.g.* applications).

3 Once again, it must be observed that present text has its main starting point in civil law. Thus, legal decisions are always seen as results of particular legal norms.

Instead, we must focus on those actors who are in fact able to apply legal norms in a manner entailing legal consequences for natural and legal persons.[4] As a result, the present article deals primarily with *public* decisions. Common examples of such decisions include public authorities approving/rejecting certain applications, as well as judges ruling in court cases. It can be said that decision-making of this kind constitutes "exercise of public power".[5]

2.2 The Sources of Legal Decision-Making

As a consequence of our definitions above, it becomes particularly important to define the various sources of legal decision-making. One should once again observe that we are not merely dealing with *legal sources* when discussing the decision-making procedure. The discussion also includes the *facts* that are to be combined with legal sources for the purpose of reaching legal decisions. Furthermore, this section concludes with a short discussion concerning certain sources that, while being relevant for the purpose of legal decision-making, are not generally thought of as "legal sources" in the traditional sense.

2.2.1 Legal Sources

The term "*legal sources*" is used to denote types of information that are generally associated with the legal field. Obvious, the most central of these sources, (at least in a civil law tradition) are the *legal norms*.[6] In defining legal norms, it first of all becomes necessary to explain the concept of "norm". Norms generally consist of an antecedent and a consequent. In other words, they are often designed in the manner of: "if X then Y". However, norms may at times be formulated in such a way, as to lead to a separation of consequent and antecedent. Specifically, some legal norms may stipulate an isolated prohibition in a certain act (e.g. "action X is illegal according to this act") which is sanctioned in another part of the act (e.g. "all violations of this act are punishable by Y"). This is an example of "law fragmentation" – i.e. the dispersion of legal norms. Norms may also consist of simple definitions – "term X is defined as Y" – even though no apparent cause and effect exists in these situations.

---

4 The term "natural person" refers to any physical individual that the law provides with legal capacity – specifically the capability to be subject to rights and obligations according to law. "Legal person" refers to an entity that, while not being a physical person, nonetheless possesses its own legal capacity (*e.g.* a corporation).

5 "Exercise of public power" can be roughly equated to the Swedish concept of "myndighetsutövning".

6 The term "legal norms" is preferred to "legal rules". This is due to a desire to avoid overlaps with the use of the term "rules" in other fields (*e.g.* that of logic). *Cf.* however Hart, ch. 5 and more recently, in Scandinavian doctrine, Wahlgren (1992), p. 146.

Continuing our definitions, norms are to be seen as *legal* only if their conception is linked to constitutionally valid legislative procedures.[7] This definition includes "law" in the traditional sense of the word (e.g. legal acts). Additionally, certain norm-making power can generally be transferred from the parliament to the government and to public authorities. These latter actors can then produce *regulations* (through procedures that we can refer to as *regulatory*).[8] Such regulations are also considered legal norms for our purposes, as they are ultimately based on measures outlined in constitutional law. Furthermore, the European Union also creates much material that member states must view as binding legal norms.

There are, of course, legal sources that cannot be considered legal norms (due to not meeting the criteria discussed above). These are referred to as *peripheral legal sources*. A typical characteristic of peripheral legal sources is that they are generally meant to illustrate, or regulate, some form of legal norm application. They can be of various kinds, e.g. precedents and preparatory works. Unlike legal norms, peripheral legal sources do not always base their validity on a particular form of conception. It is true that certain peripheral legal sources do derive their acceptance from their basis in particular political/judicial procedures (e.g. in the case of precedents and preparatory works). However, other sources, such as doctrine, are considerably more difficult to define as either "legal" or "non-legal". For instance, one could wonder exactly how "legal" certain literature must be, in order to be considered part of the legal doctrine. These reflections can also be extended to considerations concerning the required background and education of genuinely "legal" authors. Here it becomes obvious that the category of "peripheral legal sources" is not easily delimited. Nonetheless, as a rather rough definition of the term serves our purposes, there is little need to focus on these difficulties in the present context.

Certain elements of peripheral legal sources can be referred to as *quasi-legal norms*. These are characterised by exhibiting a *norm construction* while not qualifying as legal norms (i.e. not being the result of legislative/regulatory procedures). Quasi-legal norms can be found in many incarnations, from legal doctrine and preparatory works to unwritten legal principles.[9]

---

7 *See* H.L.A. Hart's discussion concerning *rules of recognition* (Hart, p. 97-107). Such rules of recognition can be characterised as a form of *secondary* legal norms. Secondary legal norms regulate the manner in which *primary* (substantive) legal norms are to be created and applied.

8 It should be noted that the capability to produce regulations can also follow directly from legislation, rather than from a specific instance of delegation.

9 It is far from uncontroversial whether principles should, in fact, be seen as part of the legal system (i.e. what is normally considered "the law"). Ronald Dworkin argues that this is the case. Specifically, his claim refers to so-called "hard cases" which do not allow for straightforward resolution (*see* Dworkin, p. 31-39 and p. 81-131). According to Dworkin's view, there are legal principles (as well as *policies*, which will not be elaborated upon here) that must be considered part of what is commonly referred to as "law". He supports this opinion with the argument that it would be considered incorrect by judges not to apply these principles in the process of certain legal decision-making. For this reason, he holds that the principles must be considered legal, rather than extra-legal. In other words, Dworkin does not perceive any true gaps in the legal system. Any apparent gaps are instead seen as failure of decision-makers to truly observe all existing legal sources (including principles). *See* Dworkin, p. 105-130 about the fictional superhuman judge Hercules. H.L.A. Hart instead

In many instances, quasi-legal norms are created by decision-making authorities themselves and subsequently utilised as internal work material. Specifically, the purpose of such quasi-legal norms is assisting the interpretation of legal norms (particularly when the latter contain vague formulations). They are often derived from one-time decisions at authorities concerning particularly complex legal issues and subsequently utilised as "moulds" when dealing with future issues of a similar character. This form of guidance is generally expected to result in benefits with regards to efficiency and consistency in legal decision-making. Here, reference can be made to the Swedish Tax Agency, which utilises a form of quasi-legal norms (termed "steering norms") in order to achieve said benefits

It should be observed that quasi-legal norms may often be of a sensitive nature. As an example, one can consider quasi-legal norms regulating the exact limits for certain tax benefits. Normally, such limits are not formulated in legislation, as it would otherwise be a simple matter for individuals to engage in tax avoidance schemes. Thus, it is vital that any supplementing quasi-legal norms are only known to specific decision-makers, who are subject to special legal control for this exact reason (see 4.1.4 for further discussion concerning the need to protect quasi-legal norms).

One should be certain to differentiate *quasi-legal* norms created by an authority from those *legal* norms that an authority can legitimately produce (e.g. through delegation of norm-making power from the parliament to lower levels of the state hierarchy).

2.2.2 Facts

Legal decision-making furthermore requires access to *facts* derived from empirical observation. In the context of legal decision-making, facts are generally retrieved for the purpose of being related to legal requirements of some sort. Commonly, relevant facts constitute *personal data*. For this reason, legislation pertaining to data protection can often place restrictions on "fact retrieval", preventing it from becoming uninhibited (see 4.2.2).

One should be observe that facts include the actual *cases* to be decided in a process of legal decision-making. If such cases are expressed in natural language, lacking legal classification, they can be referred to as *natural language cases*. With "natural language" is meant a form of casual, every-day language, which is distinct from the formulations of legal norms. For instance, it may be apparent that certain illegal tax avoidance schemes have de facto been performed (e.g. through observation of underreported income or overstated tax deductions). However, it is another matter to locate the proper legal norm required to classify this behaviour in a legally relevant way. This is particularly

---

sees the legal system as an imperfect construction of legal norms. In instances where legal norms do not offer sufficient guidance for the purpose of legal decision-making, Hart argues that judges will inevitably "fill out" the gaps through essentially non-legal elements, including moral opinions and, perhaps most significantly, individual views on the use of language (interpretation of vague terms). He refers to this as the "open texture" of law (*see* Hart, p. 120-132 and *Cf.* 2.2.3 below).

true when the legal classification entails detailed considerations concerning degrees of severity (e.g. minor/major fraud).

