EP2010 – A STRATEGIC EUROPEAN STUDY ON THE FUTURE OF ELECTRONIC PUBLISHING TOWARDS 2010

GEORG GÜNTNER¹; WERNHER BEHRENDT¹; GUNTRAM GESER¹

¹ Salzburg Research Forschungs- gesellschaft mbH
Jakob Haringer Strasse 5/III, A-5020 Salzburg, Austria
e-mail: georg.guentner@salzburgresearch.at; wernher.behrendt@salzburgresearch.at; guntram.geser@salzburgresearch.at

Our presentation gives an outlook on some of the main results of a strategic European study entitled “EP2010 – The Future of Electronic Publishing Towards 2010”. The paper puts forward the thesis that, driven by the emerging knowledge economy, “traditional” publishing industries will have to face the technological and economic challenges of knowledge-enhanced publishing: Novel publishing objects will provide different information layers (adapters) including an underlying knowledge structure to enable value adding (in a very broad sense) along various steps in the publishing workflow. We introduce the concept of “smart content” to define, along the digital content value chain, the major technological challenges and market scenarios for the future of the electronic publishing sector in the knowledge economy. Smart content objects will, among other properties, be personalised, adaptable, evolvable, unobtrusive, interoperable and ubiquitous. A proposed research map including foundational research, application-oriented research on the publishing and knowledge management workflow and integrative research, e.g. on semantic interoperability, reflects the needs of the publishing industries regarding the paradigm change from traditional “canned” publishing towards knowledge based publishing in 2010.

Keywords: smart content; knowledge management; content creation; electronic publishing; digital content value chain; strategic study; knowledge economy; innovation cycle; research policy

INTRODUCTION

STRUCTURE
Our presentation gives an outlook on some of the basic concepts and results of a study on the future of electronic publishing toward 2010. (1) The introduction summarises the background and the main objectives of EP2010. (2) We will then outline the vision of the emerging knowledge economy, analyse the current situation of the knowledge economy and raise the question if “knowledge units” could be a future challenge for the publishing industries. (3) We will further introduce the concept of smart content as a unifying vision for the knowledge management and content creation area and the electronic publishing sector. We will classify different groups of user requirements which will lead us to a functional grouping of requirements and finally to a “smart content service architecture”. (4) Finally we will discuss general research policy options and recommendations for foundational, application-oriented and integrative research.

BACKGROUND OF THE STUDY
In early 2002 the European Commission, DG Information Society, issued a call for tender for a study on the future of electronic publishing. The commissioner of the call was Unit D1, Electronic Publishing. When the work on the study finally began, it became obvious that the agenda of Unit D1, dealing explicitly with issues relating to (interactive) electronic publishing, was transferred to another unit, namely Unit E2, Knowledge Management & Content Creation (KMCC), following the new structure of the Directorate-General Information Society in the Sixth Framework Programme. It will be helpful for the understanding of the concepts and results shown in this paper to be aware that KMCC is part of a Directorate dealing with interfaces, knowledge and content creation, applications and the information market. We would say that - without an alternative - this is still the home of electronic publishing, but the area is structured along other thematic priorities. This paper also reflects this shift, introducing a set of knowledge-based or knowledge-aware extensions to the prevalent view on electronic publishing.

OBJECTIVES OF EP2010
The primary objective of EP2010 is to provide key orientations for policy makers and input to concrete implementation measures with regards to future issues in the electronic publishing area. In particular, EP2010 is expected to come to terms with three quite distinct goals:

Firstly, to provide results relevant for conceiving and fine tuning IST RTD programmes (Sixth Framework Programme, Seventh Framework Programme – see also [1]), in particular RTD related with the
sector of interactive electronic publishing (IEP). The actual success of the RTD programme rests on the willingness of researchers to take up the challenges, and on the willingness of the private sector to invest in the proposed RTD programme. The traditional industrial constituency of KMCC has been the publishing sector. For the Sixth Framework Programme, researchers and industrials in knowledge technologies are entering the scene. The key concepts and economic potential of a knowledge-based economy need to be defined in more detail before investors buy into the vision. Major shifts in research agenda may be required to address some of the challenges. For example, natural language generation may be more interesting than natural language recognition.

