The Notion of Relevance in Information Science

Everybody knows what relevance is.

But, what is it really?

Synthesis Lectures on Information Concepts, Retrieval, and Services

Synthesis Lectures on Information Concepts, Retrieval, and Services publishes short books on topics pertaining to information science and applications of technology to information discovery, production, distribution, and management. Potential topics include: data models, indexing theory and algorithms, classification, information architecture, information economics, privacy and identity, scholarly communication, bibliometrics and webometrics, personal information management, human information behavior, digital libraries, archives and preservation, cultural informatics, information retrieval evaluation, data fusion, relevance feedback, recommendation systems, question answering, natural language processing for retrieval, text summarization, multimedia retrieval, multilingual retrieval, and exploratory search.

The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really?
Tefko Saracevic

Dynamic Information Retrieval Modeling No Access
Grace Hui Yang, Marc Sloan, Jun Wang

Learning from Multiple Social Networks No Access
Liqiang Nie, Xuemeng Song, Tat-Seng Chua

Scholarly Collaboration on the Academic Social Web No Access
Daqing He, Wei Jeng

Scalability Challenges in Web Search Engines No Access
B. Barla Cambazoglu, Ricardo Baeza-Yates

Social Informatics Evolving
Pnina Fichman, Madelyn R. Sanfilippo, and Howard Rosenbaum

Click Models for Web Search
Aleksandr Chuklin, Ilya Markov, Maarten de Rijke

Information Communication
Feicheng Ma

Social Media and Library Services
Lorri Mon

Analysis and Visualization of Citation Networks
Dangzhi Zhao and Andreas Strotmann

The Taxobook: Applications, Implementation, and Integration in Search: Part 3
Marjorie M.K. Hlava

The Taxobook: Principles and Practices of Building Taxonomies: Part 2
Marjorie M.K. Hlava

Measuring User Engagement
Mounia Lalmas, Heather O'Brien, and Elad Yom-Tov

The Taxobook: History, Theories, and Concepts of Knowledge Organization: Part 1
Marjorie M.K. Hlava

Children's Internet Search: Using Roles to Understand Children's Search Behavior
Elizabeth Foss and Allison Druin

Digital Library Technologies: Complex Objects, Annotation, Ontologies, Classification, Extraction, and Security
Edward A. Fox and Ricardo da Silva Torres

Digital Libraries Applications: CBIR, Education, Social Networks, eScience/Simulation, and GIS
Edward A. Fox, Jonathan P. Leidig

Information and Human Values
Kenneth R. Fleischmann

Multiculturalism and Information and Communication Technology
Pnina Fichman and Madelyn R. Sanfilippo

Transforming Technologies to Manage Our Information: The Future of Personal Information Management, Part II
William Jones

relevance

noun rel·e·vance \ˈre-lə-vən(t)s

Definition of *relevance*

Popularity: Top 30% of words

1. (a) relation to the matter at hand (b) practical and especially social applicability: pertinence <*giving relevance to college courses*>
2. the ability (as of an information retrieval system) to retrieve material that satisfies the needs of the user

Related to relevance:

Synonyms

applicability, bearing, connection, materiality, pertinence, relevancy

First known use of relevance: 1733

Meriam-Webster Dictionary

http://www.merriam-webster.com/dictionary/relevance

The Notion of Relevance in Information Science:
Everybody knows what relevance is. But, what is it really?
Tefko Saracevic
www.morganclaypool.com

ISBN: 9781598297683 print
ISBN: 9781598297690 ebook

DOI 10.2200/S00723ED1V01Y201607ICR050

A Publication in the Morgan & Claypool Publishers series
SYNTHESIS LECTURES ON INFORMATION CONCEPTS, RETRIEVAL, AND SERVICES, #50
Series Editor: Gary Marchionini, University of North Carolina, Chapel Hill

Series ISSN 1947-945X Print 1947-9468 Electronic

The Notion of Relevance in Information Science

Everybody knows what relevance is.

But, what is it really?