Certain cases are legally classified "by default" and do not necessitate any further efforts in this direction. Important examples include cases that are initiated through formalised procedures (e.g. requests for financing). In many such instances, the legal classification – and thus, the applicable legal norm – is not in question. However, the *interpretation* of said norm may well be the subject of disagreement.

The line between factual circumstances and individual theories can often be vague – particularly since our impression of reality is inevitably based on our perceptions. However, it should be noted that legal cases generally include a large number of details that are not truly under dispute – they are simply observed and acknowledged. In such instances, when there is no reason to question the validity of certain information, it appears sufficient to accept said information as "fact" without involving any further theorising. This is often true regarding various forms of personal data (e.g. names, ages) that are required for the purpose of reaching legal decisions.

On the other hand, it is obvious that observations which involve some form of dispute regarding their interpretation (e.g. evidence issues) cannot simply be taken for either true or false facts. Consequently, the category of "facts" – viewed in the context of legal decision-making – is only relevant as a *lowest common denominator* of observed information. In other words, it merely encompasses information that finds itself on a level where there is common agreement between the different sides of a case. As soon as information is disputed in any way, it is no longer possible to talk about obvious "facts". In these latter situations, empirical observation must give way to legal analysis. Only through such an analysis is it possible to reach a conclusion – a "legal true or false" – pertaining to certain disputed information.

2.2.3 Other Relevant Sources

One can envision further elements that are relevant to the process of legal decision-making. For instance, one could argue that *interdisciplinary skills* should be seen as a form of relevant sources in this regard. Specifically, there is often a need to rely on e.g. mathematical skills and technological skills in order to achieve correct decisions in specific legal areas (such as those of taxation law and IT-law). This is especially true in our age of field convergence.[10] Furthermore, it can be claimed that *tacit sources* – information internalised in the minds of humans – constitute a further possible category of relevant sources.[11] As examples of tacit sources can be mentioned general aspects of human reasoning such as *common sense* as well as field-specific elements (e.g.

10 For further discussion concerning the interdisciplinary character of modern-day legal decision-making, *see* Helling (2004), passim.

11 Tacit sources have been described by Michael Polanyi as a form of "instrumental knowledge", which he contrasts against matters that are known "explicitly, as objects". *See* Polanyi, p. 88.

the *legal methodology* of lawyers).[12] Due to considerations concerning scope, these forms of sources will not be analysed extensively in the present article (thus, they are not included in the discussions in chapter 4).

3 Information Retrieval

3.1 Defining Information Retrieval

Initially, one should observe that *information retrieval* refers to all possible ways of collecting information. Applying this definition, one could even see the *recollection* of information in a person's mind as a form of "internal" information retrieval. This form of retrieval from human memory is of course especially relevant with regards to abovementioned tacit sources. Since such sources exist completely within the human mind (rather than in a clear external manifestation), it would seem that only "internal retrieval" from human memory can suffice for the purpose of accessing them.[13]

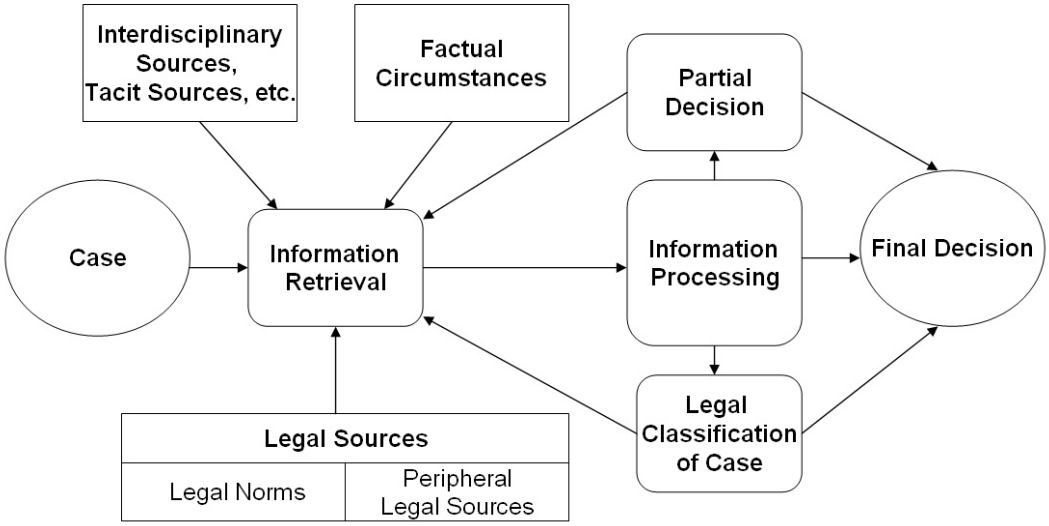
When discussing information retrieval for the purpose of legal decision-making, one can differentiate between retrieval of a *passive* and of an *active* nature. This division utilises the perspective of the decision-maker as a starting point. If information retrieval is performed on the initiative of a decision-maker, e.g. for the purpose of comparing factual circumstances with certain legal criteria, it can be termed *active*. On the other hand, if facts are received as the result of some outside initiative, the retrieval is considered *passive*. Examples of this latter situation include applications from individuals to authorities. In these instances, individuals generally submit information to the decision-maker by their own free will, as they have an interest in certain decision-making procedures being carried out. Of course, facilities for the purpose of information submission may previously have been implemented (e.g. in the shape of electronic forms on authority websites – see 4.2.1). However, this does not alter the fact that, in these instances, the decisive choice to de facto submit certain information to authorities is made by individuals

It is also important to differentiate between the phases of *retrieval* on the one hand and *processing* on the other – particularly as only the former phase is the focus of the present article. The term "processing" refers to the actual *use* of information in different ways, e.g. for the purpose of enabling legal decision-making. One should not visualise retrieval and subsequent processing of information as a one-way street. Rather, it is necessary to view both these phases as co-existing in a symbiotic relationship, where processing of information can yield possibilities of renewed and improved retrieval. Thus, we are dealing with

12 *Cf.* Helling (2003) p. 20-23 where I discuss general and specialised meta-knowledge – terms closely associated with the discussions here concerning general and field-specific tacit sources. Legal methodology is a term often utilised to describe the "intuitive" skills of the legal profession with regards to the utilisation of legal sources and the drawing of legal conclusions.

13 Of course, this does not contradict the notion that techniques in the field of artificial intelligence will eventually be able to approximate the *effects* of tacit norms and incorporate these effects into systems of automated legal decision-making. *See e.g.* infra note 22.

a form of circular movement, where stages of information retrieval and information processing may need to be reiterated many times, before a final decision (and/or a legal classification of a case) is achieved. This is illustrated by the figure below.



3.2 Information Retrieval Systems

Information retrieval systems (IR-systems) are able to enhance the speed and efficiency of legal decision-making by providing easy access to relevant sources. As a rule, such systems are connected to some form of *database*, from which they retrieve information. This retrieval is normally carried out by way of a *search engine* that accepts *queries* (inputted keywords/phrases) of greater or lesser complexity. The search engine matches the queries with information stored in the database for the purpose of presenting users with the desired results. One should observe that the sophistication of search engines varies greatly.

Many search engines allow for the utilisation of so-called *Boolean operators* (AND, OR or NOT) for the purpose of refining information retrieval. While Boolean search engines can offer valuable search possibilities, they suffer from a rigid binary construction that can lead to unpredictable results (seeming either too strict or insufficiently strict).[14] Furthermore, the results of Boolean queries are seldom ordered in a useful manner.

These disadvantages are not generally shared by so-called *probabilistic* search engines.[15] Such engines are able to rank results according to estimated

14 Karlgren, p. 33-34.

15 It should be observed that Boolean and probabilistic methods are possible to combine. *See e.g.* Cheshire II at cheshire.lib.berkeley.edu.

relevance to users, as well as handle queries that are not based on strict formulations. Thus, it can be said that probabilistic methods offer means by which to approximate natural language. As will be discussed below, certain search engines allow for formulations of queries in the form of more or less free questions.

