Part I

Relevance and retrieval

Chapter 2

The concept of relevance

2.1 The relevance of relevance

The purpose of this chapter is to explore the concept of relevance, both in general as studied in information science, and in a jurisprudential context. Common to both these fields are that they are both highly theoretical.[48] Information retrieval-based studies of relevance, which tend to be more practical, will be examined in the next chapter.

But before we start to think about what relevance *is*, it can be useful to think about *why* relevance is interesting or useful – in other words, why relevance, in the context of legal information systems, may be relevant.

A key aspect of the concept of relevance is that everyone has an intuitive understanding of what it means.[49] In particular, it's often easy for a person to determine if a particular document is relevant to a particular problem, at least if the person is knowledgeable in the particular field that the document and the problem belong to. But precisely because everyone has an intuitive understanding of what "relevant" means, it is often difficult to explain *why* the person comes to the conclusion that a document is relevant. The intuitiveness of the concept masks the fact that it can be used in different meanings. This problem is made worse by the fact that two people, with similar background knowledge in the field, may come to different conclusions as to the relevance of the particular document. Unless we have a good understanding of what we can mean by relevance, and the different aspects that make up the concept, we will not be able to explain differences in relevance judgments. If we lack tools to explain the concept of relevance, we will not be able to study the concept of relevance in a scientific way.

Scientific method involves hypothesizing about the nature of the subject

---

[48]Note that legal substantive norms regarding relevance, e.g. rules of evidence, are fairly practical, but these will not be handled in depth.

[49]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 324.

to be researched in such a way that these hypotheses can be formulated into falsifiable propositions. It must be possible to test a hypothesis, and furthermore it must be possible to prove the hypothesis wrong (thereby falsifying it). When relevance is defined in a subjective, user-dependent way (which it often is), this becomes very problematic. Some proposed hypotheses of relevance are not testable.[50]

However, when a hypothesis of relevance is interpreted, the result may be possible to express in a computer program and used in an information retrieval system. Such *interpretations* of relevance underlie every information retrieval system.[51]

2.2 The many faces of relevance

So how do we go about explaining relevance? One way might be to examine a couple of definitions that have been suggested, and start from those.

Different academic disciplines have examined relevance from different perspectives, but two closely related disciplines that have devoted particular attention to the concept is information science and information retrieval.

The definitions we will examine come from the field of Information science. This field is concerned with the scientific study of behavior and effects of information as a phenomenon.[52] It has roots in examining *scientific* communication, i.e. peer-reviewed publishing of articles in journals, which then are subject to both critique and refinement by a network of scientists. Science is constructed part by part, in fragments, not by complete treaties. This process enables science to grow at a relatively quicker pace than the rest of scholarship (including jurisprudence).[53] In order to be effective, this scientific communication must deal with relevant information.

2.2.1 Topicality

An early definition of relevance was suggested by Vickery in 1959. The definition was for the purpose of structuring a system for retrieving hardcopy documents using card based indexes, and is as follows:

> "We can now state what is meant by items "relevant" to a particular sought subject. The limits of relevance can be varied at the discretion of the designer of the retrieval system. The system can be made to retrieve items recorded for the subject sought, and, in addition, items recorded for subjects (1) which

---

[50]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 328.

[51]Ibid., p. 321.

[52]Ibid., p. 323.

[53]Ibid., p. 323.

include the subject sought, (2) which are included by the subject sought, and (3) which are coordinate with the subject sought."[54]

The definition is based on a view of relevance as a relationship between subjects and documents that are classified as being about those subjects. This view is focused around subjects or *topics*, and assumes that there exists a number (possibly infinite) of topics, and that every document, as well as every information need, can be described as belonging to one or more topics.

On closer examination this definition does not answer the most puzzling part of relevance, i.e. what makes a particular item (i.e. document) relevant to a particular sought subject (or information need). The definition assumes that there exists a hierarchical system of terms and subjects called an information lattice (what nowadays would be called a taxonomy), that all documents already have been classified using one or more subjects from this taxonomy, and that information needs correspond exactly to this taxonomy. Given all that, the definition only offers a way of *expanding* the list of immediately relevant documents (those directly classified as the subject sought) to those that either deal with more general or more specific subjects.

This view would later be known as topical relevance or topicality. This view generally assumes that relevance is an objective property, i.e. whether a document is relevant to an information need is not a question of different (subjective) opinions.[55] A relevance relation is not dependent on any particular user, background knowledge, goals, motivation etc. All that matters are subjects encoded into documents and information needs, that is, properties of the IR system. This school of thinking can be characterized as a *system oriented view*.

2.2.2 Implicit use orientation

A few years later, an influential study by Cuadra and Katter examined the process of judging relevance through empirical tests.[56]

"In this task, judges considered separately each possible pairing between the nine abstracts and the eight requirement statements (seventy-two pairs). For each pair they rated the overall

---

[54]B.C. Vickery: The Structure of Information Retrieval Systems, in: Proceedings of the International Conference on Scientific Information, 1959, p. 1277.

[55]However, it would not be correct to equate topical relevance with objective relevance. See e.g. Bing: Handbook of Legal Information Retrieval (see n. 12), p. 199 about the nature of "content relevance", a different term for topical relevance.

[56]A note on the usage of "influential": As a concept, this is as multidimensional as relevance. In the context of this thesis, I use it to describe a work or a line of thought that others have repeatedly referenced, extended or even criticized. Clues for the existance of influence can often be found in the list of references (citations) in later works – in many ways influence works similarly to relevance.

> degree of relevance of each requirement statement to the document from which the abstract was drawn. Each rating was made on a nine-point category scale varying from Absolutely No Relevance to Absolutely Maximum Relevance. Relevance was defined for the judges as 'how good an answer the journal article would be to the requirement statement"'.[57]

The study is a good representative of a *user oriented view*. As part of the instruction to judges, they were asked to evaluate relevance as an agent for a user who had a particular task to perform. Such tasks included searching for precise definitions of terminology, or compiling an exhaustive biography for a particular subject. Even though the documents and the information needs (i.e. the requirement statements) were identical, the purpose for which the end user was going to use the document (the *implicit use orientation*) affected the relevance judgments.[58]

The system oriented view of relevance starts with the retrieval system and models relevance in terms of what the system can be made to do. The user oriented view starts with the user and examines the different, sometimes incompatible, ideas of what relevance means for the user. Between these starting points there is a fundamental difference. This split between system oriented views and user oriented views have persisted to this day, although efforts have been made to bridge the gap through *integrative relevance models*.[59]

(It should be noted that Vickery at the time of publication for the paper described in the previous section also published a different paper about how to actually create the taxonomy in such a way that it corresponds to the actual needs of the user, given a particular set of documents.[60] This is another example of a user oriented view.)

2.2.3 Multidimensionality

At the same time as the Cuadra and Katter study, Rees and Schultz did a similar study of what variables that affect relevance judgments.[61] As a

---

[57]Carlos A. Cuadra/Robert V. Katter: Opening the black box of 'Relevance', in: Journal of documentation 23.4 (1967), pp. 291–303, here p. 296.

