The Information Market;

Its basic concepts, and its challenges

P. van Bommel, B. van Gils, H.A. Proper, M. van Vliet and Th.P. van der Weide

Department of Information & Knowledge Systems (IRIS), Sub-faculty of Mathematics & Informatics, Radboud University

Toernooiveld 1, 6525 ED Nijmegen, The Netherlands, EU.

{P.vanBommel,B.vanGils,Th.P.vanderWeide,M.vanVliet,E.Proper}@cs.ru.nl

Abstract

This paper discusses the concept of information market. The authors of this paper have been involved in several aspects of information retrieval research. In continuing this research tradition, we now take a wider perspective on this field, and position it as a market where demand for information meets supply for information. This paper focusses on the concept of information market.

We will start by exploring the notion of a market in general. Our considerations are inspired by economic markets, in particular micro economy. This is followed by a specialisation of these considerations to the information market, where we will also position some of the existing work.

KEYWORDS: Information Market, Information Retrieval, World-Wide-Web.

1 Introduction

Our modern day western societies are dominated by information systems. Nevertheless, already in cultures and empires long gone, information systems played an important role. The Egyptians, the Greeks as well as the Romains, already used forms of (manual!) information systems to administer trade and affairs of state. The industrial age resulted in an explosion of the amounts of information that needed to be handled. When the invention of the transistor gave us computers, these "information processing machines" were gladly accepted as a means to automate part of the information processing required.

Hand in hand with the increase of the amounts of information that need processing, the problem of *information overload* started to surface. As more and more data were amassed in information systems, it became harder and harder to find those bits of data that really mattered, i.e. that are realy of *information* to someone. This has led to the introduction of the field of information retrieval [41, 39].

The development of the Internet provided our society with the opportunity to interconnect computers, leading to networked information systems. When the Internet matured, it gave birth to the World-Wide-Web (the Web). This resulted in a multiplication of the information available to people around the globe. Information resources that are available on the Web include: Web pages, newsgroups, mailing-list archives, networked databases, applications, business services, as well as indexing services. For users of the Web, these resources are at their disposal for doing business, search other information, educational purposes, or relaxation. Since the Web literally spans the world, the number of accessible information resources is astronomical. This makes life rather difficult for the average user who shops around to discover information resources that fulfill his or her given information need. These developments have shifted the attention of information retrieval research away from "stand alone" collections to information retrieval on the Web [15, 18]

When the World-Wide-Web matured it, on its turn, gave birth to e-commerce. Given the abundance of information available via the Web, an important part of the commodities traded on the Internet are actually "carriers" of information. This brings us to the focus of this paper. This paper proposes to look at the exchange of information on the Internet as an *Information Market*, where demand and supply of information meet. As such, this paper aims to mark a transition from a traditional view on information retrieval to an *Information Market Paradigm*.

Figure 1 represents what might be called a traditional perspective on information retrieval; the *Information Retrieval Paradigm* [13, 14]. On one side (the right hand side), there are information resources that are at our disposal. These resources, which may be aggregated, are characterised in some way to facilitate their discovery. Facing the information carriers is the user with an information need. The user expresses this need in terms of an information request; a query. The query will usually only be a crude description of the actual carrier(s) needed to fulfill the given information need. The need for information can be caused by a number of reasons, but will usually be due to some *gap* in the user's knowledge. For instance, the user needs to know something in order to complete a task. Relevant information is discovered and then absorbed by the user to fill the knowledge gap. A knowledge gap may also arise out of idle curiosity. For example, some users of the Internet begin surfing the internet with no specific goal and then encounter some topic that engages their curiosity in the sense that they want to learn more about it.

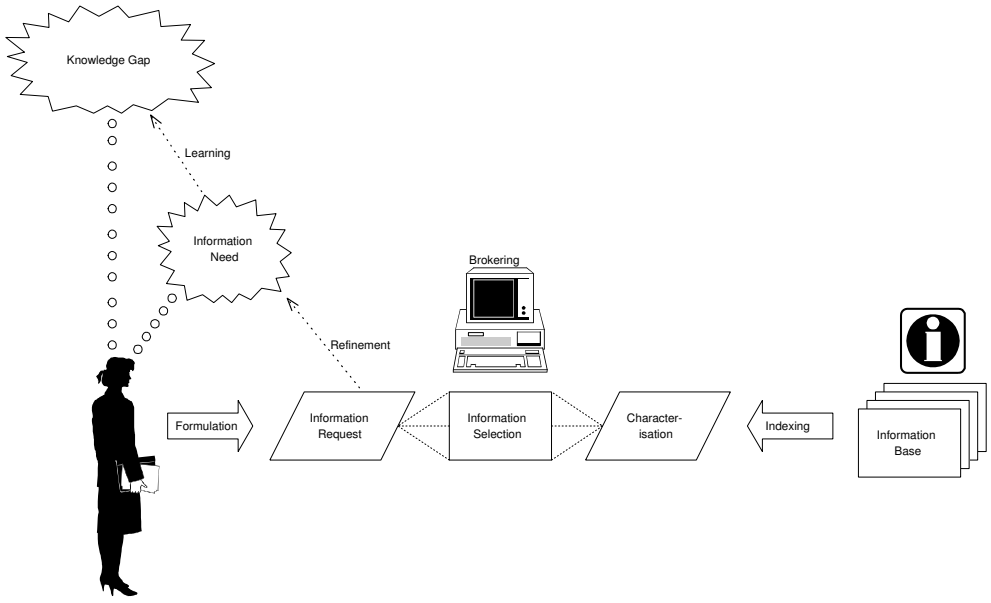

Figure 1: Information Retrieval Paradigm

Typical research challenges that follow from this paradigm are:

Formulating needs – The formulation of information requests involves two important issues. First of all, it requires some formal language in which to express the query. Secondly, a precise formulation of the *true* information need is required. Obtaining such a formulation has proven to be a non trivial task [17].