The functionality of search engines can be enhanced through many different features. For instance, *hypertext link*s can be used to associate bits of information with each other, even overlapping different databases. Thus, they can play an important role in structuring masses of text and enabling easy reference to different text sections.[16] Another useful feature concerns possibilities of *truncation*.[17] Additionally, one should not neglect the usefulness of *field-specific* searches, e.g. allowing queries pertaining to particular legal areas and/or sources. In the context of field-specific searches, considerable attention should be given to developments in the field of information standards (e.g. XML)[18]. These standards provide possibilities of defining parts of documents in a manner allowing for simple subsequent retrieval (and processing) of textual fragments.[19]

With regards to modern development, one should observe advances in the field of *conceptual* search engines. As their name implies, these search engines attempt to identify *concepts* to which queries belong, in order to delimit the type of information that is retrieved. Thus, irrelevant results can presumably be singled out more easily. The obvious difficulty relating to these conceptual engines lies in the need for accurate *concept identification*. Some systems simply let users describe the concepts to which queries belong.[20] It is also possible to utilise some form of *classifier* for the purpose of automatically attributing concepts to queries. However, such automatic classification can obviously be unreliable. Another possibility, often seen in recent developments, concerns the utilisation of *user profiles* with the aim of identifying relevant concepts pertaining to the interests of particular users. This could be characterised as a form of "personalised search".[21]

---

16 Hypertext links are an important component of the so-called "World Wide Web" (WWW). The WWW was originally designed by Tim Berners-Lee, at the European Particle Physics Laboratory (CERN) around 1990. It offers a common standard by which to share information – specifically through the use of the Hypertext Transfer Protocol (HTML). Thus, it has played an important role in the creation of internet, as we know it today. *See* also the World Wide Web Consortium at "www.w3.org".

17 "Truncation" refers to possibilities of constructing queries that use partial words (*e.g.* law- and -book). Many search engines furthermore allow the replacing of individual letters with "wildcards", if these letters are uncertain. For more information about truncation and its implications for the legal field, *see* Magnusson, passim.

18 XML stands for eXtensible Markup Language.

19 This article will not discuss details surrounding information standards. For such a detailed discussion, *see e.g.* Magnusson Sjöberg, passim.

20 *See e.g.* Gauch et al., passim.

21 *See* Trajkova & Gauch, passim.

Certain IR-systems can be characterised as "semi-intelligent".[22] For instance, some search engines allow queries formulated in natural language. This may often give the impression that a system "understands" a question, and provides answers based on that specific question.[23] However, these functions are seldom genuinely related to artificial intelligence techniques. Rather, they generally rely on locating keywords in natural language sentences and correlating them with pre-constructed indexes.

Somewhat closer to the area of artificial intelligence, we can find a particular form of legal IR-systems that take the retrieval process of a step further. Rather than merely presenting legal norms to users, they break these norms down into user questions that are possible to answer in a binary "yes/no"-fashion. The process is reiterated until the systems have retrieved all information necessary to correlate the facts of a particular case with relevant legal norms.[24] Thus, the user is gradually guided towards the legal norms that he appears to require. To exemplify this functionality, reference can be made to the so-called "Statute Expert", developed by the Australian company SoftLaw.[25]

4 Issues Concerning Retrieval of Relevant Sources

4.1 Concerning Retrieval of Legal Sources

This section discusses retrieval of information that falls under the category of *legal sources*. As seen above (2.2.1) our definition of legal sources not only encompasses legal norms but also other forms of supplementing legal material, such as precedents, preparatory works and doctrine.

4.1.1 Services Providing Legal Sources

Legal sources are commonly expressed in some form of official collection. Thus, decision-makers often rely on external online databases for the purpose of locating and collecting those legal sources that are required in decision-making processes.[26]

---

22 In fact, the term "knowledge retrieval" is sometimes utilised when referring to systems that retrieve information in a manner clearly crossing the line into information processing. *See* for instance the "Cyc project", aiming to incorporate the "background knowledge" of humans into automated search engines, in order to enable "knowledge retrieval" ("home.ku.edu.tr/~dyuret/pub/cyc96"). The term "background knowledge" can be compared to what has previously been referred to as "tacit sources" (*see* 2.2.3).

23 *See e.g.* the "robot" of the Swedish Tax Agency, at "www.skatteverket.se/deklaration/04/erik".

24 *Cf.* Karlgren, p. 35 regarding relevance feedback. *See* also the figure in 3.1 above.

25 "www.softlaw.com.au." *See* also my discussion of this technology in Helling (2003), p. 49-53.

26 The usefulness of internet-based retrieval of legal information cannot be overemphasised – particularly in our modern age of hyper-regulation. *See e.g.* Susskind, p. 138-139. Thus, while the present article focuses on problematic aspects relating to online distribution of legal sources, it is important not to let the advantages go unnoticed.

To utilise a Swedish example of a legal source database, the site *Rättsnätet* provides free access to Swedish norms.[27] Access to precedents and preparatory works is also provided (albeit requiring a paid subscription). *PointLex* is another Swedish service providing legal sources online.[28] Of course, many further examples could be presented. While online databases are quickly gaining in popularity, much information retrieval pertaining to legal sources is still performed through CD-ROM services (such as InfoLex[29]).

One should note that many of the databases utilised for retrieval of legal sources are in fact developed and managed by private companies, often contracted by the state.[30] These databases are commonly utilised by state authorities in their decision-making functions. Thus, while the state creates legal sources in its judicial and legislative role, it may simultaneously rely on privately maintained databases for the purpose of retrieving these sources. In other words, the state can often be seen as both a producer and a consumer of legal material.[31] This situation is understandably more noticeable in political settings stressing decentralisation and privatisation. However, the commercialisation of legal information is highly likely to increase even in other kinds of jurisdictions – an evolutionary process related to information constantly becoming a more valuable asset.

---

27 "www.notisum.se". Peter Wahlgren's review of this service, albeit somewhat outdated, can be of interest in this context. *See* Wahlgren (1998), p. 254-262.

28 "www.pointlex.se".

29 InfoLex is developed and maintained by Thomson Fakta ("www.thomsonfakta.se").

30 Of course, there are also instances where the state maintains databases of legal sources. *See e.g.* Rixlex ("www.riksdagen.se/debatt/lagar_forordningar.asp"), offering access to Swedish legal norms. General problems plaguing many of these state-owned information services concern a low level of user-friendliness and insufficient utilisation of existing technical possibilities. In this context, it should be noted that certain important regulatory initiatives have been carried out in Sweden. These initiatives are intended to ensure the existence of a public information system providing state authorities, as well as individuals, with access to legal sources. *See* Rättsinformationsförordning (1999:175). The beginnings of such an information system have already materialised, utilising a "portal" ("www.lagrummet.se") for the purpose of uniting disparate legal source databases of a public nature. However, much remains to be done in order to fulfil the visions underlying the project, particularly with regards to user-friendliness and the establishing of logical links across databases and between documents/textual elements. Work on the public information system is still very much in progress, supervised by the Swedish Agency for Public Management. Future developments are intended to make extensive use of information standards (specifically XML) for the purpose of marking up legal source texts and thus enhancing search capabilities (certain alternative solutions are also considered). In this context, one should observe a recently created precedent database of the Swedish judiciary ("www.rattsinfo.dom.se"), which constitutes part of this public information system. *See* also the associated Foundation for Legal Information, created as early as 1989 for the purpose of dealing with issues pertaining to the digitisation and subsequent distribution of legal material ("www.rattsinfo.se").

31 *Cf.* Bruce, p. 14, where it is said that "[c]reation of law is so dispersed that government often becomes both a consumer and a redistributors of law […] the agencies or law-making bodies that turn their output over the third parties for publication, whether these parties are inside or outside government, have the same difficulties of access that the general public has." While this discussion concerns the situation in the United States, its relevance extends beyond that jurisdiction.