[58]One could argue that the implicit use orientation is part of the information need. The important thing is that the topical relevance view, together with other system oriented views, does not allow for defining information needs using such subjective aspects

[59]Tefko Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II, in: Advances In Librarianship 30 (2006), pp. 3–71, here pp. 24.

[60]B.C. Vickery: Subject Analysis for Information Retrieval, in: Proceedings of the International Conference on Scientific Information, 1959.

[61]Rees/Schultz: A field experimental approach to the study of relevance assessments in relation to document searching, I: Final report, 1967, cited in Linda Schamber/Michael B. Eisenberg/Michael S. Nilan: A Re-Examination Of Relevance: Toward A Dynamic,

starting point, they used the following for a definition of relevance:

> "A relation between system responses and user request established by a judgment made by the user or his delegate."[62]

While not only the user request but also a user judgment is present in this definition, and thus the definition is not strictly a result of a system-oriented view, the interesting part of this definition was that Rees and Schultz also defined *usefulness* as the relationship between the information, the user's information need and also the individual characteristics of the user. In Rees and Schultz model, relevance and utility are separate, but partially overlapping concepts.[63] In this, Rees and Schultz went further in defining relevance as a *multidimensional concept*, encompassing both system-oriented and user-oriented views.

With this understanding, Rees and Schultz then proceeded to map out different variables – over 40 in all – that affected relevance judgments. These included properties both relating to the judge (i.e. education level or research experience), the documents (i.e. characteristics of the document set and document representations) as well as the judgment environment (i.e. research stage). Particularly the latter variables are reminiscent of implicit use orientation as defined by Cuadra and Katter.

2.2.4 Logical relevance

While Cuadra and Katter incorporated sociological and psychological elements, examining how the implicit use orientation affected human judgments, another early and influential paper by Cooper took a more mathematical approach to defining relevance.[64] The definition is based on a hypothetical information retrieval system capable of inference, i.e. not only storing and retrieving information but also using it for *reasoning*. The system described deals with fact retrieval, not document retrieval, in which searching is the same as deducing answers. As such, the approach is closer to expert systems than what we today mean by information retrieval systems.[65] Its definition of relevance is:

Situational Definition, in: Information Processing & Management 26.6 (1990), pp. 755–776

[62]Rees/Schultz: A field experimental approach to the study of relevance assessments in relation to document searching, I: Final report (see n. 61), pp. 17–18, cited in Schamber/Eisenberg/Nilan: A Re-Examination Of Relevance: Toward A Dynamic, Situational Definition (see n. 61), p. 763

[63]Ibid., p. 759.

[64]W. S. Cooper: A definition of relevance for information retrieval, in: Information Storage Retrieval 7 (1971), pp. 19–37.

[65]Expert systems are further described in 4.2

> "A stored sentence is logically relevant to (a representation of) an information need if and only if it is a member of some minimal premiss set of stored sentences for some component statement of that need."[66]

The definition is built upon earlier definitions in the paper. In particular, "minimal premiss set" is defined the smallest possible set of facts that yield, as a logical consequence of these facts, the answer to the search question. "Component statement" refers to the pair of questions used by the system for transforming queries into questions that the system can answer, i.e. "Is p a logical consequence of the stored data?" and "Is not-p a logical consequence of the stored data".

(Note that if none of these questions can be affirmatively answered, the end result is *uncertainty*, which conceptually is different from the similar no-results-case in traditional IR, i.e. "No documents found".)

The user's actual information need is a psychological state that is not directly visible, and as such cannot be input to a system. This definition requires that it is possible to *represent* this state in a form that can be used as input. It also requires that the meaning of all information in the system can be precisely and unambiguously defined, i.e. that the system is *fully inferential*. While this definition is a extreme example of a system-oriented view, Cooper maintains that this basic definition can be adapted to non-inferential systems by taking into account subjective properties of the user (i.e. background knowledge), and from a purely deductive system (where every fact in the system is either fully relevant for a particular question or not relevant at all) to a system built on plausible (or probabilistic) reasoning.

It is also noteworthy that Cooper, similar to Rees and Schultz, introduced a term to cover non-system oriented manifestations of relevance. Cooper calls this *utility*, and defines it as "the ultimate usefulness of the piece of information to the user".[67]

2.2.5   A relevance framework

The history of relevance research is generally divided into three periods: Before 1958, 1959 – 1976, and 1977 – present.[68]

The middle of these periods start with Vickerys influential papers cited above. It ends with a similarly influential article by Saracevic, which summarizes the then-contemporary views of relevance within a cohesive conceptual framework. In the article, a meta-definition of relevance is described:

---

[66]Cooper: A definition of relevance for information retrieval (see n. 64), p. 24.
[67]Ibid., p. 20.
[68]Stefano Mizzaro: Relevance: The Whole History, in: Journal of the American Society for Information Science 48.9 (Sept. 1997), pp. 810–832, here p. 813.

> "Relevance is the (A) *gage of relevance* of a (B) *aspect of relevance* existing between a (C) *object being judged* and (D) *frame of reference* as determined by an (E) *assessor*."[69]

It is centered around a number of views, where each view is concerned with some aspects of relevance. In particular, they consider a specific relation between an object being judged and a frame of reference, i.e. the subject knowledge view considers the relation between knowledge of a subject and a question about the subject, while the pragmatic view considers the relation between the problem at hand and the provided information. None of the views are complete, by definition.

This framework for examining relevance was later named "the communication framework".[70] Saracevic states that relevance is about increasing the effectiveness of communication. The value of examining relevance in information retrieval systems lies in the opportunity to make those systems more efficient. Efficient, in this context, is not about processing queries and returning results faster, but rather better at facilitating the *communication* between the user and the system. The need for more efficiency increases as the complexity of the world and our information about it increases.[71]

2.3   Relevance, examined

Information science is not the only academic discipline to examine relevance. Theories about the nature of relevance existed in logic and philosophy long before the advent of information science. Even after the co-dependent rise of information science and information retrieval, the field of pragmatics (a subfield of linguistics) established a theory of relevance as part of decoding meaning in human communication.[72]

However, information science as such has not formed an established theory of relevance. There have been attempts to adapt theories from all of the other three fields to the problems of information science, but none have been successful.[73]

---

[69]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 328, descriptive labels adapted from Schamber/Eisenberg/Nilan: A Re-Examination Of Relevance: Toward A Dynamic, Situational Definition (see n. 61), p. 761

[70]Tefko Saracevic: Relevance reconsidered, in: Information science: Integration in perspectives. Proceedings of the Second Conference on Conceptions of Library and Information Science (CoLIS 2), 1996, pp. 201–218, here p. 6.

[71]Idem: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 321.

[72]See e.g. Deirdre Wilson/Dan Sperber: Blackwell's Handbook of Pragmatics, in: 2004, chap. Relevance Theory, pp. 607–632

[73]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 20.

This does not mean that information science has failed as an academic field in regard to relevance. On the contrary, several different relevance models have been proposed, and a large number of manifestations, behavior and effects of relevance have been identified. There exists a large body of literature concerning all kinds of concepts associated with relevance.