Tefko Saracevic
Rutgers University

SYNTHESIS LECTURES ON INFORMATION CONCEPTS, RETRIEVAL, AND SERVICES #50

MORGAN & CLAYPOOL PUBLISHERS

ABSTRACT

Everybody knows what relevance is. It is a "y'know" notion, concept, idea—no need to explain whatsoever. Searching for relevant information using information technology (IT) became a ubiquitous activity in contemporary information society. *Relevant* information means information that pertains to the matter or problem at hand—it is directly connected with effective communication. The purpose of this book is to trace the evolution and with it the history of thinking and research on relevance in information science and related fields from the human point of view. The objective is to synthesize what we have learned about relevance in several decades of investigation about the notion in information science. This book deals with how people deal with relevance—it does *not* cover how systems deal with relevance; it does *not* deal with algorithms. Spurred by advances in information retrieval (IR) and information systems of various kinds in handling of relevance, a number of basic questions are raised: *But what is relevance to start with? What are some of its properties and manifestations? How do people treat relevance? What affects relevance assessments? What are the effects of inconsistent human relevance judgments on tests of relative performance of different IR algorithms or approaches?* These general questions are discussed in detail.

KEYWORDS

relevance, information retrieval (IR), relevance behavior, relevance effects

Contents

Figures

Acknowledgments

Under its search box, Google Scholar has a cryptic command: "*Stand on the shoulders of giants.*" A few centuries ago, Isaac Newton, referring to Galileo and Kepler, said it better:

> "*If I have seen further [than certain other men] it is by standing upon the shoulders of giants*" (letter to Robert Hooke, February 5, 1675).

And a few centuries before that, in the 12th century, Bernard of Chartres (as reported by John of Salisbury) said it even better:

> "*We are like dwarfs sitting on the shoulders of giants; we see more things and more distant things than did they, not because our sight is keener nor because we are taller than they, but because they lift us up and add their giant stature to our height*" (Metalogicon, III, 4).

In that spirit, I owe a debt of gratitude to all authors included in the references. They did the hard work. I stood on their shoulders and saw further.

My heartfelt thanks to Gary Marchionini, the Editor of the Synthesis Lectures on Information Concepts, Retrieval, and Services. He suggested this work to me some years ago, and then over the years prodded the process gently toward completion, and made helpful comments along the way. It is my fault that I am late in producing it.

Special thanks to Diane Cerra, Executive Editor, Morgan & Claypool Publishers—she was not only encouraging when encouragement was needed, but also thoroughly and joyfully professional when professional help and guidance was required.

Thanks to Yuelin Li and Ying Zhang, my assistants at the time when this work begun, who tirelessly searched the literature for sources about relevance and then organized them. My colleagues at the School of Communication and Information, Rutgers University, provided a most hospitable framework for scholarly work and musings, and my students were a source of constant challenge, not only when I was there, but also now that I am retired as Distinguished Professor Emeritus. I am not only grateful, but also happy and proud to have worked and that I am still working in that environment.

From lifelong study of relevance, I learned a lot. I also learned how little I know.

Preface

Everybody knows what relevance is. *But, what is it really?* That is the subject of this book. Actually, the book is a treatise about thinking and research on the subject of relevance in the field of information science and related fields.

Relevance is an intensely human notion. It is a people thing. Relevant information is a human assessment, a judgment, a drawing of connection between given information and a given context or problem at hand.

One more thing: like all human notions, relevance is messy. The book tries to make our understanding of relevance less messy.

Contemporary information technology tries to assist people in obtaining relevant information. A never-ending succession of systems, processes, algorithms, and displays are being devised to do so. Every one of them is assuming something about relevance—what makes relevant information relevant—and it goes from there—actually, their coverage and algorithms go from there.

Yet, what is "relevance," actually? Frankly, we do not know. We do not know it any more than we know what is information, or energy, or matter, or gravity, or life, or any number of other basic notions and phenomena. Scientist dealing with these will freely admit the proposition of "*we do not know*." Instead, of trying to answer the question directly they are busy investigating the nature, manifestations, behavior, and effects of these phenomena and notions. Newton, Galileo, Einstein did exactly that in physics, as did Watson and Crick in molecular biology.

Spurred by striking advances in applications of information technology to handling of information in general and of relevant information in particular, relevance itself became a subject of scientific investigation. *What is the nature of relevance? What are its manifestations? Behavior? Effects?*

There are two reasons for addressing these questions. Both are time honored. One is knowledge for knowledge's sake: to better understand relevance. The other is pragmatic: we assume that the more we know and understand about relevance the better we can build systems, algorithms, and processes aimed at retrieving relevant information.