Characterising supply – Good characterisation of information resources is imperative for effective information discovery, as poor characterisations inevitably lead to the retrieval of irrelevant information, or the missing of relevant information. An important question is of course which properties to include in a characterisation.

A useful property to include seems to be what an information resource is *about*. In addition, properties like authorship, price, medium, etc. may be included. In the literature standard attribute sets to characterise resources can be found in the context of metadata

standardisation efforts [9, 46, 52].

Matching demand & supply – The selection of relevant information resources for a given query is a well understood problem. The field of information retrieval has developed a number of retrieval models.

In the past, our research group has studied several aspects of these challenges (formulating needs: [12, 10, 14, 26, 32], characterising supply: [42, 24, 25], matching demand & supply: [4, 5, 6, 7]).

Information retrieval traditionally focusses on the *topical* relevance of a specific information resource to some well articulated information request. In the past we have also worked on refinements of this:

Coverage of information need – One resources is likely not to be enough to satisfy the needs of a searcher. What is really needed is a set of resources with an advice on the order in which to read/consume them [35].

Aptness of information resources – Resources may be available in a format not befitting the situation of the searcher. It may have the wrong storage format, wrong media, etc. Topical relevance, the traditional focus of information retrieval, is not enough to reason about the *aptness* of a resources to such a situation [19, 20, 22].

We are currently working on a theory to put all these pieces together in an information market context. In this paper, we therefore focus on the elaboration of the *Information Market Paradigm*, which we regard as a follow up to the *Information Retrieval Paradigm*. In this paper, we are therefore not concerned with developing yet another approach/strategy to match demand and supply of information, but rather with an attempt of more fundamentally understanding the workings of the *Information Market*.

The remainder of this paper is structured as follows. Section 2 studies the notion of a market in general. In section 3 this understanding is specialised to the information market. Here we will, as a form of a theoretical validation, position some existing (including our own) research efforts relative to the information market. The conclusion will also identify some of the open research challenges we see with regards to the information market.

2 Markets

Our generalised perspective on markets as presented here, is partially based on the concept of economic markets, in particular on the field of micro-economics [50, 28]. However, we consider economic markets to be a specific class of markets dealing with the trading of goods, services and *money*. Economic markets presume the existance of some form of currency to serve as a universal trading unit. In our study of markets as such, we generalize from this requirements. Nevertheless, our considerations are indeed inspired by literature on economic theories (without claiming to have a full overview of relevant literature from the economic field). Our considerations are primarily based on the work reported in [43, 44, 45, 49, 50, 28], as well as introspection.

2.1 Traded assets

In our view, two main classes of *assets* can be traded on a market:

Ownership of entities – Ownership of entities, such as physical goods, bank notes, part of an organisation, land, copyright, etc.

Execution of services – Services that may be applied on/to/over entities (which could actually one of the participants), that are regarded by some participant as value adding. Examples would be: the painting of a house, treatment of an illness, management of a stock portfolio, quality assessment of a car, etc.

Even though one may hold the position that, in the case of markets dealing with physical goods are traded, it are indeed the physical goods which are the entities being traded, we take the view that what is *actually* traded is the *ownership* of these entities. Trading entities on a market can, in our view, only be discussed if these entities can be regarded as being *owned* by some participant on the market. Trading an entity involves a *change* of ownership.

The class of executable services could be split further into:

Transformation of entities – Services which aim to transform some property of an entity. Examples would be: transportation of a chair from a warehouse to someone's living room, copying a file accross the Internet to a searcher's local computer, the composition of a car from its components, creation of an abstract for a lengthy book, etc.

Reduction of uncertainty – Services which are typically aimed at reducing some form of uncertainty about the entities and/or participants that are traded/trade on the market. Examples would be: matching demand and supply by a broker, quality appraisal of products on offer, appraisal of a participant's credit rating, etc.

In the next section we will discuss some information market pendants of these classes of services. Let us now, however, first explore markets in more detail. We will do so by discussing four core concepts: transactions, cost/benefits, preference and value addition.

2.2 Transactions

Let p_1 and p_2 be two participants of the market, and let o_1 and o_2 be two assets that are on offer by these two respective participants. The participants may decide to trade these assets, leading to a *transaction*. This is illustrated by the sequence diagram on the left-hand side of figure 2.

