

Demand-oriented Information Supply of Digital Content

© Kurt Sandkuhl

School of Engineering at Jönköping University, Sweden

Information overload has been a phenomenon observed and discussed in the literature since many decades. One of the pioneers of computer-supported collaborative work, Vannevar Bush, foresaw already in 1945 that it would not be possible to manage all information we collect in our “bewildering store of knowledge”. However, the attention in the scientific community in this field increased significantly during the last 10 years with the intensification of Internet, e-mail, and information systems use. The problem no longer seems to be that information does not exist electronically, but that it is difficult to find in the huge amount of available data.

The research field information logistics addresses the problem of information overload by developing concepts and technologies for improving information flow in organizations. The core idea is to use principles from material logistics, like just-in-time delivery, in the area of information supply for improved information supply. This is based on demands with respect to content, time of delivery, location, presentation, and quality of information. The scope can be a single person, a target group, a machine/facility, or any kind of networked organization. The aim is to explore, develop, and implement concepts, methods, technologies, and solutions for the above mentioned purpose.

A core subject of demand oriented information supply is how to capture the needs and preferences of a user in order to get a fairly complete picture of the demand in question. Among the different approaches for this purpose are user profiles, situation-based and context-based demand models. *User profiles* are usually created for functionality provided by a specific application. They are based on a predefined structured set of personalization attributes and assigned default values at creation time. The basic idea of the *situation-based* approach is to divide the daily schedule of a person into situations and to determine the optimal situation for transferring a specific message. This approach defines a situation as an activity in a specific time interval including topics and location relevant for the activity. Information value is a relation between a message and a situation, which is based on relevance of the topics of a message for the situation, utility of the message in specific situations and acceptance by the user. The *context-based* is based

on the idea that information demand of a person in an enterprise to a large extent depends on the work processes this person is involved in, on the co-workers or superiors and on the products, services or machines the person is responsible for. This led to the proposal to capture the context of information demand, i.e. a formalized representation of the setting in which information demand exists.

Digital content has been subject of research since several decades. An established since the 1980's is the differentiation between logical structure, layout structure and meta-data for describing content, presentation or both. A prominent example for meta-data is the Dublin Core Standard. This principal structure is valid even for contemporary document management applications and is complemented by navigation structure or time-related information. For storing meta-data two principal ways can be distinguished:

- Embedding meta-data in the document, i.e. meta-data are part of the document model,
- Managing meta-data separated from the document in the management systems.

Most contemporary content management solutions use a hybrid approach, as file formats contain a core set of meta-data for specific purposes, which are complemented with additional information on system side.

When trying to identify digital content that fits to a given demand model, matching between demand model and meta-data is considered an important element. Matching is often characterized as “searching with imprecise specification of information needs”. In this context we have to discuss the aspect of relevance of information, e.g. algorithmic, topical and cognitive relevance. The underlying concepts for algorithmic relevance, i.e. the relation between the query features and the search result, and for topical relevance, i.e. relation between aboutness of content objects and query, have to be observed when implementing matching systems. Semantic matching based on ontologies in combination with information retrieval techniques for string matching are promising approaches for information logistics in order to match information demand and content.