4.1.2 Reliance on Electronic Legal Sources and Corresponding Exposure to Liability

In view of what has been said above, it is important to observe possible difficulties relating to *reliance upon digital versions* of legal sources.[32] Companies providing legal source databases do not generally accept any responsibilities for errors in their digital texts. In fact, they commonly provide disclaimers emphasising the need to consult printed versions of legal sources in order to remove any doubts regarding possible inconsistencies. Of course, even in the absence of such a disclaimer, state authorities are responsible for ensuring that the sources they utilise in their decision-making processes are correct.

Assuming that an erroneous digital representation of a legal source is utilised in a process of legal decision-making, the state can incur liability for – in a sense – misunderstanding itself. Obviously, the "bouncing" of legal sources back and forth between public and private actors does not constitute an ideal situation from the perspective of legal information integrity. Choosing trustworthy information providers is naturally of great importance in this context. However, this alone may not be enough to ensure proper information security – specifically if possibilities of *data intrusion* are kept in mind.[33] One must concede that achieving complete information security in *any* manner of network setting (be it public or private) is a close to impossible task. This is particularly true for databases that are connected to external networks such as internet – a situation common for legal databases of today.

For the reasons discussed above, it is important to focus on possibilities that allow information to "verify itself" through technical means. In other words, if information can provide some sort of "stamp" as to its qualities, the exact path of this information (through more or less obscure intermediaries) becomes less of a concern. For one thing, transferred information must possess *integrity*, in the sense of not being corrupted on its travels between authorities, companies and individuals. Furthermore, it is generally essential to know the identity of the original producer of the information – information must exhibit *authenticity* and allow for *accountability*.

---

32 Nonetheless, it is possible to envision an interesting market niche for legal professionals as developers and managers of legal information systems. Through extensive involvement of legal professionals in the creation of legal information databases, high quality could more easily be ensured. Some researchers are optimistically predicting such development. *See e.g.* Susskind p. 98-100. There is, however, an apparent risk of form triumphing over substance. In other words, many legal source providers may be more focused on technical, and even artistic, innovation than on achieving correct presentation of legal material (with the usual exactness – some would say pedantry – of the legal field). In such a scenario, the role of legal professionals may easily become marginalised. This argument carries particular weight bearing in mind that modern *end users* of legal information can often be rather disassociated from the legal field (*see* infra note 56). Thus, they are likely to have other priorities with regards to the design of legal source databases than most legal professionals do.

33 The Convention on Cybercrime (ETS No. 185) of the Council of Europe [cit. Cybercrime Convention] is of particular interest in this context. *See* especially Articles 2-6. The legal regulations herein are generally of a reactive character, thus doing little to *prevent* data intrusion. However, it should be noted that Article 6, concerning the "misuse of devices", may play some part in ensuring that data intrusion is not in fact carried out.

Technical advances in the field of *electronic signatures* can be of great value for the purpose of fulfilling abovementioned goals. In this context, the so-called "Electronic Signature Directive" is of great relevance.[34] The primary purpose of this EU directive is establishing a digital alternative to handwritten signatures. Electronic signatures can be described as electronic data utilised to authenticate other electronic data.[35] If the signatures meet certain criteria (including the abovementioned information security goals) they are considered *advanced*.[36] Advanced electronic signatures can help offset many difficulties pertaining to electronic transfers of legal sources – difficulties that can be expected to increase as a result of increased reliance on online information, coupled with political initiatives pertaining to decentralisation and privatisation.[37]

4.1.3 Intellectual Property Issues Related to Legal Source Databases

Considering the manner in which legal sources are often distributed between authorities, companies and individuals, issues of intellectual property are obviously of great importance. Copyright protection is of particular interest in this context. Such protection does not require registration in most jurisdictions, but enters into force automatically as a result of the creation of a work.[38] Generally, works protected by copyright may not be replicated or distributed to the public without the consent of the copyright-holder (the creator of the work or one to whom copyright has been transferred). The copyright-holder can therefore be said to possess *exclusive* rights pertaining to various manipulations of the protected material.

As adherence to the rule of law requires that legal sources be freely available to all, one could argue that such sources should fall within the realm of public

34 Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community Framework for Electronic Signatures [cit. Electronic Signature Directive].

35 Electronic Signature Directive, Article 2(1). Specifically, electronic signatures consist of hash values that are encrypted with "keys" for the purpose of digitally signing specific information. The rather well known Public Key infrastructure (PKI) is characterised by the use of a private key (held only by a specific party) and a public key (freely available). In order to verify the possession of certain keys, one relies on certificates from so-called trusted third parties (TTPs). As much has been written about the technical details of digital signatures elsewhere, they will not be elaborated on further here. However, reference can be made to the Secure Legal Information Management Project (SLIM) – *see* "www.juridicum.su.se/slim".

36 Electronic Signature Directive, Article 2(2). The term *qualified* is sometimes utilised to refer to electronic signatures that reach a particularly high standard of information security. While no reference is made in the Electronic Signature Directive to "qualified electronic signatures", the term is *e.g.* encountered in Swedish legislation pertaining to the directive.

37 In relation to this discussion can be mentioned the new Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-use of Public Sector Information. This directive, to be implemented in member states before 2005, aims to increase the flow of official information across sectors and jurisdictional boundaries. Such initiatives could compound difficulties related to quality control of public information.

38 *See* the 1886/1979 Berne Convention for the Protection of Literary and Artistic Works [cit. Berne Convention], Article 5(2).

domain and thus not be subject to copyright.[39] Additionally, it may seem contradictory that the state, in its role of legal decision-maker, should face copyright restrictions pertaining to the utilisation of state-created material. It is true that various jurisdictions exempt "official material" (such as legal sources) from copyright. However, one should note a difference between copyright pertaining to legal sources themselves, and copyrights pertaining to the actual database in which such sources exist. Specifically, copyright can concern the manner in which material is selected, structured and made accessible (which can be referred to as the *design* of a database). Such copyright can e.g. materialise as a result of database development by private companies, even if these companies do not possess the copyright to the actual materials included in the database.

In the European Community, copyright pertaining to databases is regulated by the so-called "Database Directive" from 1996, as well as the "Infosoc Directive" from 2001.[40] The Database Directive defines "database" as "a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means".[41] In instances where the selection or arrangement of database contents can be seen as an author's "own intellectual creation", the database in question is protected by copyright.[42] In other words, a database must exhibit *originality* in order to be entitled to copyright protection. Legal databases conforming to the originality requirement could be believed somewhat rare, considering the rather form-bound manner in which legal material is traditionally ordered. However, it should be noted that technical features of databases (e.g. innovative search methods) may lead to the originality requirement being met, even in instances where the databases rely heavily on traditional legal divisions of material.

Furthermore, it should be observed that even non-original databases can be subject to a form of intellectual property protection known as *sui generis*. This protection requires that a database has been created as the result of *substantial investment* of either a quantitative or a qualitative nature (specifically, in the obtaining, verification or presentation of contents).[43] Thus, in a sense, the sui generis right concerns the contents of a database. It co-exists with the actual copyright to these contents (which may often be owned by someone else) and/or copyright pertaining to the database design. As a rule, neither *the whole* nor *substantial parts* of databases under sui generis protection may be extracted or re-utilised without the right-holder's consent. Thus, rather strong intellectual property rights can be derived simply through collection and re-distribution of

39 The Berne Convention does not offer any guidance in this respect, but leaves it up to member states to decide about the copyright status of official legal material. *See* the Berne Convention, Article 2(4). It should also be observed that the main argument in favour of copyright, ensuring incentives for further creation, does not carry weight when discussing legal sources.

40 Directive 1996/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases [cit. Database Directive] and Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society [cit. Infosoc Directive].