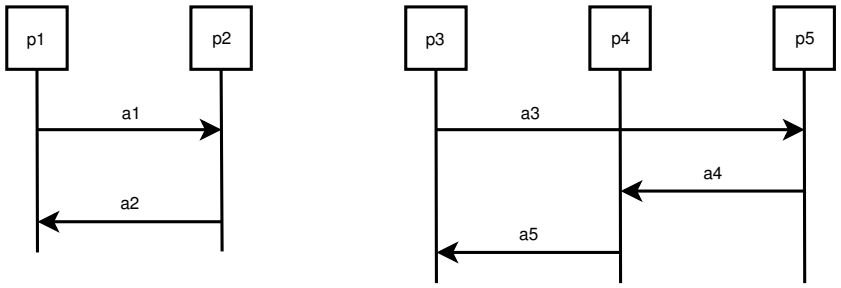

Figure 2: Two transactions

In economic markets, one is used to referring to participants as either being a *selling* or a *buying* participant. In our view, the notion of selling and buying can only be defined relative to a specific asset that is involved in the transaction. In the example given above, one could state that participant p_1 sells asset o_1 to participant p_2, making p_1 the selling and p_2 the buying party. By the same token, however, one could state that participant p_2 sells asset o_2 to p_1. We argue that there is nothing wrong with this duality. As a more concrete example of this duality, consider the following example:

> *When some person visits a flower shop to obtain a bouquet of roses, they may do so by handing over €10.– to the person behind the counter. In this situation, we would state that the person behind the counter* sells *the flowers to the person in exchange for €10.–. However, one could also state that the person visiting the flower shop,* sells *a note of €10.– in exchange for a bouquet of roses.*

It is only our day-to-day use of the terms *sell* and *buy* in the context of the *economic markets* that have lead to a uni-directional view on selling and buying. The "pre-occupation" of economic

market with the role of *money* as a universal means of trading, has produced to the default interpretation of seller and buyer.

The sales of an asset by a participant to another participant, will be referred to as a trans*actand*. Let t be a transactand, then we will use $t : s \xrightarrow{a} b$ to denote the fact that in transactand t participant s sells asset a to participant b. A transaction can now be regarded as being a set of transactands. If T is a transaction, then we can define:

$$s \xrightarrow{a} b \in T \;\triangleq\; \exists_{t \in T} \left[t : s \xrightarrow{a} b \right]$$

As a rule we will require:

$$t_1, t_2 \in T \wedge t_1 : s \xrightarrow{a} b \;\wedge\; t_2 : s \xrightarrow{a} b \;\Rightarrow\; t_1 = t_2$$

In other words, the involved participants and asset uniquely determine the transactand in a transaction. This will allow to denote the transaction depicted on the left-hand side of figure 2 as:

$$\{p_1 \xrightarrow{a_1} p_2,\; p_2 \xrightarrow{a_2} p_1\}$$

and the one on the right-hand side as:

$$\{p_3 \xrightarrow{a_3} p_5,\; p_5 \xrightarrow{a_4} p_4,\; p_4 \xrightarrow{a_5} p_3\}$$

The set of participants involved in:w a transaction are defined as:

$$\mathsf{Participants}(T) \;\triangleq\; \left\{ p_1 \;\middle|\; p_1 \xrightarrow{a} p_2 \in T \;\vee\; p_2 \xrightarrow{a} p_1 \in T \right\}$$

Now we have broken down the transactions that may take place on a market into more elementary parts, we can turn our attention to the drivers behind the transactions. Why do the transactions take place in the first place? There is usually some *benefit* to the participants of a transaction, even though this may be an *'artificial'* benefit such as holding on to one's life in the case of a *you're money or your life* situation.

We presume the participants of the market to behave in a goal-driven manor. These goals might be explicit in the reasoning of the participants, but may also be more implicit and based on emotions. For the moment we presume $\mathcal{GL}$ to be the set of possible goals. Let furthemore, $\mathcal{PA}$ be the set of participants on the market and $\mathcal{ST}$ be the set of states a participant may hold. We presume the function: $\mathsf{Id} : \mathcal{ST} \rightarrow \mathcal{PA}$ to identify which states belong to which participant. Given the state s of a participant $\mathsf{Id}(s)$, we can view the satisfaction of the goals which the participant (in a certain state!) may have as a function: $\mathsf{Satisfaction} : \mathcal{ST} \times \mathcal{GL} \rightarrow [0..1]$. For each goal, the level of satisfaction is expressed as a number between 0 and 1.

The consumption of some asset by a participant in a transaction, will result in a change of state of that participant. If T is a transaction, and s is a participant state, then $s \ltimes T$ is the state which results after the participation of $\mathsf{Id}(s)$ in transaction T. We require the resulting state to belong to the original participant: $\mathsf{Id}(s) = \mathsf{Id}(s \ltimes T)$ and the participant to indeed be a participant of the transaction: $\mathsf{Id}(s) \in \mathsf{Participants}(T)$.