Unfortunately, the nomenclature, or the terms for those concepts, is not universally agreed upon.[74] One authors' "pertinence" may be another authors' "utility", and drawing the line between a manifestation of relevance and a model of relevance is not always easy. For the overview of the field in this section, the nomenclature from a recent review by Saracevic has been adopted.[75]

Relevance is a *phenomenon*. A common way of studying a phenomenon is to examine how it manifests itself, how it behaves, and what effects it has. This empirical evidence can be used to test or inspire models and theories. As stated above, there exists no established theory of relevance in information science. However, there does exist several models.

2.3.1 Manifestations of relevance

Many authors that have written about relevance have observed that while relevance is universally defined as a relation between two things, the nature of those things, and thus the resulting relevancy, can vary wildly – in other words, the phenomenon of relevance manifests itself in different ways.

At one end, consider the relation between a boolean query and a document matching that query – we say that the document is relevant to the query.[76] At the other end, consider the relation between a problem and the new knowledge a specific person needs to solve the problem – we say that the knowledge is relevant to the problem. But these are two very different relationships, and thus two very different manifestations of relevance.

Many authors have attempted to enumerate the different manifestations of relevance. These attempts have not always used the term "manifestation of relevance" – common other terms include "type of relevance", "kind of relevance" or "view of relevance".

A common, but simplistic,[77] enumeration is to separate "topical relevance" from "user relevance". The former is the relation between a particular topic (subject, concept) and documents that, in one way or another, are about that topic. The latter is the relation between the problems and needs

---

[74]Stefano Mizzaro: How many relevances in information retrieval, in: Interacting with computers 10 (1998), pp. 303–320, here p. 305.

[75]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59).

[76]Boolean queries and the family of boolean retrieval models are further described in sec. 3.1.1

[77]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 33.

of a particular user and information that can satisfy that need. Topical relevance is mostly an objective relation – if two people disagree on whether a document is relevant for a certain topic, one of them is probably wrong.[78]

But there has been many more relevancies suggested. And once the number of possible entities in a space exceeds two, there is a need to structure these entities in some way. Mizzaro has suggested the perhaps most sophisticated model.[79] By starting from the universally accepted observation that relevance is a relationship between two things, he suggests that all manifestations of relevance can be classified along four axes:[80]

1. **Information resource:** This can be either a document, a surrogate for a document (such as an index card for it), and information (typically received from a document).

2. **Representation of user problem:** This can be either the actual problem, an information need (the representation of a problem as perceived by a user), a request (the information need expressed in natural language) or a query (the request expressed in a "system language" such as search terms and modifiers).

3. **Components:** This can be the topic (or subject area) that the user is interested in, the task that the user is performing, or the context (everything that is not part of the topic or task, but that affects the search process and relevance evaluation)

4. **Time:** The point in time, or the stage of the problem resolving process, that the user is in. What is relevant can change as the user learns more about the problem at hand. Unlike the other axes, which have a small number of possible values, this axis is continuous.

Using Mizarros classification system, it becomes possible to classify a given manifestation of relevance as a point (or set of points) in a four-dimensional space, such as "the relevance of a document to an information need for what concerns the task at the arising of the problem".

While this system is undoubtedly sophisticated and flexible, it is certainly complex. A simpler way of classifying relevance manifestations is to

---

[78]Bing: Handbook of Legal Information Retrieval (see n. 12), p. 199 discusses the nature of "content relevance", which is another term for topical relevance, and concludes that in a strictly formal system, content relevance is objective (Bing uses the term "absolute"), but that the legal system must afford for differing opinions regarding topical relevance. However, these differences are not the norm and do not mean that topical relevance in law is entirely subjective.

[79]Mizzaro: Relevance: The Whole History (see n. 68).

[80]Labels for these are derived from Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 33

place them on a one-dimensional scale. Saracevic has done just that, ranging the manifestations from objective to subjective:[81]

1. **System relevance:** Also known as algorithmic relevance, this denotes the relationship between a query, expressed in a way directly applicable by an IR system, and the objects (typically documents or pointers (surrogates) to texts) retrieved by the system. Logical relevance, as defined by Cooper,[82] is a stringent type of algorithmic relevance, as are the retrieval models described in sec. 3.1.1.

2. **Topical relevance:** Also known as subject relevance or topicality, this is the relationship between the topic expressed by a query, and documents that are about that topic. Determining this "aboutness" is done by humans, and is not entirely objective. Since many forms of algorithmic relevance also attempt to classify information into topics, by definition in a completely objective way, it has been suggested that this manifestation should be denoted "intellectual topicality" to clarify the difference.[83]

3. **Pertinence:** Also known as cognitive relevance, this is a relationship between a users information need, taking into account the user's current background knowledge, and information addressing that information need in the system. As the users' background knowledge changes over time, this relevance manifestation is dynamic in a way that the two previous aren't.

4. **Utility:** Also known as situational relevance, this is a relation between a situation at hand and information in the system. This differs from pertinence in that it covers more than just the specific information need, and takes into account e.g. the extent to which the user can make use of the information and the extent to which the information reduces uncertainness regarding the situation.

5. **Motivational relevance:** Also known as affective relevance, this is a relationship between the users' ultimate goal, intents and motivation and the information in the system. If the user is satisfied, the goal is accomplished etc., then the information is motivationally relevant.

Note that this is not to say that there exist exactly five relevance manifestations. Clearly there are many more. But the above classification provides a way of categorizing manifestations on a simple one-dimensional scale.

---

[81]Saracevic: Relevance reconsidered (see n. 70), p. 12.

[82]Cooper: A definition of relevance for information retrieval (see n. 64).

[83]Pia Borlund: The Concept of Relevance in IR, in: Journal of the Americal Society for Information Science and Technology 54.10 (2003), pp. 913–925, here p. 915.

This classification has been adapted by other authors, although not without criticism.[84] Borlund has suggested that the "drive" to want information (motivational relevance) is not an independent manifestation of relevance but rather an inherent characteristic of relevance behavior in general.[85] Cosjin and Ingwersen has argued in a similar way, and proposes that the fifth category should instead be *socio-cognitive relevance*, defined as the relationship between the situation at hand as perceived by the user in the users' socio-cultural context and information in the system. Compared to motivational relevance, this introduces an element of non-subjectiveness at the subjective end of this scale, as the user will attempt to understand the values and methods (and thus the rules for inferring relevance) of the users socio-cultural context in an objective way.

We initially stated that the tendency to divide relevance into just two kinds is overly simplistic. However, it is possible to describe two distinct *classes* of relevance manifestations: Those that are entirely objective, (also known as system-based manifestations), and those that include subjectivity (also known as human or user-based manifestations).[86] In Saracevics classification schema above, only system relevance is entirely objective.