41 Database Directive, Article 1(2).

42 Database Directive, Article 3(1).

43 Database Directive, Article 7(1).

legal sources, even though no significantly original substance or functionality is added.[44]

According to what has been said above, most any utilisation of databases for the purpose of retrieving legal sources must be preceded by consent from any relevant intellectual property right-holders (e.g. as the result of a license situation).[45] If such consent does not exist and the databases are nonetheless utilised, sanctions generally come into effect.[46] However, one should note that abovementioned directives allow member states to stipulate exceptions to copyright and sui generis in certain situations. One of these situations concerns the performance of *judicial proceedings*.[47] Thus, there is a possibility to exempt information retrieval related to public decision-making from the effects of intellectual property rights. Information retrieval by private individuals may also be exempted, if this retrieval is performed for purposes of "private use" and for ends that are not commercial.[48]

Providers of databases generally utilise technical measures – e.g. in the form of password systems – in order to prevent unauthorised access to their services/works. Attempts to *circumvent* these measures are in fact seen as separate forms of copyright violations according to the Infosoc Directive.[49] Additionally, certain regulations in the directive concern *devices* that are primarily used for circumvention purposes or promoted for these purposes.

---

44 *See* also Reichman & Samuelson, p. 86.

45 In its normal sense, copyright regulates either copying or distribution to the public (as opposed to mere viewing). *See e.g.* the Infosoc Directive, Articles 2-4. However, it should be observed that any usage of electronic databases generally entails copying. This is due to the technical characteristics of digital information, which must be temporarily copied in some form (even if only to a system's internal memory) in order to be viewable. Such temporary copies are commonly referred to as *transient*. According to EC legislation, transient copying is excluded from copyright when it entails "lawful use" of a work and has no "independent economic significance". *See* the Infosoc Directive, Article 5(1b).

46 *See e.g.* the Infosoc Directive, Article 8(1).

47 Database Directive, Article 9(c), Infosoc Directive, Article 5(3e).

48 Infosoc Directive, Article 5(2b). This regulation requires "fair compensation" to be given to the copyright holders in the event of lawful reproduction for private use. It should furthermore be observed that the Database Directive, Article 8(1), contains a regulation excepting extraction and re-utilisation of *insubstantial* parts of databases from sui generis protection. This exception requires the user in question to be "lawful" and that the usage does not conflict with normal exploitation of the database (or unreasonably prejudices the legitimate interests of the maker of the database). In contrast, the extraction and re-utilisation of *substantial* parts of a database can only be performed in specific regulated instances. One of these, found in Article 9(a) of the Database Directive, concerns extraction of substantial parts of *non-electronic* databases – if this is performed by lawful users and for private purposes. No similar exception exists concerning extraction of substantial parts of *electronic* databases, which means that the sui generis protection in these instances is particularly strong. Considering the differences in legal effects between extractions pertaining to insubstantial/substantial parts of a database, it seems unsuitable that no clear definitions of these terms are to be found.

49 Infosoc Directive, Article 6. The regulation even includes sui generis protection, as seen in Article 6(3). One should note that only circumvention of "effective" protection measures is regulated. The rather uninformative definition of "effective" in Article 6(3) does not provide much guidance as to the intended meaning of this term, merely stating that protection is effective if it achieves the protection objective.

Member states are required to ensure through adequate legislative measures that such devices are not imported, distributed, sold, rented, advertised for sale/rental or possessed for commercial purposes.[50] The potential problems of this legal construction are apparent. Consider a situation where certain access is legal, but is nonetheless prevented by technical measures. Should then the circumvention (which is, in fact, required for access) be considered illegal, while the access itself is not? If so is the case, works can be effectively "locked up" and thus completely shielded from any legitimate access demands of a public or private character.

In order to somewhat remedy this conflict, the Infosoc Directive stipulates certain demands on member states concerning *some* of the exceptions to intellectual property rights discussed above – including the exception concerning *judicial proceedings.* Specifically, if a member state has implemented one or several of these exceptions, the state *must* consequently ensure that the exceptions can be exercised in spite of any technical measures that may be in place (provided the situation is not remedied through voluntary agreements).[51] These obligations upon member states do not extend to the exception concerning *private use*. Thus, there is no requirement to ensure that implemented exceptions concerning private use are reflected in the design of technical measures.[52] This situation must be characterised as a weakening of possibilities to exercise fair use rights in practice.

Furthermore, requirements to conform technical safeguards to existing intellectual property exceptions are not applicable for so-called *on-demand services*. These are described in the Infosoc Directive as "works or other subject-matter made available to the public on agreed contractual terms in such a way that members of the public may access them from a place and at a time individually chosen by them".[53] This definition can be taken to include all manners of online databases. The use of the expression "contractual terms" should not be given exaggerated importance, as it is generally a simple matter to incorporate so-called "click wrap" licenses in online services.[54] Thus, it appears that the directive allows online databases to be "locked up" in the manner described above, regardless of whether the access is allowed (or even required) by other legislation. Considering the possibility that online legal databases maintained by private companies may eventually be the only high-quality

50 Infosoc Directive, Article 6(2).

51 Infosoc Directive, Article 6(4). It should once again be emphasised that member states are required to ensure access to certain information *only* if this access corresponds to an exception concerning intellectual property rights *that the state has chosen to implement*. Thus, while the requirement to ensure access is mandatory, the underlying preconditions for such access (in the form of exceptions to exclusive intellectual property rights) have been made voluntary – a legal construction that can be described as somewhat peculiar.

52 Compare Infosoc Directive 6(4) with 5(2b). However, one should note that member states *may* take it upon themselves to ensure possibilities of private use copying. Thus, the intervention of states is voluntary in this case.

53 Infosoc Directive, Article 6(4), fourth indent.

54 The expression "click-wrap license" simply refers to an electronic contract that allows users to agree to terms by way of a mouse click. *See* further 4.2.1 below (final section) concerning legal issues related to electronic contracts.

providers of legal sources, this legal construction could lead to negative repercussions in many regards.

Thus, to summarise, intellectual property rights pertaining to legal databases can effectively "put a price" on legal sources, even in instances where the legal sources per se are not subject to copyright. Consequently, the modern environment of increasing dependence on private online databases may well result in a form of "information divide". Possibilities of accessing legal information can be expected to become increasingly dependent on the possession of financial/technical resources.[55] Private legal databases commonly offer many "bells and whistles" (e.g. overdone graphical interfaces) in order to further increase charged licensing fees while at the same time maintaining a competitive edge. Nonetheless, it must be said that private legal databases generally maintain a higher level of user-friendliness than do those provided by the state. [56]

Obviously, efficient access to legal sources is crucial both to state authorities and to private individuals for the purpose of establishing a clear "complete picture" of the legal system.[57] Thus, it can be argued that intellectual property rights to legal databases, and the resulting privatisation and commercialisation of legal sources, can ultimately hamper adherence to the rule of law. One apparent manner in which to solve this problem is for states to take on more active roles as distributors of their own official information.[58] While some states have taken

---

55 *See* also Reichman & Samuelson, p. 137-138.

56 In this context, it is important to consider the increased *value* of legal information in the modern age. Whereas legal information has traditionally been of interest mostly to lawyers, technical developments (*e.g.* pertaining to information retrieval methods – *see* 3.2) have increased the *accessibility* of such information. In other words, the user-friendliness of modern legal products has provided the general public with possibilities to locate legal sources easily (without requiring particularly extensive legal skills). Consequently, the general demand for legal material is greater than it used to be. *See* also Susskind p. 99 where it is claimed that electronic publishing has resulted in products that "demand less legal expertise on the part of the users and so appeal to a wider user base".

57 Here, reference can be made to the principles of "ignorantia iuris nocet" (ignorance of the law harms) and "ignorantia iuris neminem excusat" (ignorance of the law is no excuse). If individuals cannot be excused by claiming ignorance of the law, it appears reasonable that they should be provided with easy and unconditional (or at least close to unconditional) access to this law. In defining the term "law", there is no reason to focus merely on legal norms, as many peripheral legal sources (*e.g.* precedents and preparatory works) are essential for the purpose of understanding legal formulations. Of course, one could argue that legal sources require legal skills to be interpreted properly, and that free access to such sources would therefore be of little value to the common individual. Nonetheless, it appears important as a matter of principle that the sources of legal decision-making are easily accessible to all individuals who are bound by a particular legal system. Every individual then has to take it upon him-/herself to achieve the educational state required to understand these sources and thus be able to abide by them.