On closer consideration, our statement: $p_1 \xrightarrow{a} p_2$ as an abbreviation for: "Participant p_1 sells asset a to participant p_2" is not specific enough. An actual transaction will take place between participants who hold a specific *state*. For our considerations in the next subsections, we will need this more refined view. We will therefore use $t : s_1 \xrightarrow{a} s_2$ as an abbreviation for: "In transactand t, participant $\mathsf{Id}(s_1)$ in state s_1 sells asset a to participant $\mathsf{Id}(s_2)$ in state s_2". We do require:

$$t : s_1 \xrightarrow{a} s_2 \;\Rightarrow\; t : \mathsf{Id}(s_1) \xrightarrow{a} \mathsf{Id}(s_1)$$

The set of states involved in a transaction is identified as:

$$\mathsf{States}(T) \;\triangleq\; \left\{ s_1 \;\middle|\; \exists_{s_2, a} \left[s_1 \xrightarrow{a} s_2 \in T \;\vee\; s_2 \xrightarrow{a} s_1 \in T \right] \right\}$$

2.3 Costs and benefits

The actual benefit of an asset is difficult to measure. This also makes it hard for participants to asses whether they wish to purchase/consume the resource or not: the only way to assess the true benefit is by consuming it! See for example [45] for an overview of issues pertaining to pricing and valueing resources from an economic point of view.

In [2] Ken Alder writes "Our methods of measurement define who we are and what we value". In his book, he describes the quest or a universal measure for distance in the late 1790's by two astronomers. Their task was to establish this new measure –"the meter" as one ten-millionth of the distance from the North Pole to the equator. This is, by the standards deployed in these days, as well as by modern standards, a daunting task. Where the astronomers Delambre and Méchain's quest was to find a measure for distance, the "queste" for markets in general is to present a measure for value (cost/benefits) of assets.

We presume that the benefits of an involvement in transaction can be defined as the positive impact on the satisfaction levels of a participant:

$$\mathsf{Benefit}(s,T) \triangleq \lambda_{g\in\mathcal{GL}}.\text{MAX}(\mathsf{Satisfaction}(s \ltimes T, g) - \mathsf{Satisfaction}(s,g))$$

We have employed the Lambda calculus notation [8] to denote a function ranging over $\mathcal{GL}$. The costs of an involvement in a transaction can be defined as the negative impact on the satisfaction levels of a participant:

$$\mathsf{Cost}(s,T) \triangleq \lambda_{g\in\mathcal{GL}}.\text{MAX}(\mathsf{Satisfaction}(s,g) - \mathsf{Satisfaction}(s \ltimes T, g))$$

Given a relative prioritisation of the different goals, a weighed level of satisfaction could be computed. Let $\mathsf{Priority} : \mathcal{ST} \times \mathcal{GL} \rightarrow [0..1]$ therefore be a function which identified the level of priority a participant (in a specific state) gives to the specified goal. We presume the priority function to be a distribution totalling to one for each of the states:

$$\forall_{s\in\mathcal{ST}}\,[\Sigma_{g\in\mathcal{GL}}\,\mathsf{Priority}(s,g) = 1]$$

With this weighing function, we can define the overall satisfaction as follows:

$$\mathsf{Satisfaction}(s) \triangleq \Sigma_{g\in\mathcal{GL}}\,\mathsf{Satisfaction}(s,g) \times \mathsf{Priority}(s,g)$$

It seems sensible to presume that the level of satisfaction of all participants of a transaction should not decrease:

$$\forall_{s\in\mathsf{States}(T)}\,[\mathsf{Satisfaction}(s) \leq \mathsf{Satisfaction}(s \ltimes T)]$$

An increment in satisfaction does not have to be a *hard* goal such as the quantity of possession of goods, money, etc, but could also be a *soft* goal such as social esteem or appreciation by friends.

2.4 Preference

The concept of benefits is very much related to that of preference: people prefer things that have a greater benefit over things with less benefit to them. It is not always apparent why some asset is preferred and, thus, why it is of greater benefit to the person who made the choice. The notion of preference (and the relation between preference and utility) is discussed in the field of micro economics (see e.g. [50, 28]). In micro economics it is assumed that:

- consumers exhibit rational behavior in the sense that they prefer (bundles of) assets with a high benefit over assets with a low benefit,
- a binary (complete, transitive and irreflexive) preference relation $\succ$ such that $a \succ b$ means that bundle a is preferred over bundle b,
- a utility function U may be constructed such that $U(a) > U(b)$ iff $a \succ b$.

More elaborate schemes for preference in a micro economical context exist as well. For example,[47] describes a reference-dependent approach to utility and preference. The core of the described approach is that preference is dependent on a current "position". In other words, $a \succ b|p$ should be read as "a is weakly preferred to b viewed from position p.

We can relate this directly to our discussions above, when (1) equating the notion of a "bundle" as the "goods" received by a participant in transaction and (2) equating position to participant state. A position (state) aware utility function over bundles (transactions) might be defined as:

$$U(T|s) \triangleq \mathsf{Satisfaction}(s \ltimes T) - \mathsf{Satisfaction}(s)$$

leading to:

$$T \succ T'|s \triangleq U(T|s) > U(T'|s)$$

2.5 Value addition

Given some tradeable asset, one may decide to perform a transformation to it. For example, a TV may be transported from a warehouse to your living room, a document may be downloaded to your PC, a car may be washed and a document can be converted from one format to another. From the perspective of some participant (state), the transformation may, or may not, be value adding. If a is an asset, and $S(a)$ is the asset which results after performing some transformation S to it, then the added value of performing S to a, *in the context of* a participant in state s and a transaction T, can be defined as:

$$\mathsf{AddedValue}(S, a)[s, T] \triangleq \mathsf{Satisfaction}(s \ltimes T) - \mathsf{Satisfaction}(s \ltimes T^a_{S(a)})$$

where $T^o_{S(a)}$ is the transaction which differs from T *only* in that all transactands involving a have been changed to involve $S(a)$. More formally:

$$\begin{aligned} R^a &\triangleq \left\{ s_1 \xrightarrow{a} s_2 \;\middle|\; s_1 \xrightarrow{a} s_2 \in T \right\} \\ A^a_{a'} &\triangleq \left\{ s_1 \xrightarrow{a'} s_2 \;\middle|\; s_1 \xrightarrow{a} s_2 \in T \right\} \\ T^a_{S(a)} &\triangleq (T - R^a) \cup A^a_{S(a)} \end{aligned}$$