It appears that the system-based manifestations are by necessity simpler and more primitive than the user-based manifestations, and that they cannot contain the entire complexity of the ultimately human notion of relevance. And that may very well be. But algorithmic relevance is the only kind of relevance that a designer of an IR system knows how to deal with.[87] And while the stated goal of research into the user-based manifestations is to inform the design of IR systems,[88] telling systems designers *how* to incorporate user-based manifestations (and just not *why* they should) is a lot harder than it appears.[89] It may be, to paraphrase Donald Knuth, that user-based manifestations of relevance is still more art than science.[90]

2.3.2 Determining relevance

A common way of studying a phenomenon in a scientific way is to study its manifestations, behavior and effects. We have covered relevance manifestations, but relevance does not, in itself, behave. However, people behave

---

[84]See e.g. Erica Cosijn/Peter Ingwersen: Dimensions of relevance, in: Information Processing & Management 36 (2000), pp. 533–550 and Borlund: The Concept of Relevance in IR (see n. 83)

[85]Ibid., p. 915.

[86]Ibid., p. 914.

[87]Cooper: A definition of relevance for information retrieval (see n. 64), p. 36.

[88]Schamber/Eisenberg/Nilan: A Re-Examination Of Relevance: Toward A Dynamic, Situational Definition (see n. 61), p. 773.

[89]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 25.

[90]See the quote at the beginning of this thesis

around relevance, particularly when determining whether relevance exists.[91]
The question of how people determine relevance deals both with cognitive
processes within the user as with the sociological context in which the user
and the information need exists. The determination process can be de-
scribed as inferential - the user observes aspects of the problem, available
information and also background knowledge, and uses this to infer whether
there exists some sort of relevance manifestation.

Relevance clues

We can distinguish between two different ways of studying the relevance de-
termination processes. The first, which has roots in librarianship, concerns
topical relevance only. In this process, a user has to determine whether a
given document is relevant to a certain fixed topic. In this process, there
is not a lot of room for individual differences e.g. in taste or background
knowledge. Given a shared understanding of the boundaries of a topic, two
different individuals will often come to the same conclusion as to whether
a document is relevant. However, some of the time judgments will differ
due to a number of reasons. We have covered the concept of *implicit use
orientation*,[92] the effect of the assumed document usage on the relevance
determination.[93]

The second approach is applicable for all manifestations of relevance. It
focuses on characteristics or "clues" found in information objects that users
use to determine relevance. This approach, outlined by Schamber et. al in
1990,[94] became known as "clues research".[95]

(It should be noted that the observation that users use a large number of
characteristics to determine relevance is much older – Cooper suggested in
1971 that "[a] document may be relevant to an information need but useless
because the user has no faith in its accuracy")[96]

A number of general studies later confirmed that it was possible to iden-
tify a number of classes of clues.[97] However, the number and exact list of
such clues seem to differ from study to study.

Still, some general patterns can be discerned. The number of clues classes
is generally small (less than 20), and users make use of several of them in

---

[91]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on
the Notion in Information Science. Part II (see n. 59), p. 37.

[92]See sec. 2.2.2

[93]See sec. 2.3.4 for some other reasons why individual judgment differ, even for some-
thing as nominally objective as topicality.

[94]Schamber/Eisenberg/Nilan: A Re-Examination Of Relevance: Toward A Dynamic,
Situational Definition (see n. 61), p. 773.

[95]Saracevic: Relevance reconsidered (see n. 70), p. 12.

[96]Cooper: A definition of relevance for information retrieval (see n. 64), p. 36.

[97]See Saracevic: Relevance: A Review of the Literature and a Framework for Thinking
on the Notion in Information Science. Part II (see n. 59), p. 39 for a summary of these
studies

conjunction. In general, clues can be grouped into the following groups (sectioned into two partitions):[98]

Characteristics of the information object	Characteristics dependent on the information seeker
Content characteristics, such as its topic, quality, recentness or clarity. *Information object* characteristics such as type, format, availability and costs. *Validity* characteristics such as accuracy, authority and verifiability.	*Situational match*: Appropriate to the current situation or task. *Cognitive match*: Understandable, containing new information, amount of effort needed. *Affective match*: Emotional reactions such as entertainment, frustration, uncertainty. *Belief match*: Confidence in the truth of the material.

While users in general do not think of relevance manifestations using Saracevics five categories described above, they are perfectly capable of distinguishing between e.g. topical relevance and utility. A user can for instance recognize that a document is about the topic that they are researching (i.e. being topically relevant), but not useful for solving the problem at hand (i.e. not being situationally relevant).[99]

Characteristics of the information object are, as opposed to characteristics dependent on the information seeker, sometimes possible to express and embed in information systems.[100] What characteristics are possible and appropriate to use for this? The obvious choice is the textual content of the documents. If a word that exists in the query also exists in the document, they might have a relevance relationship. In most information retrieval models, the document is not considered as an entity with a structure but rather a simple "bag of words".[101]

But documents are more than the sum of their words. Many kinds of documents, particularly in scientific publishing and legal information, contain a rich set of references or citations. These citations may also act as relevance clues.[102]

The use of citations as evidence of systemic relevance is not new. The

---

[98]Ibid., p. 43.

[99]Mizzaro: Relevance: The Whole History (see n. 68), p. 823.

[100]c.f. Schamber/Eisenberg/Nilan: A Re-Examination Of Relevance: Toward A Dynamic, Situational Definition (see n. 61), p. 773, research questions 9-11

[101]Christopher D. Manning/Prabhakar Raghavan/Hinrich Schüze: Introduction to information retrieval, 2008, sec. 6.2.

[102]Jussi Karlgren: Information Retrieval: Statistics and Linguistics, Excerpt from the PhD Thesis "Stylistic Experiments in Information Retrieval", 2000, sec. 3.6.1.

first scholar concerned with Citation indexes was Garfield who founded the field of bibliometry.[103] This field uses analysis of citation networks to, amongst other things, create impact rankings of academic journals.[104]. This line of research soon asserted that citations seem to be indicative of scholarly impact and quality. This assessment is not without controversy – in particular, citations can and often is made for other reasons than to indicate quality, including negative citations (citations to works being criticized) which may lead to inflated citation counts for works of low quality. Still, correlations between reputational ratings (representing a measurement of scholarly quality) and citation rates have found to be very high.

Relevance dynamics

Relevance is not determined once and then never changed. The search process is dynamic, and as the user learns more about the problem, he/she reinterprets the task and is thus able to reformulate the information need statement into new queries. Furthermore, during this cognitive process, documents that were previously believed to be relevant may turn out to be non-relevant, or vice versa. During the search process, both the task and the users' cognitive state are constantly changing.[105]

Starting with Rees and Schultz in 1967, this aspect of relevance has been examined in numerous studies. As the studies have differed in scope and method, general conclusions are fairly basic. It can be observed that the clues that the user makes use of for inferring relevance do not change as the process move forward, but that different criteria are given different weight over time. As time progresses, the relevance judgments becomes more stringent.[106]

Changes in relevance inference do not just occur during the search process. Over the course of years and decades, additional information may change a relevance relation without changing neither the problem/topic, the information itself, nor even individual characteristics. If a particular paper, which at one time was thought topical and valuable and thus relevant, later gets discredited, it will probably not be thought of as cognitive or situationally relevant (although it may still be topically relevant).