58 The argument can be made that it would be more expensive for states to develop their own advanced systems of legal source distribution, rather than simply relying on private parties for the retrieval of such information. Nonetheless, considering the ever-increasing cost of privately maintained information, as well as issues of information security (*see* 4.1.2) and goals pertaining to increased access to legal sources (*see* supra note 57), it would at least be advisable for states and their authorities to weigh options more carefully than is often done.

steps in this direction, the development must generally be characterised as rather sluggish.[59]

4.1.4 Issues Specifically Concerning Quasi-Legal Norms

As described above (2.2.1), the elements of legal sources which we have termed *quasi-legal norms* are often created by specific authorities themselves – either through explicit formulations or through generally accepted oral traditions. Quasi-legal norms can be of a sensitive nature, as they may reveal many aspects of the work methods of legal decision-makers. Outside individuals possessing this knowledge could more easily manipulate "the system" to their benefit (e.g. through engaging in tax avoidance schemes).[60] Thus, issues of information security – specifically the need for *confidentiality* concerning certain material – come into the foreground. For this reason, authorities may often rely on *intranets* (internal networks) for the purpose of sharing quasi-legal norms of a sensitive nature, ensuring that no computer in the intranet has an "opening" towards external networks (such as internet).

It should be noted that the use of intranets constitutes a defence only with regards to *external* data intrusion. There is still a risk that disloyal employers utilise sensitive quasi-legal norms for their own benefit. This is of particular concern if these employers are not the intended users of the quasi-legal norms in question (and thus are not subject to special controls or any other measures by which to balance the situation). For this reason, the utilisation of *access systems* is of great importance in authorities that deal with intranets carrying sensitive information. "Access systems" can be described as technical protections limiting access to certain information (and/or permitted operations on this information) to particular categories of professionals. Thus, through these safeguards, authorities can distribute quasi-legal norms in such a way as to make them accessible only to those decision-makers who truly require them in their professional roles.

4.2 Concerning Retrieval of Facts

Retrieving correct facts is obviously of great importance from the perspective of legal decision-making. As explained above (2.2.2), any observed circumstances, which are relevant to the process of legal decision-making, fall under the category of "facts". Here, we will primarily focus on retrieval of facts constituting *personal data*.

---

59 Nonetheless, as discussed above (supra note 30) certain states have introduced legislative measures of significance in this context.

60 Of course, certain insight into the workings of state authorities *should* be provided to the public for reasons of legal transparency. However, a balance must be struck between such "reasonable" insight on the one hand, and potentially damaging insight on the other – something that is obviously easier said than done but nonetheless necessary.

4.2.1 Practical Issues Surrounding Retrieval of Facts

As an introduction to this section, attention will be drawn to an apparent paradox concerning fact retrieval in the legal field.[61] On the one hand, retrieval of facts is seldom possible to accomplish satisfactorily before a legal classification of a case is achieved (i.e. when the relevant legal norm has been identified). On the other hand, it may be equally difficult to classify a case in a legally relevant manner unless sufficient facts have been retrieved. While decision-makers apparently handle this paradox in some fashion, it is not clear exactly how this is achieved. It may be that the application of tacit sources (see 2.2.3) plays some part in steering the minds of decision-makers through such hurdles. Nonetheless, a clear description of the process remains elusive. Possibly, it can be viewed in terms of alternations between retrieval of facts and attempts at legal classification, so that the path towards such a classification is gradually cut out (see 3.1 above – particularly the figure).

The retrieval (and subsequent storage) of various facts have traditionally been closely related to the concept of *filing systems*.[62] In order to define this latter term in the context of personal data, we can turn to the EU "Data Protection Directive" from 1995.[63] Here it is stated that filing systems constitute "any structured set of personal data which are accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis".[64] In other words, the definition of "filing system" requires information to be *structured* in some fashion that allows straightforward establishing of links between different bits of personal data (facts regarding an individual). As an example, one can consider a paper list that links names and personal numbers across different columns, so that one can be derived from the other.

When we, as individuals, wander through modern society, we generally leave personal data traces behind during our entire lifetimes (starting even at birth). In our age of e-commerce, even simple acts such as purchasing minor items in online stores can lead to the creation of elaborate personal data profiles.

In this context, it should be observed that the traditional notion of "filing system" has lost considerable importance over the last few decades. Instead, *electronic databases* have gradually gained footing as the primary method of information storage. Such databases do not rely on particular textual structuring for the purpose of establishing links between different bits of information. Search capabilities pertaining to electronically stored information are so considerable that even sentences in plain, unstructured text can serve purposes similar to those of traditional (paper-based/manual) filing systems.

Fact databases related to legal decision-making are seldom developed solely by decision-making authorities themselves. Rather, they often expand and evolve as a result of outside initiatives. Here, the abovementioned division into

---

61 This paradox has been observed in research. *See e.g.* Bing, p. 228.

62 The terms "register" or simply "file" are often utilised synonymously with "filing system".

63 Directive 1995/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal data and on the Free Movement of Such Data [cit. Data Protection Directive].

64 Data Protection Directive, Article 2(c).

active and passive information retrieval (as defined in 3.1) becomes relevant. The most apparent form of passive fact retrieval concerns situations where individuals initiate decision-making processes by their own initiative. This is commonly done through some form of *application* to state authorities (e.g. concerning financial grants). In these situations, electronic forms – common to several authorities – can enhance possibilities of achieving efficient, consistent and correct information retrieval in an online environment. Such forms may draw upon advances in the field of *information standards* (particularly XML, see 3.2) in order to ensure that received segments of information can be properly categorised, easily shared between authorities and effectively introduced into various electronic processes (e.g. pertaining to automated legal decision-making).

Advanced electronic signatures (previously discussed in 4.1.2) may also play an important part in enabling efficient sharing of information between individuals and authorities. Specifically, such signatures may be utilised for the purpose of providing individuals with so-called *e-identities*. The term "e-identity" refers to electronic measures that allow identification of individuals, thereby enabling the use of online services that necessitate secure identities. In some jurisdictions, the introduction of e-identities has increased possibilities of performing rather elaborate information exchanges online – including the filling out of tax forms and the registration of one-man companies.[65]

Obviously, legal regulations can be of great value for the purpose of promoting efficient sharing of facts between individuals and state authorities. This issue is closely related to possibilities of concluding contracts electronically.[66] The so-called "Electronic Commerce Directive" of the EU is of relevance in this context.[67] Said directive seeks to remove barriers to the conclusion of electronic contracts in EU member states.[68] However, one should observe that the directive allows member states to stipulate exceptions to these demands for an "electronic-friendly" contracting environment – specifically with regards to certain categories of contracts.[69] These exceptions appear to stem from a quite pessimistic outlook regarding technical possibilities of ensuring secure data transfers (e.g. through the utilisation of advanced electronic

65 This refers specifically to the Swedish situation, which allows e-identities to be utilised for said purposes. In Sweden, e-identities are provided by banking, telecom and postal services. For more information regarding the use of e-identities in Sweden, specifically in a public setting, reference can be made to the Swedish Agency for Public Management ("www.statskontoret.se"). This body is responsible for co-ordinating efforts pertaining to the modernisation of the Swedish public sector (*Cf.* supra note 30).

66 As will be seen below (4.2.2), the existence of a contract can constitute a valid ground for personal data retrieval/processing according to EU legislation. The same is true with regards to *consent* given by the person whom the data concerns. Possibilities of concluding contracts and/or submitting consent electronically can thus play an important role in ensuring efficient sharing of personal data.

67 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on Certain Legal Aspects of Information Society Services, in Particular Electronic Commerce, in the Internal Market [cit Electronic Commerce Directive].