The following example illustrates the role of a value adding transformations:

> *Consider the market for antiquities such as paintings. Consumers can either buy a painting from another person, or via an intermediary at an auction. In the first case, value is transferred from the seller to the buyer in the form of the painting, and back in the form of a payment. In the latter case, the seller expects that selling his painting at the auction will result in a higher price. Even more, this price has to exceed the fee that hey (probably) has to pay to be able to sell at this auction. Also, from the consumer point of view, buying at an auction may have a higher value because of the painting is first checked by experts (is it really a* Van Gogh*), because of extra insurance etcetera.*

3 The Information Market

This section is concerned with a specialisation of the ideas presented in the previous section to the context of the information market. It will also position some of the pre-existing research relative to the notion of the information market.

3.1 The assets

In accordance to [22, 19] the entities traded on the information market are dubbed *information resources*, or *resources* for short. In the context of the Web, an information resource can be

defined as [34]: any entity that is accessible on the Web, and which can provide information to other entities connected to the Web. A definition which truly supports the open character of the net. Examples of information carriers included are: web pages (including free text, sound, images, and video fragments), free text databases, traditional (relational, object-oriented, etc.) databases and people's e-mail addresses.

Even though the trading is about information resources, it are actually different levels of ownership/usage rights that are traded. One could distinguish between four main classes:

1. Right to read/consume the information resources for a fixed period of time (For example, renting a DVD).
2. Right to show the contents of the information resources to other parties (For example, having the right to show the DVD at a public party).
3. Right to redistribute, i.e. produce copies (For example, having the right to make copies of a DVD and re-sell these copies).
4. Full transfer of ownership

Most information resources that are available to us via the Web, in particular Web pages, fall in the first two classes. Quite often, the information will actually be available for free. That is to say, the amount of Euro's one has to hand over in exchange is close to zero. As we will see below, the costs/benefits of information resource involves more than the amount of money that is handed over.

In addition to trading of ownership/usage of information sources, services pertaining to these information sources are traded as well. Such services may include:

- Transformation of an information resource's storage format (e.g. from OpenOffice to PDF, BMP to JPEG).
- Translation of an information resource from one language to another (e.g. Dutch to Danish or from French to German).
- Transfer of information resources from one location on the Internet to another location (e.g. downloading a file).
- Changes to the contents of an information resource (e.g. creating an abstract for a lengthy document).
- Quality appraisal of the content of an information resource (e.g. selection of high quality papers for an E-Journal).
- Aiding searchers in articulating/formulating their specific need for information (e.g. offering searchers a query-by-navigation interface [1, 14]).
- Listing information resources that match some pre-set criteria (e.g. a traditional search engine [3, 23]).

Information resources and related services are not the only assets traded on the market. Producers (and transformers) of information resources will only do so if they have a reason. In other words, there must be some flow of assets back to the producers. This backward flow will have to originate from the consumers of the information resources. This flow could consist of money, but could equally well deal with intangiable assets such as intellectual esteem, personal achievement, social standing, etc. In this sense, the workings of the open source cumminities are quite illustrating. In the open source community [38], software is developed (one could actually argue that software, and in particular the sources, are forms of information resources as well) that is consequently used by people who do (usually) not pay any money for the right to use this software. The "payback" for the developers of this software is in social esteem (within society and/or the hackers world) and personal enjoyment [38]. Quantifying the backwards flow on an information market is also a major issue in the field of knowledge management [31, 29, 27]. One of major blockages for knowledge management seems to be the willingness for people to share knowledge, which usually boils down to the question "what will people get in return?".

3.2 Transactions

Transactions on the information market as such, will not differ dramatically from markets in general. However, in the case of the information market, we can elaborate more on the goals which drive the consumers of information resources.

A future consumer of an information resource should have a need for information. This need for information can be caused by a number of reasons. At the moment we distinguish between two types of goals: *increment of knowledge* and *change of mood*. The former corresponds to a situation where someone finds that they are lacking some information/knowledge. This "knowledge gap" [34] could pertain to something fairly specific such as learning *the latest price of 19 micron wool*, to the very broad, such as learning about the *theory of relativity*. Note that viewing the consumption of information with the aim of filling a gap in ones knowledge could benefit from fields such as goal-driven learning [36, 37] as it enables for explicit reasoning about the goals with which consumers look for information.

When a consumer aims to achieve a *change of mood*, then this probably indicates a situation where an information resource is needed such as a book, movie, or music, to influence the mood. This may range from music that is uplifting, a music that is relaxing, a movie to take someones mind of unpleasant affairs, etc.