A problem with using citation analysis for inferring relevance over the long run is the problem of "obliteration through incorporation", when a work is so influential and well known that writers no longer feel the need

---

[103]Eugene Garfield: Citation Indexes for Science - A New Dimension in Documentation through Association of Ideas, in: Science 122 (July 1955), pp. 108–111.

[104]Fred R. Shapiro: The Most-Cited Law Review Articles, in: California Law Review 75 (1985), pp. 1540–1554, here p. 3.

[105]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 41.

[106]Ibid., p. 44.

to explicitly cite it.[107] This problem is common in legal information as well, particularly when landmark cases are reinterpreted as general principles of law and thus not cited anymore.

2.3.3 Measuring relevance

In Saracevics 1975 meta-definition (summarized in sec. 2.2.5), relevance is defined as being a *gage of relevance*, and this gage can be exemplified using terms such "measure" or "degree".[108] This hints at a part of relevance that is intuitively understood, but which may be subject to closer scrutiny: Is relevance a binary measure (either something is relevant or it is not), a continuous measure, or does it have several, but discrete values?

In information retrieval, different retrieval models assume different answers to the question. Standard boolean retrieval models have - as the name implies - a binary view of relevance, due to the set theoretic and first-order logic foundation of boolean retrieval. Either a document satisfies the constraints that the query expresses, or it doesn't. Probabilistic models attempt to measure the *probability* that a given document is relevant, and as such assume a binary scale of relevance (but a continuous scale of probability). Vector space models attempt to measure the *similarity* between a document and a query, and assume that a high degree of similarity correlates with a high degree of relevance. Therefore, the vector space models assume a continuous scale of relevance. The different retrieval models are described in more detail in 3.1.1.

In evaluation of IR systems, it is common to assume a binary view of relevance measures. In practice, evaluation takes place against a predetermined set of binary relevance judgments, most often made from a topical relevance perspective. This basic view of relevance measure is a prerequisite for the traditional evaluation metrics of precision and recall, and is a part of the so-called "Cranfield paradigm" of IR evaluation.[109] The binary view of relevance have been defended by Cooper for his definition of logical relevance, stating that the argument for introducing different degrees of relevance is confusing the definition of relevance with the difficulty of creating useful search interfaces.[110]

However, even Cooper concedes that there is ground for treating relevance as a continuous measure when moving from strict logical relevance (where relevance is deduced) to a system where plausible or probabilistic reasoning is used to infer relevance. Since probability has degrees, relevance

---

[107]Shapiro: The Most-Cited Law Review Articles (see n. 104), p. 3.
[108]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 328.
[109]See further sec. 3.1.3
[110]Cooper: A definition of relevance for information retrieval (see n. 64), p. 26.

as used in such systems needs to have degrees as well.[111]

And as we move further from the system view of relevance, we find more arguments for treating relevance as a continuous measure. Saracevic states in his communication model of relevance that relevance is considered as a measure of the effectiveness of the contact between a source and a destination in a communication process.[112] Effectiveness can be stronger or weaker. Even concerning evaluation of IR systems, Kekäläinen mean that topical relevance should be understood as a continuous variable.[113]

2.3.4 Difference in relevance judgments

Any empirical experiment concerning the concept of relevance will need some sort of assessment about whether a particular document is relevant in relation to a particular information need. However we choose to define "relevant", the assessment needs to be done by a human assessor. We know that, in practice, human judges differ in their opinion on relevance. It's therefore interesting to find out more about why this is, how large the differences are between different judges, and how these can be explained.

Apart from giving further insight into the concept of relevance, this knowledge is also necessary to do a meaningful evaluation of IR systems. If we know that relevance agreement in a particular domain is 95% (meaning that judges disagree in their relevance assessments 5% of the time), it is not really possible for an IR system to score better. Even if the system agrees with judge A 100%, it will only with judge B 95% of the time – and we have no way of determining which of judge A's or B's assessments are the "correct" ones.[114]

The fact that human judges disagree on what is relevant calls into question whether relevance can, in fact, be used as a meaningful metric when evaluating information retrieval systems. If we do not understand why human judges disagree, and we treat relevance judgments as a given, derived metrics such as precision and recall may not be useful for evaluating IR performance.[115]

In Cudra and Katter's 1967 study, some of the possible factors that influence human judges were examined.[116] The breakthrough of this study was to

---

[111]Cooper: A definition of relevance for information retrieval (see n. 64), p. 30.

[112]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 326.

[113]Jaana Kekäläinen: Binary and graded relevance in IR evaluations - Comparison of the effects on ranking of IR systems, in: Information Processing and Management 41 (2005), pp. 1019–1033, here p. 1020.

[114]In the field of natural language processing (NLP), this upper bound for system performance is known as the human ceiling. Jurafsky/Martin: Speech and language processingn (see n. 44), p. 189

[115]Cuadra/Katter: Opening the black box of 'Relevance' (see n. 57), p. 292.

[116]Ibid.

show that a number of factors outside of the information need, the document and the judge affect the judgment of relevance. The human judges in the study were given a set of documents and a set of information requirements, and asked to provide relevance scores (on a 9-point scale) for each pair. The judges were then asked to act as agents for persons with specific usage needs or tasks, and redo their relevance judgments based on this. This implicit use orientation affected relevance judgments. In taking into account and trying to measure the impact of a human judges mental state, the Cuadra and Katter study provides a useful theoretical framework for user models of relevance.

But it is not only the implicit use orientations that create differences in relevance judgments. Studies show that high subject knowledge of the judges corresponds with a low number of documents judged relevant. The higher level of subject knowledge also corresponds with higher agreement among judges.[117]

A different aspect that also affects relevance judgments is the concept of marginal relevance. A document that, by itself, would be judged as highly relevant, may be thought of as only partially relevant because it is read after another document with overlapping subject content.[118] Since a typical IR interaction presents a series of (supposedly) relevant documents in response to a query, maximizing marginal relevance requires that this result set is diverse, not just closely matched to the query.[119]

2.3.5 Models of relevance

After having examined various aspects of relevance, we can now try to look at the bigger picture. At the start of this section, it was stated that there exists no established complete theory of relevance in information science.[120] However, there exists several models (sometimes called "frameworks").[121]

Existing models can generally be divided into system models and user models. In a system model, creating systems that can process information and match this with user queries is emphasized. These relevance models must be possible to express in the form of algorithms, and there must be hard data consisting of documents and queries (or information about documents or queries) available as input to these algorithms. They are called system

---

[117]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 336.

[118]Goffman and Nevil, cited by ibid., p. 336

[119]Manning/Raghavan/Schüze: Introduction to information retrieval (see n. 101), sec. 8.5.1.

[120]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1), p. 337.

