How many relevances in information retrieval?

Stefano Mizzaro[a],*

[a]*Department of Mathematics and Computer Science, University of Udine, Via delle Scienze, 206
Loc.Rizzi-33100, Udine, Italy*

Abstract

The aim of an information retrieval system is to find relevant documents, thus relevance is a (if not 'the') central concept of information retrieval. Notwithstanding its importance, and the huge amount of research on this topic in the past, relevance is not yet a well understood concept, also because of inconsistently used terminology. In this paper, I try to clarify this issue, classifying the various kinds of relevance. I show that: (i) there are many kinds of relevance, not just one; (ii) these kinds can be classified in a formally defined four dimensional space, and (iii) such classification helps us to understand the nature of relevance and relevance judgement. Finally, the consequences of this classification on the design and evaluation of information retrieval systems are analysed. © 1998 Elsevier Science B.V.

Keywords: Information retrieval; Relevance; Kinds of relevance; Relevance judgement; System design; System evaluation

1. Introduction

Relevance is a crucial concept of information retrieval (IR) [1–3] as the aim of an IR system is to find relevant documents. Many researchers studied this issue; the main events (for an extensive survey see Ref. [4]) were probably:

- Vickery [5,6] presented at the 1958 ICSI debate a distinction between 'relevance to a subject' which refers to what the IR system judges 'relevant', and 'user relevance' refering to what the user needs.
- Rees and Schultz [7] experimentally studied the effect of different scaling techniques on the reliability of judgements. They noted that relevance judgements are inconsistent and affected by about 40 variables.

---

* Corresponding author. Tel: +39 (432) 55.8456; Fax: +39 (432) 55.8499; e-mail: mizzaro@dimi.uniud.it; WWW: http://www.dimi.uniud.it/~mizzaro