Collectively, one can refer to these two types of goals as *cognitive* goals. In addition to a cognitive goal, a consumer of information will have some *operational goal* as well. This latter goal relates to the tasks the consumer has/wants to perform. These tasks may put requirements (such as timeliness) on the information consumption (and searching!) process. For example, timeliness is a good reason for turning to a (good) search engine when searching for some specific information.

3.3 Costs and benefits

The costs and benefits of an information resource are particularly difficult to measure. We shall adopt a multi-dimensional view on measuring the potential benefit of a resource:

Utility – dealing with the information that may be provided by a resources and the timeliness. This refers to the contents of the information resources; the topic/aboutness of a resource.

Structure – concerned with the form (report, painting, movie, audio) and format (PDF, MP3) of a resource.

Emotion – dealing with the emotional effect (pretty/inspiring) that a resource may have when it is "consumed".

These value domains actually closely correspond the three aspects of *architecture* as introduced by Vitruvius, a Roman writer, architect and engineer, active in the 1st century BC. These aspects were called *utilitas* (our utility domain), *firmitas* (our structure domain) and *venustas* (our emotional domain). See e.g. [40, 51, 53] for details.

This benefits-taxonomy can be used to explore the potential benefits of resources in diverse situations. For instance, in case of art the emotional benefit is likely to be more important than in the case of technical reports, for searchers using a WAP or I-mode based connection the structural aspects are likely to be if high importance.

The costs associated to a resource also fits the above discussed multi-dimensional domain. For a searcher these costs would, for example, include:

Utility – The costs of actually obtaining the resource such as search costs (time and money), costs for the Web-connection etcetera.

Structure – The amount of disk space needed to store the information resources at a convenient location, computing capacity needed to display the information resource, etc.

Emotion – The costs associated to actually conceiving the resource (i.e. the cognitive load associated with interpreting and understanding the resource. These are costs from the informational domain. See e.g. [48] for more details.

For a publisher these costs would, for instance, include:

Utility – The costs associated to creating the resource such as time and effort.

Structure – The costs associated to storing the resource such as disk space, as well as required computing power in creating the resource.

Emotion – Intellectual energy needed to create the contents of the resource. This may also be referred to as cognitive load [11].

3.4 Preferences

On the information market, consumers will have preferences just as on a market in general. To a large extend, these preferences will be dominated by the specific state the consumer is in. However, some factors might be discerned in a consumer's preferences that will remain stable over time. For example, a consumer's long term interests, hobby's, intellectual capacities, domain of work, etc.

When made explicit, in terms of so-called profiles, these preferences can make a consumer's life on the information market easier. Already in [30, 16, 33, 21] it was recognised that information retrieval systems can be personalised for users by means of profiles. Such profiles may be used by search engines to better select apt information sources [21]. They can also be employed to aid in the actual formalation of information needs [34]. For example, some person may, whenever using the term *surfing*, actually refer to *surfing the waves on the Sunshine Coast*. This "default" interpretation may be part of the profile.

Consumer preferences do not only pertain to the content of information resources. They may also pertain to, for example, the format/media of the resource they may be able (want to) consume. When using a PDA over a wireless connection to connect to the Web, movies are a rather awkward medium. In this case, text-based information resources will be preferred over movies.

3.5 Value addition

Value addition on the information market may be achieved by the earlier discussed services, such as: (1) transformation of an information resource's storage format, (2) translation of an information resource from one language to another, (3) aiding searchers in articulating/formulating their specific need for information, and (4) matching demand for information to supply of information. Traditionally, information retrieval focusses on (4) with extensions towards (3). In our opinion, theories are needed to underpin all services that enable the working (value addition) of the information market. Such theories will need to take the goals of all participants of the market into consideration.

4 Conclusion

At the start of this paper we have discussed how an evolution can be observed moving beyond the traditional information retrieval paradigm to an information market paradigm. We have provided a discussion on the general notion of a market where assets are traded. This was then specialised to information resources, leading to an information market.

At present, we are working on a more fundamental understanding of markets in general and information markets in particular. Based on these insights, we will evolve our existing theories for different aspects of information retrieval. We expect that models for goal-driven reasoning of participants on the information market, will in particular be fruitfull in improving the workings of the information market. Most importantly the retrieval of relevant information by searchers. This, however, requires a thorough and fundamental understanding of the goals of the participants on the market and how the traded information resources may contribute towards these goals, i.e. what their costs/benefits are with regards to these goals.

References

[1] M. Agosti, A. Archi, R. Colotti, R.M. Di Giorgi, G. Gradenigo, B. Inghirami, P. Matiello, R. Nannuci, and M. Ragona. New prospectives in information retrieval techniques: a hypertext prototype in environmental law. In *Online Management 89, Proceedings 13th International Online Information*, pages 483–494, London, United Kingdom, EU, 1989.

[2] K. Alder. *The measure of all things: The Seven-Year Odyssey and Hidden Error That Transformed the World*. Free Press, New York, New York, USA, 2002. ISBN 0-743-21675-X

[3] Altavista search engine project, 1999.
`http://www.altavista.com`

[4] A.T. Arampatzis, T. Tsoris, and C.H.A. Koster. IRENA: Information Retrieval Engine based on Natural language Analysis. In *Proceedings of the RIAO'97 Conference*, pages 159–175, McGill University, Montreal, Canada, 1997.