[121]A model is an abstraction that forms general ideas from specific examples. idem: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 22

models because they are to be used in an automated system. Ideally, the user is represented by a query that is directly acceptable by the system. In this model, relevance is considered a property of the system.[122]

In a user model, the emphasis is on the cognitive and social dimensions of the user as a whole, as well as the process of interaction. Apart from that, there have been a number of different models proposed in this space, with no clear consensus on what the proper model for relevance that puts the user in the center should look like.[123]

The most influential of these models are:

- **The communication framework:** This model was proposed in Saracevics 1975 review,[124] and was later named such. This model considers relevance to be a measure of effectiveness in a communication process between a user and an IR system.

- **The dynamic model:** Also known as the situational framework, this was proposed by Schamber et al.[125] While the dynamic aspects of relevance had been discussed before, this model has the dynamic and situational nature of relevance as its starting point.[126]

- **The psychological framework:** Alongside suggestions within the dynamic model, some researchers in the early 1990s suggested that the user's cognitive processes with respect to the information seeking process should be the focal point of the relevance. These suggestions were synthesized in the "psychological relevance" framework suggested by Harter in 1992.[127]

- **The stratified model:** This model was first suggested in Saracevics 1996 review of relevance research, and was further described in his 2006 review. It focuses on the information retrieval interactions and notes that interactions take place on a number of planes or *strata*. The starting point for the interaction is the surface level between user and computer (i.e. the interface to an IR system), but relevance manifests itself in strata below (information resources, algorithms and hardware) as well as above (cognitive, affective and situational processes).[128]

---

[122]Saracevic: Relevance reconsidered (see n. 70), p. 6.

[123]Idem: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 25.

[124]Idem: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science (see n. 1).

[125]Schamber/Eisenberg/Nilan: A Re-Examination Of Relevance: Toward A Dynamic, Situational Definition (see n. 61).

[126]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 23.

[127]Idem: Relevance reconsidered (see n. 70), pp. 7.

[128]Idem: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 27.

There is a correspondence between system-based relevance manifestations, and the system model of relevance. Similarly, there are lots of similarities between the user-based relevance manifestations and the various user-oriented models of relevance. It is important not to confuse relevance manifestations with models. The former deals with particular types of relevance, and there is room for several different manifestations within one coherent description of the relevance concept. The latter are alternative models for describing relevance as a whole.

2.4 Relevance in jurisprudence

2.4.1 Legal information as compared to scientific information

In the previous section, we learned that information science, having its roots in the natural sciences, is concerned with *effective communication of scientific knowledge*, and that relevance is a key factor in making communication processes effective. The communication framework model of relevance emphasizes that relevance is not an independent phenomenon, but must be considered as a measure of effectiveness.

Legal practitioners will appreciate the need for effective communication of the knowledge of the field of law. Law as an information corpus contains massive amounts of information contained in fixed texts such as statutes or legal cases. It also contains less visible but crucial knowledge in the form of meta-norms (see sec. 1.2). This knowledge is structurally organized; we can therefore talk about *legal knowledge structures*.[129]

Legal knowledge structures are not constructed like the scientific knowledge structures. While the (natural) sciences only deal in *persuasive authority* (the truth of a proposition does not depend on who states it, only if empiric data supports is and/or it is internally consistent), law also deals with *binding authority*, that is, statements from a particular source whose truth depend on that source (and other formal aspects, such as the law having been promulgated or statement being part of a verdicts *ratio decidendi*).[130]

Still, there are similarities in how the two entities, scientific knowledge and legal knowledge, are constructed. This is particularly noticeable in case law. No court ever sits down to construct the grand unified model of e.g. product liability. Instead, it is constructed case by case, in pieces. This rule-creating function of the courts is particularly prominent in common law legislations, and is generally seen as a necessary and desirable function.[131]

---

[129]Wahlgren: Automation of Legal Reasoning - A Study on Artifical Intelligaence and Law (see n. 15), p. 176.

[130]The ratio decidendi (lat. "reason for deciding" is the part of a legal verdict that expresses the reasoning for the judgment)

[131]Melvin Aron Eisenberg: The Nature of the Common Law, 1988, p. 5.

Not all law is constructed by fragment – jurisprudential scholars tend to create and present their works in the form of large treaties, and legislators of statutory law can at least in part attempt to create broad-ranged codifications of entire parts of the law (even though these tend to be based, at least in part, on earlier statutory law).

In law, the concept of relevance comes into play both in substantive norms as well as meta-norms. Meta-norms can specify sources one may draw on when constructing a legal argument (the doctrine of legal sources), what constitutes a legally sound argument, rules for resolving the situation when two applicable norms conflict with each other (such as the *lex specialis* rule),[132] and norms for estimating the precedent value of a court verdict.

Substantive norms may specify e.g. what may be permitted as evidence, and therefore be relevant, in a legal process. In Scandinavian jurisdictions, the principle of free assessment of evidence places very few limits on permissible evidence. In the Anglo-American legal tradition, much more substantive law deals with questions regarding relevance of evidence.[133] In the following, we will only address the concept of relevance with respect to meta-norms, as relevance rules in substantive norms do not help us building better general legal IR systems (although they can be useful e.g. in case management systems).

2.4.2 Jurisprudence and legal method

Jurisprudence is the theoretic analysis of law, the nature of rights and duties, and judicial reasoning.[134] The subject matter of jurisprudence is the nature of law and its workings. Jurisprudence has historically dealt with the study of legal concepts such as "property" or "rights", but now is mainly about the philosophy of law and legal theory.[135]

Jurisprudence studies the legal system as such, i.e. not what the law says about any particular question, but rather how to understand the concept of law. Included in this are theories of what a legal system is, how it derives its legitimacy, how different institutions (legislators, courts, scholars and the academic discipline of jurisprudence) relate to each other, but ultimately the question of what a valid legal reasoning process is, and how to compare the soundness of opposing legal arguments.

In order to answer this question, the legal method is developed. The legal method is a process, controlled by meta-norms, for finding answers for creating legal arguments (propositions) and in particular determining the

---

[132]The *lex specialis* rule specifies that of two applicable and conflicting norms, the most specific norm should take precedence

[133]Bing: Handbook of Legal Information Retrieval (see n. 12), p. 11.

[134]Raymond Wacks: Understanding jurisprudence: an introduction to legal theory, New York 2005, p. 12 at note 1.

[135]Penner: Textbook on Jurisprudence (see n. 46), p. 1.

validity of those arguments, from substantive norms.

Formulating propositions of law

Legal questions in the real world start with actual situations where it can be assumed that the law has a role to play. A first definition of what relevance can mean in a juridical context must be the demarcation between the facts of the case that have legal significance and facts that do not have any legal significance. The first step in the legal method can be thought of as the identification of such actual facts that have legal relevance. This is done by finding norms that in one way or the other fits the circumstances of the cases. Based on this a legal proposition or argument can be developed.

Methods of interpretation

Abstract questions in jurisprudence seldom have a direct application for the practicing lawyer who seeks the answer to the question if a particular proposition of law is correct (or if it can hold up in court). Ideally, there exists one and only one clearly formulated and uncontroversial rule stated in legislation or case law which is directly applicable to the case at hand. But if no such rule exists, a rule exists but its formulation is unclear with respects to the facts of the case, or if two or more partially applicable rules lead to different results, the legal information at hand needs to be interpreted in light of the facts of the case. Since different methods of interpretation can lead to different results there is in practice room for different answers to the basic question - whether a given proposition of law is true. There exist different schools of thought as to what method of interpretation is the proper one to use in different scenarios. It can therefore be misleading to talk about *the* legal method.