From its start shortly after the end of the Second World War, one of the main orientations of information science is toward developing and deploying ways and means for retrieval of relevant information by using advances in information technology. Not surprisingly then, investigation of the notion of relevance became a research topic on its own.

Note that the book is in large part based on previous studies of relevance by the author:

Saracevic, T. (1975). Relevance: A Review of and a framework for the thinking on the notion in information science. *Journal of the American Society for Information Science*, 26(6), 321–343.

Saracevic, T. (2007a). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 58(3), 1915–1933.

Saracevic, T. (2007b). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology*, 58(13), 2126–2144.

Saracevic, T. (2008). Effects of inconsistent relevance judgments on information retrieval test results: A historical perspective. *Library Trends*, 56(4), 763–783.

Saracevic, T. (2016). Relevance: In search of a theoretical foundation. In: Sonnenwald, D. (ed.), *Theory Development in the Information Sciences*. Austin: University of Texas Press.

Tefko Saracevic
July 2016

CHAPTER 1

Introduction

Everybody searches. Everybody googles. In doing so, everybody is looking for something that is relevant to whatever current question, state, problem. And everybody knows what relevance is. It is a "y'know" notion, concept, idea—no need to explain whatsoever. This is true from one end of the globe to another. And this is the basic, underlying reason why search engines and other systems with similar goals are so wildly successful throughout the globe. They need no explanation as to what they are all about. They are about relevance.

Searching for relevant information using information technology became a ubiquitous activity in contemporary information society. *Relevant* information means information that pertains to the matter or problem at hand—it is directly connected with effective communication. A significant economic component is involved in these activities. Searching is not only ubiquitous; it became big business as well. Relevance is involved in all that.

The basic objective of all information retrieval (IR) systems, including contemporary search engines and abstracting and citation systems of all kinds, is to provide relevant information to users in response to their queries or profiles and (hopefully) information needs. Garcia-Molina et al. (2011) refer to them under a generic name "*information providing mechanisms*" and conclude that:

> "… *one of the fundamental problems in computer science has become even more critical today: how to identify objects satisfying a user's information need. The goal is to present to the user only information that is of interest and relevance, at the right place and time*." (ibid., 21)

Queries and profiles could be explicit or implicit; retrieval methods and algorithms could vary widely; displays could differ; coverage could fluctuate in breadth and depth … but no matter what the operation, still the basic intent is to present users not with any old kind of information, but with relevant information. The basic goal of all research and development (R&D) in IR systems (or information-providing mechanisms) in computer science is improving the relevance of the results to the query. In that sense, relevance is the basic notion underlying any and all IR systems. It underlies Google, Baidu, and all other search engines; Amazon and other systems with recommendations; advertising systems geared toward individuals or identified groups; Web of Science, Scopus, and all indexing and citation databases; and searching the content of complex websites, such as the Library of Congress. Relevance is "the invisible hand" that governs these systems.

The notion of the "invisible hand" is taken from Adam Smith (1723–1790), regarded as the father of economics. In his classic 1776 book *An Inquiry into the Nature and Causes of the Wealth* *of Nations* Smith set out the mechanism by which he believed economic society operated; among others, he explained market decisions as often being governed by an "invisible hand." In the same spirit, while the hand of relevance is invisible, it is governing.

1.1 INFORMATION TECHNOLOGY AND RELEVANCE

In human history, relevance has been around forever, or as long as humans tried to communicate and use information effectively. Modern computers have been around since the 1950s, the Internet started its rise in the 1970s, the Web in the 1990s, and social media in the 2000s. In this short time, contemporary information technology (IT), which includes modern communication technology, social media, and information systems based on IT, changed and transformed a great many things in society—from education to health care, from earning a living to leisure, from physics to classics, from government to being governed, from being young to being old. IT changed information activities dramatically, namely the way we acquire, organize, store, preserve, search, retrieve, communicate, interact with, and use information. In all of those information activities relevance plays a most significant, underlying, and yet elusive role.

> A note on information objects, i.e., entities potentially conveying information: They can take many forms, such as documents (historically first treated in IR and still in great many cases), but also extended now to facts, music, spoken words, still and moving images, artifacts, and other forms of human expression.