[5] A.T. Arampatzis, T. Tsoris, C.H.A. Koster, and Th.P. van der Weide. Phrase-based Information Retrieval. *Information Processing & Management*, 34(6):693–707, December 1998.

[6] A.T. Arampatzis, Th.P. van der Weide, C.H.A. Koster, and P. van Bommel. Text Filtering using Linguistically-motivated Indexing Terms. Technical Report CSI-R9901, Computing Science Institute, University of Nijmegen, Nijmegen, The Netherlands, EU, January 1999.

[7] A.T. Arampatzis, Th.P. van der Weide, C.H.A. Koster, and P. van Bommel. *Linguistically-motivated Information Retrieval*, volume 69, pages 201–222. Marcel Dekker, New York, New York, USA, 2000.

[8] H.P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, volume 103 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, The Netherlands, EU, Revised Edition, 1984.

[9] T. Berners-Lee. Universal Resource Identifiers in WWW. Technical Report RFC1630, IETF Network Working Group, June 1994.

[10] P.D. Bruza. Hyperindices: A novel aid for searching in hypermedia. In A. Rizk, N. Streitz, and J. Andre, editors, *Proceedings of the European Conference on Hypertext – ECHT 90*, pages 109–122, Cambridge, United Kingdom, EU, 1990. Cambridge University Press.

[11] P.D. Bruza, S. Dennis, and R. McArthur. Interactive internet search: keyword directory and query reformulation mechanisms compared. In *Proceedings of the 23rd Annual ACM Conference of Research and Development in Information Retrieval (SIGIR'2000)*. ACM Press, 2000.

[12] P.D. Bruza and Th.P. van der Weide. Two level hypermedia - an improved architecture for hypertext. In A.M. Tjoa and R.R. Wagner, editors, *Proceedings of the Data Base and Expert System Applications Conference (DEXA 90)*, pages 76–83, Vienna, Austria, EU, 1990. Springer-Verlag, Berlin, Germany, EU.

[13] P.D. Bruza and Th.P. van der Weide. The modelling and retrieval of documents using index expressions. *ACM SIGIR FORUM (Refereed Section)*, 25(2), 1991.

[14] P.D. Bruza and Th.P. van der Weide. Stratified Hypermedia Structures for Information Disclosure. *The Computer Journal*, 35(3):208–220, 1992.

[15] H. Chen, A. Houston, R. Sewell, and R. Schatz. Internet browsing and searching: User evaluations of category map and concept space techniques. *Journal of the American Society for Information Science*, 49(7):604–618, 1998.

[16] Pei-Min Chen and Fong-Chou Kuo. An information retrieval system based on a user-profile. *The Journal of Systems and Software*, 54(1):3–8, september 2000.

[17] C.W. Cleverdon. The Significance of the Cranfield Tests on Index Languages. In A. Bookstein, Y. Chiarmarella, G.E Salton, and V.V. Raghavan, editors, *Proceedings of the 14th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 3–12, Chicago, Illinois, USA, October 1991. ACM Press.

[18] B.C. Desai. Supporting Discovery in Virtual Libraries. *Journal of the American Society for Information Science*, 48(3):190–204, 1997.

[19] B. van Gils, H.A. Proper, and P. van Bommel. A conceptual model for information suppy. *Data & Knowledge Engineering*, 51:189–222, 2003.

[20] B. van Gils, H.A. Proper, and P. van Bommel. Towards a general theory for information supply. In C. Stephanidis, editor, *Proceedings of the 10th International Conference on Human-Computer Interaction*, pages 720–724, Crete, Greece, EU, 2003. ISBN 0-805-84930-0

[21] B. van Gils, H.A. Proper, P. van Bommel, and E.D. Schabell. Profile-based retrieval on the world wide web. In P.M.E. de Bra, editor, *Proceedings of the Conferentie Informatiewetenschap (INFWET2003)*, pages 91–98, Eindhoven, The Netherlands, EU, 2003.

[22] B. van Gils, H.A. Proper, P. van Bommel, and Th.P. van der Weide. Transformations in information supply. In J. Grundspenkis and M. Kirikova, editors, *Proceedings of the Workshop on Web Information Systems Modelling (WISM'04), held in conjunctiun with the 16th Conference on Advanced Information Systems 2004 (CAiSE 2004)*, volume 3, pages 60–78, Riga, Latvia, EU, June 2004. Faculty of Computer Science and Information Technology, Riga Technical University, Riga, Latvia, EU.

[23] Google search engine, 2001.
`http://www.google.com/`

[24] F.A. Grootjen. Employing semantical issues in syntactical navigation. In *Proceedings of the 22nd BCS-IRSG Colloquium on IR Research*, pages 22–33, Cambridge, United Kingdom, EU, 2000.

[25] F.A. Grootjen. Indexing using a grammerless parser. In *2001 IEEE International Conference on Systems, Man & Cybernetics (SMC2001)*, Tucson, Arizona, USA, 2001. ISBN 0-780-37089-9

[26] A.H.M. ter Hofstede, H.A. Proper, and Th.P. van der Weide. Query formulation as an information retrieval problem. *The Computer Journal*, 39(4):255–274, September 1996.