68 Electronic Commerce Directive, Article 9(1).

69 Electronic Commerce Directive, Article 9(2).

signatures). While such technical means may not yet be fully developed, the exceptions nonetheless appear somewhat excessive – precluding the directive from truly paving way for an electronic "information market". This statement is particularly true with regards to our topic of legal decision-making, as one of the exceptions concerns "contracts requiring by law the involvement of courts, public authorities or professions exercising public authority". [70] Thus, it must be said that the directive does not play a significant role in promoting the efficiency of electronic transfers between individuals and state authorities.

4.2.2 Issues of Personal Data Protection

As seen above, many facts that can be crucial for the purpose of legal decision-making belong to the category of *personal data*.[71] Access to various forms of personal data is often required for the purpose of applying legal norms – this due to the fact that many legal formulations concern personal attributes to a greater or lesser degree. Thus, there is an obvious need on the part of public authorities to retrieve personal data. At the same time, digitisation of information coupled with advances pertaining to computer networks have significantly enhanced possibilities of collecting and manipulating such data in various ways.

This development is often portrayed as beneficial to society as a whole. As touched upon above (4.2.1), extensive sharing of personal data between authorities assists in diminishing the need for individuals to resubmit their information repeatedly to disparate facets of public administration. In other words, such sharing allows the state to exhibit a "united face" in its relations with the public. Furthermore, one could argue that increased flow and sharing of personal data promotes the quest for a cost-efficient public sector.[72] Nonetheless, these advantages must be balanced against potential dangers related to increased availability of personal data – particularly with regards to the privacy of individuals. In order to strike such a balance, certain legislative initiatives have been carried out.

An important piece of legislation regulating the use of personal data in a European perspective is the abovementioned "Data Protection Directive".[73] The directive defines personal data as "any information relating to an identified or identifiable natural person".[74] Obviously, this definition is extensive, encompassing all manners of facts relating to private individuals. The individual whom certain personal data concerns is seen as the "data subject" in relation to

---

70 Electronic Commerce Directive, Article 9(2b).

71 Of course, there exist legally relevant facts that do not constitute personal data. These may also require various legal considerations (*e.g.* pertaining to trade secrets). However, due to the limited scope of the present article, such issues are better dealt with elsewhere.

72 *See* also Blume, passim.

73 Directive 1995/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal data and on the Free Movement of Such Data [cit. Data Protection Directive].

74 Data Protection Directive, Article 2(a).

this data, while the entity exercising control over specific personal data processing is referred to as the "data controller".[75]

The main purpose of the Data Protection Directive is to regulate *processing* of personal data.[76] "Processing" is also defined broadly as including "any operation or set of operations which is performed upon personal data".[77] This definition differs from that utilised in the present article, where information retrieval and information processing are seen as two distinct – though closely related – phases of legal decision-making (see 3.1). As previously stated, the present work is primarily concerned with aspects of information retrieval.[78] Thus, the following discussions regarding the Data Protection Directive will expressly discuss "retrieval" of personal data, even though the actual scope of the directive is in fact wider. Nonetheless, some discussions will also touch upon *storage* of personal data, as this is an obvious result of the data being retrieved.

The Data Protection Directive only permits retrieval of personal data in certain stipulated instances. It lists certain *criteria*, at least one of which has to be met if personal data retrieval is to be legitimate. One of these criteria concerns *consent* given in an unambiguous manner.[79] Consent is defined as "any freely given specific and informed indication of [the data subject's] wishes by which the data subject signifies his agreement to personal data relating to him being processed".[80] This criterion is of particular importance in the context of cases initiated by private individuals. Formalised procedures (e.g. applications to public authorities) generally require consent to retrieval/processing of personal data. In practice, such consent can be given through some form of digital or printed condition, which is accepted by means of a signature, mouse click or a similar method.[81]

---

75 Data Protection Directive, Article 2(e).

76 One should observe that the Data Protection Directive only regulates data processing that is performed "wholly or partly" by automatic means, as well as manual processing of data which forms (or is intended to form) part of a "filing system" (*see* 4.2.1). This follows from Article 3(1).

77 Data Protection Directive, Article 2(b).

78 Due to the delimitations expressed above (and also presented in 1.2), regulations in the Data Protection Directive that only concern specific *uses* of personal data ("processing" according the more narrow definition in 3.1), rather than the preceding retrieval phase, are not discussed here. This includes Article 15 in the Data Protection Directive, regarding decisions-making performed in an automated fashion.

79 Data Protection Directive, Article 7(a).

80 This definition is found in the Data Protection Directive, Article 2(h). From the definition it is possible to conclude that consent cannot be implied (*e.g.* through passivity) but must concern a *specific* instance of personal data processing. By the same token, so-called "blanket consents", in the sense of a general acceptance of processing pertaining to a certain *type* of personal data, must generally be considered ineffective.

81 There has been some controversy among EU member states concerning whether mouse clicks should be seen as genuine "consent" (*e.g.* in Germany). However, if electronic acceptance of data processing is not permitted by law, online services have to be combined with paper-based acceptance forms, something that obviously counteracts the purpose of an electronic "information market". E-identities (*see* 4.2.1) can obviously play an important role in allowing for greater possibilities of submitting valid consent electronically. The abovementioned "Electronic Commerce" directive is also relevant in this context, as it

If consent does not exist, it is nonetheless possible to base personal data retrieval on the existence of a *contract* entered into by the data subject – provided that performance of the contract necessitates such retrieval.[82] As an example, employment contracts may require certain retrieval of personal data for their functionality (e.g. pertaining to the managing of salaries). Contractual relations can obviously exist between individuals and the state, thereby giving the latter a valid legal ground for retrieving certain personal data. However, state authorities may not utilise this data for unrelated purposes – at least not without renewed retrieval based on another legal ground. This follows from the so-called "purpose principle" relating to the quality of data handling (see below).

Personal data retrieval can also be motivated by the existence of a *legal obligation* with which a data controller is required to comply.[83] Legal obligations on state authorities can often concern the performance of decision-making functions. However, as will be seen below, the public sector – acting in its decision-making role – can generally base its personal data retrieval on a specific criterion pertaining to tasks of public interest.

The directive furthermore allows personal data retrieval if it is required to protect the *vital interests* of a data subject.[84] A literal interpretation of the term "vital" seems to imply a life and death situation, a reading that also conforms to the preamble of the directive.[85] Consequently, personal data retrieval can seldom be motivated through this legal ground.

As hinted at above, the directive also permits personal data retrieval that is required for the performance of a task carried out in the *public interest*, or which relates to the exercise of *official authority*.[86] Thus, if a state authority needs to retrieve/process certain personal data for the purpose of legal decision-making, it is consequently entitled to do so. Considering that many important state functions are dependent on efficient access to personal data, this criterion obviously plays a significant role in ensuring a functional public sector,

The final criterion concerns *legitimate interests* of a data controller.[87] If such interests exist, and are found to weigh more heavily than the interests of private subjects to retain their privacy, they constitute a basis for personal data retrieval. There is seldom a need for public decision-makers to rely on this regulation for the purpose of motivating personal data retrieval, as such retrieval can generally be performed according to other mentioned criteria. Rather, the "legitimate interests" regulation focuses mainly on personal data retrieval in the context of business relationships – specifically for the purpose of enabling "effective

---

removes obstacles to electronic conclusion of contacts. However, as discussed in 4.2.1, the scope of this directive is restricted by extensive exceptions.

82 Data Protection Directive, Article 7(b). The same is true if the retrieval is required in order to "take steps at the request of the data subject prior to entering into a contract".

83 Data Protection Directive, Article 7(c).

84 Data Protection Directive, Article 7(d).

85 Recital 31 in the Preamble to the Data Protection Directive refers to retrieval (processing) that is "carried out in order to protect an interest which is essential for the data subject's life".