Since the subject matter for this study is EU law, it is based on the EU legal method that has been formulated primarily through the case law of the European court of justice. Distinguishing for this method, compared to the method practiced in Sweden, is a stronger focus on case law and general principles of law, and a stronger inclination to declare void both national and community legislation.

Doctrine of legal sources

Through this we arrive at the doctrine of legal sources. This augments the different methods of legal interpretation by stating from which sources we may search for norms and interpretation data for these norms when determining the truth of a legal proposition. In EU law, *general principles of law* is normally considered a valid source of law, even though these by definition are not written down (except partially, in vague terms, in the founding treaties).

Preparatory works, which are produced during the drafting of statutory community law, has generally not been considered as a source for a legal proposition, although this is changing.[136]

The issue is of interest when constructing legal information retrieval systems. The doctrine of legal sources contributes by limiting what information needs to be in our system by setting an objective outer boundary for what can possibly be relevant.[137] But even within this boundary, there are differences between sources. Even though two different sources may apply to that case at hand, they may not have the same weight in the reasoning process. A legal argument based on one source may trump a different argument based on another, simple because the first source is more important.

Considering that EU law has a strong case law tradition, and that ECJ in several cases have formulated basic principles of EU law (that later have been upheld),[138] it is not surprising that different cases have different weight in legal reasoning. Cases in the General Court (EGC) have less weight than cases decided in the court of justice (ECJ).[139]

Some of the cases in ECJ are said to be landmark cases, i.e. cases that have a strong effect on how EU law is interpreted. For any given information need, a landmark case will always be more relevant than a non-landmark case, provided that both cases cover the material that the information need concerns.

2.4.3   The notion of legal relevance

As discussed in sec. 2.3, relevance in information science can be understood as a multidimensional concept. There are numerous attempts to explain or organize the different manifestations of observable relevancies into a coherent model. This begs the question: What manifestations of legal relevance is there? Is there a legal model of relevance?

We started our discussion of relevance in general by examining a couple of influential attempts to define the term. In jurisprudence, similar definitions are rare. Most writings on relevance focus on the boundary conditions of relevance, i.e. what are the permissible sources to draw upon when constructing a legal argument. In this way, relevance is defined in terms of the doctrine of legal sources.

Perhaps the most explicit definition of legal relevance was formulated by Bing in 1984:[140]

---

[136]Jörgen Hettne/Ida Otken Eriksson (eds.): EU-rättslig metod. Teori och genomslag i svensk rättstillämpning, 2005, p. 24.

[137]Bing/Harvold: Legal Decisions and Information Systems (see n. 7), p. 39.

[138]For example *Van Gend en Loos*, 26/62, which established that the EEC treaty created binding legal rights directly enforceable in member states

[139]Hettne/Eriksson (eds.): EU-rättslig metod. Teori och genomslag i svensk rättstillämpning (see n. 136), p. 39.

[140]A similar, yet substantially different, definition appeared in Bing/Harvold: Legal De-

"A legal source is relevant if:

- (1) The argument of the user would have been different if the user did not have any knowledge of the source, i.e. at least one argument must be derived from the source; or

- (2) legal meta-norms require that the user considers whether the source belongs to category (1); or

- (3) the user himself deems it appropriate to consider whether the source belongs to category (1)."[141]

This is a definition that is reminiscent of Cooper's definition of logical relevance.[142] In essence, both these definitions state that relevance of information in relation to a question stems from the power of contributing to an answer to that question. Bing's definition contains auxiliary categories that allows for legal sources to be evaluated and dismissed, while still counting as relevant. This is to allow the user to practice the legal method instead of having the system do it.

Is legal relevance binary or graded?

Like Cooper, Bing is of the view that relevance is binary. This does not mean that all legal sources are equal. They differ in rank, age and specificness, and that this should be accounted for when different legal sources acts as base for conflicting legal arguments. Bing states that while this results in arguments with relative weight, this should not be confused with relevance of the source itself.[143]

This position differs from other interpretations of relevance in legal information. For example, Burson states that we think of documents being more or less relevant to a particular issue.[144] Some relevance studies in legal IR have established a discrete scale for relevance, usually with steps like "on point", "relevant", "related" and "irrelevant".[145] Such a scale also assumes a non-binary view of legal relevance (although not necessarily a continous scale)

---

cisions and Information Systems (see n. 7), p. 40

[141]Bing: Handbook of Legal Information Retrieval (see n. 12), p. 28.

[142]See sec. 2.2.4

[143]Bing: Handbook of Legal Information Retrieval (see n. 12), p. 30.

[144]Scott F. Burson: A Reconstruction of Thamus: Commends on the Evaluation of Legal Information Retrieval Systems, in: Law Library Journal 79 (1987), pp. 133–143, here p. 141.

[145]Such a study is cited in Dabney: The Curse of Thamus: An analysis of Full-Text Legal Document Retrieval (see n. 7), p. 15

Legal relevance manifestations

Unlike Bing's definition, other interpretations of relevance in legal information tend to be fairly simplistic and often defer to the intuitive understanding of relevance. As an example, the influential 1985 empirical study by Blair and Maron, which used the scale "vital", "satisfactory", "marginally relevant" and "irrelevant", left the question of what these labels actually meant up to the individual test subjects in the study.[146]

Such intuitive understandings shed no light on the crucial issue of what legal relevance actually is. We have learned that relevance, as studied in information science, have multiple manifestations. Bing offers a more detailed view of relevance that is based upon a classification by Königova,[147] with three manifestations (although the term "kind of relevance" is used instead of "manifestation of relevance").[148] These are:

- **Formal relevance**, defined as the syntactic similarity between two texts (where one can be a search request). This similarity may be based on word occurrence, but also things like document type, age and author. Formal relevance acts as an approximation of content or subjective relevance, defined below. This manifestation is equivalent to what Saracevic calls system relevance.

- **Content relevance**, defined as the adequacy of the content as a response to the request. It does not depend on the user's particular background knowledge. It is therefore most closely related to what Saracevic calls topical relevance.

- **Subjective relevance**, which can be described as usefulness to the user, taking into account the user's previous knowledge. This definition is a bit more vague, and might cover both "pertinence" and "utility" as defined by Saracevic.

On human disagreement of relevance

As we have seen in sec. 2.3.4, even experts on a subject sometimes disagree on whether a document is relevant or not for a given information need. This general finding is evident within legal IR as well. Since relevance is the basic criteria of the precision and recall metrics,[149], and that these metrics are commonly used to evaluate IR systems efficiency, it is important to know

---

[146] Blair/Maron: An Evaluation of Retrieval Effectiveness for a Full-text Document Retrieval System (see n. 37).

[147] Marie Königová: Mathematical and statistical methods of noise evaluation in a retrieval system, in: Information Storage and Retrieval 6 (1971), pp. 437–444.