When people use IT for information activities, particularly searching, relevance is the primary underlying reason. As far as people are concerned, relevance is tacitly present and inescapable. Correspondingly, relevance plays a significant, underlying role when these activities are performed by IR systems (or information-providing mechanisms) as well; because these systems are designed primarily to respond with information or information objects that are potentially relevant to people. In this lies the significance of relevance.

The reason we position people and IT together in the discussion above is to point out some basic premises, distinctions, problems, and conflicts. IR systems use a variety of methods, algorithms, and technologies aimed at organization of information or information objects and subsequently aimed at searching for and retrieval of information (or information objects) that may be potentially relevant to users; for shorthand, we label this "*systems relevance*." People go about their ways and assess their own version of relevance; we label this "*human relevance*." Both treat relevance as a relation (as explained more fully in Section 3.3). However, each may have different premises for what is involved in the relation and in assessing that relation. There are two interacting worlds—the systems world and the human world—and two basic categories of relevance—systems' and humans.' The two worlds interact with various degrees of problems and conflict, from none to a lot.

Our concern here is with the human world of relevance. Relevance is treated here as a human condition, which it is. While we can never get far from systems and associated technology, this book does *not* cover how systems deal with relevance. It does *not* deal with algorithms. Treatments of relevance in information retrieval—in algorithms, measures, evaluation—are beyond the scope of this book. There are plenty of books, articles, proceeding papers, lectures, and musings that deal with these topics. However, relevance is at the heart of such systems and algorithms. All are geared toward relevance.

1.2 PURPOSE, OBJECTIVES, ORGANIZATION

Spurred by advances in information retrieval (IR) and information systems of various kinds in handling of relevance, a number of basic questions were raised:

> *But what is relevance to start with? What are some of its properties and manifestations?*
> *How do people treat relevance? What affects relevance assessments?*

Following these basic questions, over the past half a century or so the nature, manifestation, behavior, effects, and problems related to human relevance were investigated in a number of fields, but primarily in information science and to some degree in computer science. Relevance research is multidisciplinary. The purpose here is to trace the evolution and with it the history of thinking and research on relevance in information science and related fields from the human point of view. To repeat: The book does not deal with the technological and algorithmic point of view, namely, it does *not* address how systems deal with relevance as related to information or information objects.

The objectives are to present the results of this thinking and research in a way that may be useful not only for enhancing a general understanding of the notion of relevance but also, hopefully, for attempts to incorporate some of that understanding in design and operations of information retrieval systems, search engines included (more about that in Section 9.3). The history of science and technology is full of instances where a better understanding of the basic notions or phenomena underlying a technology led to development of more effective and successful technologies and systems. A fruitful, though sometimes convoluted and arduous, translation was realized. I hope that a better understanding of relevance may lead to better information systems. This clearly illustrates the significance of relevance research. Considering and understanding relevance as a notion is still relevant, if not even more so, to building and operating information systems—now ever more complex in the Web environment—that effectively provide information to users pertaining to their problems at hand.

This book evolved also in part from a series of relevance lectures: individual chapters can be referred to and used as a series of individual lectures covering various aspects of thinking, research, and experimental findings on relevance.

In addition to the Introduction (Chapter 1), discussing the relation between contemporary developments in information technology and relevance, the book as a whole is further organized in eight chapters covering different topics.

- Chapter 2 provides a *historical overview* with concentration how information retrieval was linked to relevance.

- Chapter 3 focuses on general and specific understandings about *manifestations and attributes of relevance*.

- Chapter 4 reviews various proposed *models of relevance* and describes some of them in detail.

- Chapter 5 is devoted to *theories of relevance* in several fields and examines their application in information science.

- Chapter 6 enumerates experimental studies that provided observation and data on *behavior of relevance* concentrating on determination of what makes information or information objects relevant.

- Chapter 7 enumerates experimental studies that provided observation and data on *effects of relevance* concentrating on relevance judges and judgments and their effects on results of IR tests.

- Chapter 8 enumerates experimental studies that analyzed effects of *inconsistent relevance judgments* on IR test results.

- Finally, Chapter 9 provides some general *conclusions* about relevance research and trends.

Each chapter, including this one, concludes with a synthesis that in effect provides a digest and interpretation of contemporary thinking on the topic treated or suggests hypotheses for future research.