86 Data Protection Directive, Article 7(e).

87 Data Protection Directive, Article 7(f).

competition".[88] Nonetheless, one should note that the balance between "legitimate interests" and personal integrity can affect decisions concerning whether public authorities should, upon request, present their stored personal data to the *public*. Even in countries with a strong tradition concerning free access to public records, data protection legislation is often given priority.[89]

Although one or several of the above criteria may be applicable, one must often consider additional regulations concerning *sensitive* personal data. Such personal data are given special protection, as they concern the private and intimate aspect of individuals' lives. Examples include data relating to political beliefs, health, sex life and ethnic origin.[90] As before, certain criteria (out of which at least one has to be satisfied) apply when deciding whether retrieval/processing of sensitive data is to be considered legitimate. The data subject's *consent* can once again play a role in overriding data protection safeguards.[91] However, when dealing with sensitive data, the consent must be *explicit*.[92] Another criterion concerns sensitive personal data that the data subject has made *public*.[93] Furthermore, retrieval can be permitted if it is required for the "establishment, exercise or defence of legal claims".[94] It should also be noted that retrieval of personal data relating to "offences, criminal convictions or security measures" may, as a rule, only be performed under the control of official authorities.[95]

Even in instances where processing of personal data is allowed by the Data Protection Directive, according to the criteria discussed above, this processing must nonetheless adhere to certain *principles relating to data quality*.[96] Personal

88 *See* Recital 30 in the Preamble to the Data Protection Directive.

89 Observe for instance the Swedish Secrecy Act (1980:100) chapter 7, section 16 which makes access to public records dependent on the conditions set forth in the Swedish Personal Data Act (1998:204).

90 Data Protection Directive, Article 8(1). One should observe that "health" has been interpreted extensively in practice. Personal data can thus be viewed as sensitive simply if it concerns some manner of health issue. There is no further requirement that this issue is in any way "sensitive" from society's point of view (moral or otherwise). To illustrate this matter, reference can be made to a preliminary ruling by the European Court of Justice concerning a Swedish case (C-101/01). Here it was held that even personal data concerning a broken leg qualify as "sensitive" under the Data Protection Directive.

91 Data Protection Directive, Article 8(2a).

92 The difference between *unambiguous* consent (required in the case of "normal" personal data processing) and *explicit* consent (required in the case of sensitive personal data processing) has not been well clarified in the directive – something which obviously causes negative effects in terms of harmonisation across member states. All that can be discerned is that "explicit consent" refers to a manner of consent that is stronger and clearer than that which is merely unambiguous.

93 Data Protection Directive, Article 8(2e).

94 Data Protection Directive, Article 8(2e).

95 Data Protection Directive, Article 8(5).

96 Data Protection Directive, Article 6(1a-e). In this context, one could also observe the rights of individuals to receive information regarding data about them being retrieved/processed. This includes information concerning the identity of the data controller as well as regarding the purpose for which the data are collected and subsequently put to use. These obligations are also relevant when personal data are retrieved from someone other than the individual

data must be processed *fairly and lawfully*. Additionally, they must be kept *accurate* and *updated*. Another important principle requires data controllers to specify the *purpose* behind certain retrieval of personal data – not deviating from this purpose while the data is retained. Concerning this "purpose principle", one should also note that retrieved data must *not be excessive* in relation to the purpose for which they are collected. They must also not be stored for a longer period of time than is necessary for the purpose to be realised.

The data controller is obliged to comply with the above principles of data quality, something that can be difficult in our society of swift network transfers and corresponding threats of interception and data distortion. Obviously, technical safeguards play an important role in ensuring that personal data are stored as securely as possible. For instance, firewalls (security precautions limiting network access) and anti-virus programs (protection against malicious code) can considerably diminish risks pertaining to theft and corruption of personal data. If these measures are insufficient, the data controller can even be forced to disable access to external networks (primarily internet) on any systems that are utilised for the purpose of handling personal data with high requirements concerning secrecy. Depending on the confidentiality level of the personal data in question, this could occasionally be the recommended option.

Other important technical means by which to safeguard oneself from risks of data intrusion (and general malfunctioning of systems) include creating security copies of information – thereby enabling subsequent restoration of the information in the event disaster strikes. Information technology has vastly simplified the process of creating identical copies of specific information. With the advent of new technologies such as writable CDs and DVDs, as well as hard drives with larger capacity, it has become a simple matter to store enormous amounts of text information on small physical media.[97] However, an interesting paradox of information security in the information age relates to the creation of security copies. On the one hand, such copies play an important role in ensuring the secure storage of information and, consequently, improving its integrity. On the other hand, when security copies are produced in inordinate amounts, there is an apparent risk that principles prohibiting excessive and over-long data storage are neglected. In the predominantly paperbound era of years past, it was often necessary from a practical point of view to clear out entire rooms of manual filing systems in order to make room for new information. The seemingly inconspicuous CDs/DVDs and hard drives of our modern age may be more easily overlooked – consequently posing a greater threat to legal adherence.

---

whom they concern. *See* the Data Protection Directive, Articles 10-11. Individuals also have the right to directly request information concerning whether data relating to them is being processed (as well as concerning the purposes of the processing). This latter right follows from the Directive's Article 12.

97 CD stands for Compact Disc. The meaning of the acronym DVD is more disputed. Initially, it stood for Digital Video Disc. As DVD-technology progressed, some agreement was reached concerning the more accurate (but rather awkward) name Digital Versatile Disc. Today, it appears as if "DVD" should in fact not be seen as an acronym at all, but simply be accepted as the name of the medium.

Literature

Bing, Jon, *Legal Decisions and Computerized Systems*, in Seipel, Peter (editor), From Data Protection to Knowledge Machines: The Study of Law and Informatics, pp. 223-250, Kluwer Law and Taxation Publishers, Deventer, 1990.

Blume P, *The Citizens' Data Protection*, in The Journal of Information, Law and Technology (JILT), 1998(1) "http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/1998_1/blume/"

Bruce, Tom, *Legal Information and the Internet in USA*, in Legal Information and the Internet – Experiences and Challenges (SOU 2002:102), p. 7-26, Fritzes, 2002.

Dworkin, Ronald, *Taking Rights Seriously*, Duckworth, London, 1977.

Gauch, Susan et al., *KeyConcept: A Conceptual Search Engine*, Information and Telecommunication Technology Center (Technical Report), University of Kansas (USA), 2004. [cit. Gauch et al.]

Hart, H. L. A., *The Concept of Law*, Clarendon Press, Oxford, 1961.

Helling, Erik, *Obstacles to the Development of Legal Knowledge-Based Systems* (IRI report 2003:2), The Law and Informatics Research Institute, Stockholm, 2003. [cit Helling 2003]

Helling, Erik, *Proactivity and Interdisciplinary Co-operation in the Context of the Legal Role*, in Nordic Yearbook of Law and Informatics 2003, Jure, Stockholm, 2004. [cit Helling 2004]

Karlgren, Jussi, Stylistic Experiments for Information Retrieval, SICS, Stockholm, 2000.

Magnusson, Cecilia, *Trunkering vid sökning i juridisk text. Praktisk genomgång* (IRI report 1985:7), The Law and Informatics Research Institute, Stockholm, 1985. [see also Magnusson Sjöberg, Cecilia]

Magnusson Sjöberg, Cecilia, *Critical Factors in Legal Document Management*, Jure, Stockholm, 1998.

Polanyi, Michael, *Personal Knowledge: Towards a Post-Critical Philosophy*, Routledge, London, 1958/1974.

Reichman, Jerome & Samuelson, Pamela, *Intellectual property rights in data?*, in *Vanderbilt Law Review 50(1)*, p. 51–166, Vanderbilt University Law School, (January issue) 1997.

Susskind, Richard, *The Future of Law: Facing the Challenges of Information Technology*, Oxford University Press, 1996 (Revised Paperback Edition 1998).

Trajkova, Joana & Gauch, Susan, *Improving Ontology-Based User Profiles*, paper submitted for RIAO'2004 (held in Avignon, France). [cit Trajkova & Gauch]

Wahlgren, Peter, *Automation of Legal Reasoning, A Study on Artificial Intelligence and Law*, Kluwer Law and Taxation Publishers, Stockholm, 1992. [cit Wahlgren 1992]

Wahlgren, Peter, *Notisum – Rättsnätet via Internet*, in Juridisk Tidskrift, 1997-98 (issue 1), p. 254-262, Stockholm University. [cit Wahlgren 1998]