[148] Bing: Handbook of Legal Information Retrieval (see n. 12), p. 197–203.

[149] Turtle: Text Retrieval in the Legal World (see n. 17), p. 10.

how these relevance judgments are made and why they differ. In particular, are there any specific patterns to relevance judgments for legal IR systems?

The typical evaluation methodology in the Cranfield paradigm is to create a set of information requirement statements (typically based on real user needs) together with a set of domain expert judgments for a corpus of documents. The judgments state, for each document, whether it is relevant (and, in some settings, to what degree it is relevant) for each information needs. Creating this gold standard set of judgments requires that experts are at least somewhat in agreement with each other.[150]

Numerous studies have shown that this is often not the case for legal IR:

- In one of the earliest studies about legal IR,[151] experts were in agreement on how relevant a document was (on a four point scale) only 4.3% of the time.

- Some years later, a study by Tapper showed that lawyers would disagree about the correct classification of cases (which can be described as a form of topical relevance assessment), and even that the same lawyer would describe a single case differently when presented in another context.[152]

- In the NORIS programme run between 1972-1983 at the Norwegian Research Center for Computers and Law,[153], a single expert in administrative law had to reevaluate his "gold standard" after running a experiment, from an initial set of 162 documents to a final set of 207 documents.[154]

It is easy to agree with Dabney's statement that "[t]he judgment of experts seem to be an unsound way to determine relevance"[155]. But we are not yet closer to understanding why expert judgment fails in this regard.

Legal citations as relevance clues

The information science research concerning relevance clues (described in sec. 2.3.2) suggests that relevance of all types can be inferred from characteristics of the information object. Legal information is rich with characteristics that can be of potential use as relevance clues.

---

[150]The concept of a "Gold standard" is described in sec. 3.1.3

[151]Eldridge: An Appraisal of a Case Law Retrieval Project, in: D Johnston (ed.): Proceedings of the Computers and the Law Conference, 1968, cited by Dabney: The Curse of Thamus: An analysis of Full-Text Legal Document Retrieval (see n. 7), p. 15

[152]Cited by Bing/Harvold: Legal Decisions and Information Systems (see n. 7), p. 41

[153]Bing: Handbook of Legal Information Retrieval (see n. 12), p. 236.

[154]IdemThird Generation Text Retrieval Systems, in: Journal of Law and Information Science 1.3 (1981), pp. 183–209, here p. 155.

[155]Dabney: The Curse of Thamus: An analysis of Full-Text Legal Document Retrieval (see n. 7), p. 15.

We can examine this using the grouping of relevance clues suggested by Saracevic:[156]

- **Content characteristics**:

  - **Topic**: Many types of legal information, particularly legal cases, are routinely categorized into a set of fixed or fluid categories (e.g. in headnotes)

  - **Recentness**: Most if not all legal information has information about publishing or enactment dates. This information is typically part of a document's metadata when available electronically. Note however that there is no direct correlation between recentness and relevance.

  - **Quality** and **Clarity**: These characteristics are applicable to legal information as well, but they are of a more subjective nature than Topic or Recentness.

- **Information object characteristics**:

  - **Type**: Legal information generally have a clearly indicated document type (e.g. statute, court case, preparatory work, doctrinal thesis)

  - **Format**: Documents are often available in electronic form in full text, but e.g. for court cases sometimes only headnotes are available. For jurisprudential doctrine, often only the bibliographic record is available in electronic form.

  - **Availability** and **Costs**: These characteristics are generally a function of whatever information database one is using. For EU law, the EURLEX database is complete and freely available, although in practice a lawyer often needs access to other sources such as the publication Common Market Law Review (CMLR).

- **Validity characteristics**:

  - **Accuracy** and **Verifiability**: Legal information that has binding authority can be considered to be accurate by virtue of itself. However, there may still be questions about whether the information is accurately converted into electronic form, or whether the content is up-to-date. If the user lacks ability to verify this it might result in lower relevance assessments.

  - **Authority**: Notions of authority is built into many types of legal information, i.e. a constitution has higher authority than regular

---

[156]Saracevic: Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II (see n. 59), p. 47.

statutes, and supreme court decisions have higher authority than lower-level court decisions.

In law, authority is primarily derived from the creator of, or the process of creating, the information. In this view, the authority of one document is not changed by other documents. But there is a complementary view of authority where authority for information is derived from other actors reactions (particularly confirming) to that information.

In case law, the usage of citations to earlier cases is common. There have been several suggestions that these citations are useful as clues in legal IR systems. In 1970, Marx suggested that even though traditional legal information retrieval (that is, before the advent of electronic legal IR systems) uses the process of "Shepardization" (described in 3.2.1) to follow citations between cases, automated systems of that time did not make full use of the structure derived from citations between cases.[157]

In early 1980's, Tapper did several experiments with citation networks, notably creating an algorithm for vector-based correlation of cases, with elements of the vector being based upon citation patterns.[158]

In recent years, particularly after the introduction of the Google search engine and its PageRank algorithm,[159] interest in using case citations for legal IR has sprung up. Smith suggests using PageRank-like algorithms in legal IR, but adapted for the differing objective authority of e.g. courts of different levels.[160] Harvold suggests that browsing may be a better retrieval paradigm than (free text) search, particularly if using the structure of backlinks (inbound links) between documents.[161] Geist proposes that the usage of citation analysis can improve the relevance ranking in legal IR systems.[162]

There seems to be an overwhelming opinion that techniques successful in general web search can be useful in legal IR as well. By proposing that case citations can act as relevance clues, this suggestion is consistent with relevance theory as examined in information science. By observing how citations are used within the legal method as practiced today, the suggestion is connected to jurisprudence as well.

However, one of the key aspects of modern link analysis algorithms (such as PageRank) is that not all links are of equal value. On the web, a page

---

[157]Stephen M. Marx: Citation Networks in the Law, in: Jurimetrics Journal 10 (1970), pp. 121–137, here p. 123.

[158]Colin Tapper: An experiment in the use of citation vectors in the area of legal data, 1982.

[159]Lawrence Page et al.: The PageRank Citation Ranking: Bringing Order to the Web. Technical Report 1999-66, Stanford InfoLab, Nov. 1999

[160]Thomas A. Smith: The Web of Law, in: San Diego Law Review 44 (2007), pp. 309–354, here p. 351.

[161]Harvold: Is searching the best way to retrieve legal documents? (See n. 38).

[162]Anton Geist: Using Citation Analysis Techniques For Computer-Assisted Legal Research In Continental Jurisdictions, MA thesis, Edinburgh: The University of Edinburgh, 2009, p. 50.

gains authority when other pages refer to it. For legal information, an influential court case (case A) will, over time, be cited much more often than a non-influential case (case B) by the same court at the same time. But does that mean that citations from case A should count for more than citations from case B when determining the influence for an even older case?

There is no jurisprudential explanation for why that should be the case. But there are other aspects which may give case citations different weight. The most important of these are the inherent authority of the citator, and citation patterns over time. We will revisit these aspects in sec. 3.3.