1.3 SYNTHESIS: BASICS ABOUT RELEVANCE

Here are some basic issues and realities about relevance:

◊ Searching for relevant information using technology became a ubiquitous activity in contemporary information society.

◊ As far as people are concerned relevance is a "y'know" notion, concept, idea—no need to explain or define it whatsoever. This is true from one end of the globe to another.

◊ The basic objective of all information retrieval (IR) systems, including contemporary search engines and information retrieval systems of all kinds, is to provide relevant information to users in response to their queries or profiles and, hopefully, needs.

◊ There are two interacting worlds—the systems world and the human world—and two basic categories of relevance—systems' and humans.' The two worlds interact with various degrees of problems and conflict, from none to a lot.

◊ Relevance is "the invisible hand" that governs all kinds of information retrieval systems and information-providing mechanisms.

CHAPTER 2

A Bit of History

"Those who don't know history are doomed to repeat it."
Edmund Burke (1729–1797)

2.1 INFORMATION SCIENCE

Debate and research about relevance have a pragmatic foundation. Historically, contemporary concerns with relevance as a notion stem from operational and down-to-earth problems related to retrieval of relevant information in information science.

Dictionaries and encyclopedias routinely define information science[1] as the science dealing with the effective collection, storage, and retrieval of information. More specifically, information science is a field of professional practice and scientific inquiry addressing the problem of effective communication of knowledge records—no matter the form: print, electronic, image, sound—among humans in the context of the social, organizational, and individual need for and use of information (Saracevic, 1999). The key orientation here is the problem of need for and use of information, as involving knowledge records of all kinds in all media. To provide for that need, information science deals with specifically oriented information techniques, procedures, systems, and technologies. However, for understanding of information science, as of any other field, of interest is not only a lexical definition but also (and even more so) the description of problems addressed. A field is defined by problems addressed.

Information science is a field that emerged in the aftermath of the Second World War, along with a number of new fields, with computer science being but one example. The rapid pace of scientific and technical advances that were accumulating since the start of the 20th century produced by midcentury a scientific and technical revolution. A most visible manifestation of this revolution was the phenomenon of "information explosion," referring to the unabated, exponential growth of scientific and technical publications and information records of all kinds. The term "information explosion" is a metaphor, as is "population explosion," because nothing really exploded but just grew at a high rate, if not even exponentially. The phenomenon of information explosion is continuing and even accelerating to this day, particularly in the digital and Web environments. Addressing the problems of dealing with information explosion is at the heart of information science.

[1] https://en.wikipedia.org/wiki/Information_science

The impetus for the development of information science, and even for its very origin and agenda, can be traced to a 1945 article, "As we may think,"[2] by Vannevar Bush (1890–1974),[3] a respected MIT scientist and, even more importantly, the head of the U.S. scientific effort during WWII (Bush, 1945). In this influential article, Bush did two things: (a) he succinctly defined a critical and strategic problem that was on the minds of many, and (b) proposed a solution that was a "technological fix," and thus in tune with the spirit of time. Both had wide appeal and Bush was listened to because of his stature. He defined the problem in almost poetic terms as:

> "*The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships*" (ibid. Section 1).

In other words, Bush addressed the problem of information explosion. The problem is still with us, even growing.

His solution was to use the emerging computing and other information technology to combat the problem. However, he went even further. He proposed a machine named Memex,[4] incorporating (in his words) a capability for "*association of ideas*," and duplication of "*mental processes artificially*." A prescient anticipation of information science and artificial intelligence is evident. Memex was never built, but to this day is an ideal and, some think, a utopia. We are still challenged by the ever-worsening problem of information explosion, now universal and in a variety of digital formats. We are still trying to fix things technologically. We are still aspiring to incorporate Memex's basic ideas in our solutions of finding what Bush called the "*momentarily important item*," meaning a relevant item.

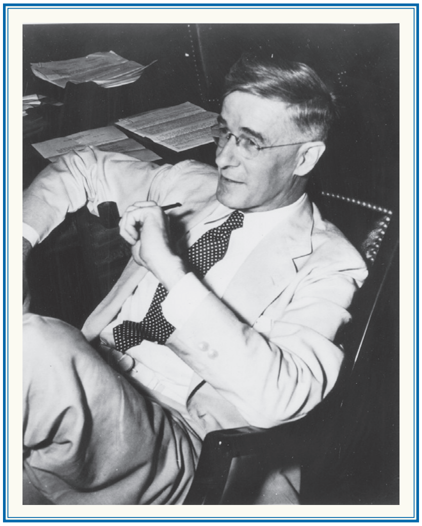

Figure 2.1: Vannevar Bush (1890–1974). Courtesy MIT Museum.