[27] S.J.B.A. Hoppenbrouwers and H.A. Proper. Knowledge discovery - de zoektocht naar verhulde en onthulde kennis. *DB/Magazine*, 10(7):21–25, November 1999. In Dutch.

[28] M.L. Katz and H. S. Rosen. *Microeconomics*. Irwin, 2nd edition, 1994. ISBN 0-256-11171-5

[29] D. Leonard-Barton and D. Leonard. *Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation*. Harvard Business School Press, Boston, Massachusetts, USA, 1998. ISBN 0-875-84859-1

[30] S.H. Myaeng and R. R. Korfhage. Towards an intelligent and personalized retrieval system. In *Proceedings of the ACM SIGART international symposium on Methodologies for intelligent systems*, pages 121–129, Knoxville, Tennessee, United States, 1986. ACM Press. ISBN:0-89791-206-3

[31] I. Nonaka and H. Takeuchi. The knowledge-creating company. *Harvard Business Review*, (November-December):97–130, 1991.

[32] M.P. Papazoglou, H.A. Proper, and J. Yang. Landscaping the information space of large multi-database networks. *Data & Knowledge Engineering*, 36(3):251–281, 2001.

[33] S. Pierra, C. Kacan, and W. Probst. An agent-based approach for integrating user profiles into a knowledge management process. *Knowledge-Based Systems*, 13(5):307 – 314, October 2000.

[34] H.A. Proper and P.D. Bruza. What is Information Discovery About? *Journal of the American Society for Information Science*, 50(9):737–750, July 1999.

[35] H.A. Proper and Th.P. van der Weide. Information coverage – incrementally satisfying a searcher's information need. In C. Stephanidis, editor, *Universal Acces in HCI: Towards an Information Society for All*, pages 719–722, New Orleans, Louisiana, USA, August 2001. Lawrence Erlbaum Associates, Mahwah, New Jersey, USA. ISBN 0-805-83609-8

[36] A. Ram and D. Leake. A framework for goal-driven learning. In M. des Jardins and A. Ram, editors, *Proceedings of the 1994 AAAI Spring Symposium on Goal-Driven Learning*, pages 1–11, 1994.

[37] A. Ram and D.B. Leake. *Goal-Driven Learning*. MIT Press, Cambridge, Massachusetts, USA, 1995. 0-262-68083-1

[38] E.S. Raymond. *The Cathedral & The Bazaar*. O'Reilly, Sebastopol, California, USA, 1999. ISBN 1-565-92724-9

[39] C.J. van Rijsbergen. *Information Retrieval*. Butterworths, London, United Kingdom, EU, 1975.

[40] D. Rijsenbrij. *Architectuur in de digitale wereld (versie nulpuntdrie)*. Nijmegen Institute for Information and Computing Sciences, Radboud University Nijmegen, Nijmegen, The Netherlands, EU, October 2004. In Dutch. ISBN 90-90188285-3

[41] G.E Salton and M.J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, New York, New York, USA, 1983.

[42] J.J. Sarbo, J.I. Farkas, F.A. Grootjen, P. van Bommel, and Th.P. van der Weide. Meaning Extraction from a Peircean Perspective. *International Journal of Computing Anticipatory Systems*, 6:209–227, 2000.

[43] M. Sarkar, B. Butler, and C. Steinfield. Cybermediaries in electronic marketspace: Toward theory building. *Journal of Business Research*, 41(3):215–221, March 1998.

[44] C.H. Scott and J.E. Scott. On models for the operation of a class of electronic marketplaces. *Omega*, 32(5):373–383, October 2004.

[45] C.E. Shannon and H.R. Varian. *Information Rules, a strategic guide to the network economy*. Harvard Business School Press, Boston, Massachusetts, USA, 1999. ISBN 0-975-84863-X

[46] K. Sollins and L. Masinter. Functional requirements for uniform resource names. Technical Report RFC1737, IETF Network Working Group, http://www.ietf.org/rfc/rfc1737.txt, December 1994.

[47] R. Sugden. Reference-dependent subjective exptected utility. *Journal of economic theory*, 11(2):172–191, 2003.

[48] H. Tardieu and V. Gyselinck. Working memory constraints in the integration and comprehension of information in a multimedia context. In H. van Oostendorp, editor, *Cognition in a digital world*, chapter 1, pages 3–24. Lawrence Erlbaum Associates, Hillsdale, New Jersey, USA, 2003. ISBN 0-805-83507-5

[49] E. Turban, J. Lee, H.M. King, and D. Ching. *Electronic Commerce, a managerial perspective*. Prentice-Hall, Upper Saddle River, New Jersey, USA, 2000. ISBN 0-13-975285-4

[50] H.R. Varian. *Intermediate Microeconomics, a modern approach*. Norton, New York, USA, 4th edition, 1996. ISBN 0-393-96842-1

[51] M. Vitruvius. *Handboek Bouwkunde*. Athenaeum – Polak & Van Gennep, Amsterdam, The Netherlands, EU, 1999. Translated by: T. Peters. ISBN 90-253-5870-5

[52] S. Weibel, J. Godby, E. Miller, and R. Daniel. Metadata workshop report. Dublin, Ohio, USA, March 1995.

[53] Vitruvius, October 2004. Last checked: 26-Oct-2004.
`http://en.wikipedia.org/wiki/Vitruvius`