A number of scientists and professionals in many fields around the globe listened and took up Bush's challenge. Governments and politicians listened as well and provided funding. The reasoning went something like this: Because science and technology are strategically important for society, efforts that help them, information activities in particular, are also important and need

2 http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/
3 http://www.livinginternet.com/i/ii_bush.htm
4 https://en.wikipedia.org/wiki/Memex

support. In the U.S., this led to support of research and development in information retrieval and to some extent relevance research as well, mostly through the National Science Foundation (NSF).

Talking about NSF support: NSF led a multi-agency funding under the same mentioned mandate run by Division of Information & Intelligent systems under Michael Lesk entitled Digital Library Initiatives (DLI); DLI 1 ran 1994–1998; DLI 2 1999–2004. In 1994, the first six awards were made. One of them supported a Stanford University project led by Hector Garcia-Molina and Terry Winograd. Within that project Garcia-Molina's students Larry Page and Sergei Brin collaborated on a system named BackRub that became Google. In the original paper about Google, Brin and Page (1998) acknowledged support from NSFs Digital Library Initiative for development of Google (note that there are various versions of this paper, some have the acknowledgment, others don't). NSF has a story "On the Origins of Google"[5] on its site. There is a continuum from Bush to Google.

> Bush was also instrumental in establishing the National Science Foundation (NSF) in the U.S.
>
> The National Science Foundation Act of 1950 (P.L. 81–507) provided a number of mandates, among them "to foster the interchange of scientific information among scientists in the U.S. and foreign countries" (Section 3(a)3) and "to further the full dissemination of [scientific and technical] information of scientific value consistent with the national interest" (Section 11(g)). The 1958 National Defense Education Act (P.L 85–864) (the "Sputnik act") enlarged the mandate: "The National Science Foundation shall [among others] undertake program to develop new or improved methods, including mechanized systems, for making scientific information available" (Title IX, Section 901). By those mandates, an NSF division, which after a number of name and direction changes is now called the Division of Information and Intelligent Systems (IIS), has supported research in these areas since the 1950s.

2.2 INFORMATION RETRIEVAL (IR)

Right after the Second World War a variety of projects started applying a variety of technologies to the problem of controlling information explosion, particularly in science and technology. In the beginning, the technologies were punched cards and microfilm, but soon after computers became available the technology shifted to and stayed with computers. Originally, they begun and evolved within information science and specific fields of application, such as chemistry. By the mid-1960s computer science joined the efforts in a big way.

Various names were applied to these efforts, such as "machine literature searching," or "mechanical organization of knowledge" but by the mid-1950s "information retrieval" prevailed.

[5] http://www.nsf.gov/discoveries/disc_summ.jsp?cntn_id=100660&org=IIS

Actually, the term "information retrieval" (IR) was coined by mathematician and physicist Calvin N. Mooers,[6] a computing and IR pioneer, just as the activity started to expand from its beginnings after World War II. He posited that:

> "*Information retrieval is … the finding or discovery process with respect to stored information … useful to him [user]. Information retrieval embraces the intellectual aspects of the description of information and its specification for search, and also whatever systems, technique, or machines that are employed to carry out the operation*" (Mooers, 1951, p. 25).

Over the next half century, information retrieval[7] (IR) evolved and expanded widely. In the beginning, IR was static; now it is highly interactive. It dealt only with representations—indexes, abstracts—now it deals with full texts as well. It previously concentrated on print only; now it covers every medium, and so on. Advances are impressive. However, in a basic sense, IR continues to concentrate on the same fundamental things Mooers described. Searching was and still is about retrieval of relevant information or information objects.

It is of interest to note what made IR different from many other techniques applied for control of information records over a long, historical period. The key difference between IR and related methods and systems that long preceded it, such as classifications, subject headings, various indexing methods, or bibliographic descriptions, including the contemporary Functional Requirements for Bibliographic Records (FRBR) (IFLA, 1998), is that IR specifically included "specification for search." The others, including FRBR, did not. In these long standing techniques what user needs are and should be fulfilled were specified (rather briefly), but how the search will be done was neither specified, nor mentioned at all. Data in bibliographic records were then organized in a way to fulfill the specified needs. Searching was assumed and left to itself—it just happens. In IR, the user's needs are assumed as well, but the search process is specified in algorithmic detail, and data is organized to enable the search. Search engines are about searching to start with; everything else is subsumed to that function.

The fundamental notion used in bibliographic description and in all types of classifications or categorizations, including those used in contemporary databases, is aboutness. The fundamental notion used in IR is *relevance*. As mentioned, it is not about any kind of information, and there are a great many, but about *relevant* information. Fundamentally, bibliographic description and classification concentrate on describing and categorizing information objects; IR is also about that but, *and this is a very important "but*," in addition IR is about searching as well, and searching is about relevance. Very often, the differences between databases and IR were discussed in terms of differences between structured and unstructured data, which is OK, but the fundamental difference

[6] http://www.nsf.gov/discoveries/disc_summ.jsp?cntn_id=100660&org=IIS
[7] https://en.wikipedia.org/wiki/Information_retrieval

is in the basic notion used: *aboutness* in the former and *relevance* in the latter. Therein similarity and difference lie. *Relevance entered as a basic notion through the specific concentration on searching*.

By choosing relevance as a basic, underlying notion of IR, related information systems, services, and activities—and with it, the whole field of information science—went in a direction that differed from approaches taken in librarianship, documentation, and related information services, and even in expert systems and contemporary databases in computer science.

For example, the basis of expert systems is uncertainty (or rather reduction of uncertainty based on if–then rules). As a result, expert systems are very different from IR systems. In comparison to IR, expert systems are not as widely adapted and used; actually, they disappeared; presently the designation "expert systems" is used rarely or not any more at all. One of the reasons may be due to the choice of the underlying notion. Relevance is a human notion, widely understood in similar ways from one end of the globe to the other. Uncertainty is not. Besides, the assumption that information decreases uncertainty does not hold universally; information may also increase uncertainty.

2.3 TESTING OF IR SYSTEMS AND FIRST CONCERNS WITH RELEVANCE

Historically, relevance actually crept in unannounced. At the start of IR in the 1950s, nobody actually made a big point about it. IR systems were constructed to do relevance, but nobody talked about it. Still, principles posited then are valid to this day. It was, and still is, accepted that the main objective of IR systems is to retrieve information relevant to user queries, and possibly needs, or as Mooers observed in the above definition, "useful to user."

Actually, the first discussions directly involving relevance in the early 1950s were not about relevance, but about non-relevance, that is about "false drops" or noise—unwanted information retrieved by IR systems. The concerns were about getting too much non-relevant information; subsequently efforts were directed toward development of methods for decreasing of false drops. In other words, the concerns over relevance started with the large amount of non-relevance that endangered effective communication. The problem is still with us.

In mid the 1950s Allen Kent (1922–2014) and James W. Perry (1907–1971), both chemists and pioneers in information science, wrote a series of articles about techniques of IR. In one of the articles, they suggested measures for evaluating performance of IR systems. They called them "precision" and "relevance" (later because of confusion renamed "recall") (Kent et al., 1955). This was the first full recognition of relevance as an underlying notion of retrieval—relevance was the criterion for these measures. Precision and recall measure the probability of agreement between what the system retrieved/did not retrieve as relevant (systems relevance) on the one hand and what the user assessed as relevant (user relevance) on the other hand, with user relevance being the base for deriving probabilities of agreement. *User relevance was and still is treated as the gold standard for* *performance evaluation*. Over time, many other measures were suggested, but did not take. Precision and recall remained standard measures of effectiveness to this day, with some variations on the theme. Relevance became and remained the underlying criterion for measuring the effectiveness of IR. By now, it is cemented there.

Very soon, after IR systems appeared, the perennial questions asked of all systems were raised: *What is the effectiveness and performance of given IR approaches? How do they compare?* It is not surprising that these questions were raised in IR. At the time, most developers, funders, and users associated with IR were engineers, scientists, or worked in related areas where the question of testing was natural, even obligatory.

The first attempt at IR testing was reported by Gull (1956); relevance judgments were critical in collapse of that test (since the aftermath of results affected IR testing to this date, the experiment is described in some detail in the sidebar). Relevance came to the fore. Ever since all who were engaged with IR testing or their critiques became very sensitive about relevance and relevance judgments. In other words, what is used today as to determining relevance in IR tests has its roots in a collapse of a study in the early 1950s. IR testing continued to this day. Precision and recall (with a number of variations) remained standard measures of effectiveness. A lesson from the first test was learned, although today hardly anybody knows its source. All IR tests include a single judge (or sometimes a group with a consensus) that provide a golden rod of relevance of documents (information sources) against which a system or algorithm performance is assessed.

Gull (1956) reported on an attempt to test the performance of two competing IR systems developed by separate groups: one developed by the Armed Services Technical Information Agency (ASTIA) using subject headings, and the other by a company named Documentation Inc., using uniterms (keywords searched in a Boolean manner). Each group searched 98 requests using the same 15,000 documents, indexed separately, in order to evaluate performance based on relevance of retrieved documents. *However, each group judged relevance separately*. Then, not the systems' performance, but their relevance judgments became contentious. The first group found that 2,200 documents were relevant to the 98 requests, while the second found that 1,998 were relevant. There was not much overlap between groups. The first group judged 1,640 documents relevant that the second did not, and the second group judged 980 relevant that the first did not. They tried to reconcile, considering each other's relevant documents, and again comparing judgments. At the end, they still disagreed; their rate of agreement, even after peace talks, was 30.9%. That did it. The first ever IR evaluation did not continue. It collapsed. The lesson was learned: *Never, ever use more than a single judge (or a single object, such as source document) for establishing the gold standard for comparison. No IR test ever does.*

IR testing began in the late 1950s within a certain context as described by Cyril Cleverdon[8] (1914–1997), an IR testing pioneer, in his acceptance speech for the 1991 Gerard Salton Award given by the Special Interest Group on Information Retrieval (SIGIR), Association for Computing Machinery (ACM):

> "*These new techniques [uniterms, coordinate indexing] generated considerable argument, not only between the proponents of the different systems, but also among the library establishment, many of whom saw these new methods as degrading their professional mystiques … Most of this could be ignored, such as the comment,* 'You had no right to be so intelligent with the Uniterm system; it is meant to be used by persons of low intellect.'*… Controversy over the new methods was still raging, with extravagant claims on one side being countered by absurd arguments on the other side, without any firm data being available to justify either viewpoint* (Cleverdon, 1991, pp. 3, 5)."

Testing continued with pioneering and even legendary tests at the Cranfield Institute of Technology in the UK (thus "Cranfield tests"). There were two periods for these tests: Cranfield 1 ran from 1957 to 1961 and Cranfield 2 from 1963 to 1966 (Cleverdon, 1967; 1991). Reports from various Cranfield tests are reproduced in the SIGIR Museum.[9] The most important result from Cranfield tests was the finding that complex, controlled, classificatory indexing schemes were not necessary for effective retrieval; keywords worked as well. This justified pioneering development of algorithm-based IR by Gerard ("Gerry") Salton (1927–1995) and his SMART system with numerous experiments (Salton, 1969).

Among others, Cranfield tests established the traditional model of IR (presented in Figure 4.1) and an associated test methodology still in use today in the majority of IR tests. IR testing culminated with Text Retrieval Conference (TREC)[10] established in 1992 and continuing to this day.

IR tests are about comparing effectiveness of different retrieval methods. (Excellent historical syntheses of various IR tests from the Cranfield tests to TREC are in Robertson (2008) and in a Morgan & Claypool book by Harman, 2011). All these tests used relevance as the basic criterion and user relevance judgments as the gold standard for comparison. However, they also had an important unintended consequence. Relevance assessments used in early tests became contentious; it quickly became clear that relevance itself is a problem. Relevance assessments were revealed, among others, not

> TREC is a long-term, ongoing effort at the (U.S.) National Institute for Standards and Technology (NIST) that brings various IR teams around the world together annually to compare results from different IR approaches under laboratory conditions (Vorhees and Harman, 2005).

8 https://en.wikipedia.org/wiki/Cyril_Cleverdon
9 http://sigir.org/resources/museum/
10 http://trec.nist.gov/