Secondly, to provide a vision that combines knowledge technologies and content creation. The research goals of the publishing sector and those of the EU/IST programme guided by ISTAG (IST Advisory Group; see also [2], [3] and [4]) need not necessarily be compatible. While the notion of “knowledge resources” as the new form of content may be agreeable to both parties, the notion of support for an all-digital, ambient intelligence space may be too visionary for those who need profitable products in the near to medium term, rather than medium to long term.

Thirdly, the study should also be relevant for the publishing industries which may only to a certain extent be prepared to participate in longer-term RTD in a future IEP sector. In this context the notion of “Smart Content” needs to be elaborated so that researchers can use it as a tool for systematic formulation of research objectives, technologists can use it as a high-level architectural metaphor for future information systems; and content providers can use it as a unit of value when they think of new knowledge/information services and products. Ideally, the properties of smart content are such that the consumer begins to associate tangible qualities and benefits with “smart” as opposed to “dumb” content.

**TOWARDS A “KNOWLEDGE ECONOMY”**

There are many ways in which one can define a “knowledge economy”, and attempt to analyse whether a society is based on such an economy. One formula that is considerably developed and shared in the research community, is the “knowledge-based” economy ([5]).

As such, we define knowledge economy in terms of what humans can do with knowledge-enhanced objects (of all kinds); what knowledge-enhanced objects can do with other knowledge-enhanced objects, and what (interrelated) knowledge-enhanced objects can do for humans. The emerging knowledge economy will build to an increasing degree on knowledge that is used for building tools that support the value chain of knowledge creation and delivery. This implies advanced technologies for knowledge creation and improvement, for designing and creating objects (physical and biological manufacturing of “knowledge-rich” objects), for programming objects (i.e. provide them with “capabilities”, for example, the capability to interact with other objects and/or humans, to reason, ...), for trading and licensing of “knowledge units” and knowledge-enhanced objects, for customising them for specific environments (e.g. adapt and “embed” them in work, education, leisure environments), and profit from other benefits associated with “knowledge units” and knowledge-enhanced objects, e.g. enhancing creative activities, or improving productivity.

Yet, the following points should be noted: Firstly, the knowledge economy is not fundamentally different from the existing one, i.e. “knowledge units” will not replace “real money”. Secondly, the knowledge economy will be dynamic and strong (in competitive terms) if it has people and organisations that are capable of building, marketing, adapting, and using knowledge tools and knowledge-enhanced products and services. Thirdly, there is a whole range of benefits of using “knowledge units” and knowledge-enhanced objects, that need to be explored (particularly for the end-users [B2C]). These include, to name just a few examples: achievements in learning, quality of social/community life (participation, inclusion), leisure and entertainment, privacy, personal freedom and security.

**A SITUATION ANALYSIS FOR THE KNOWLEDGE ECONOMY**

EP2010 provides an overview of the situation of the interactive electronic publishing sector in 1998, at present and a hypothetical preview for 2010. The analysis describes three stages from a “boom” (1998), to a “depression” (2003), and further to a hypothetical “up-take” in 2010 of the knowledge economy. With regards to knowledge concepts, knowledge-enhanced objects and their economic potential, we envision the following development.
TABLE 1 – SITUATION ANALYSIS: KNOWLEDGE ECONOMY

<table>
<thead>
<tr>
<th>1998</th>
<th>2003</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge concepts</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Data – information – knowledge – wisdom</td>
<td>- Ontologies come to the fore</td>
<td>- Knowledge units, k-agents.</td>
</tr>
<tr>
<td>- Data mining</td>
<td>- No clear picture on merging of ontologies</td>
<td>- First emergence of commercial ontologies</td>
</tr>
<tr>
<td></td>
<td>- Web mining for semi-structured documents</td>
<td>- First emergence of commercial knowledge banks</td>
</tr>
<tr>
<td><strong>Type of knowledge objects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Hyper-linked documents</td>
<td>- Semantic tagging in its infancy</td>
<td>- Semantically marked-up knowledge units.</td>
</tr>
<tr>
<td></td>
<td>- Some document models begin to include knowledge structures</td>
<td></td>
</tr>
<tr>
<td><strong>Knowledge trading (KUs: knowledge units)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- KUs (added) hidden in products or embedded in “canned content”.</td>
<td>- Ever more companies start focusing on their knowledge assets and put them on their balance sheet. (Edvinsson/Malone 1997; Roos et al. 1997)</td>
<td>- Increasing degree of creation, accumulation, and trading of KUs, e.g. concepts, specific ontologies, visualisations, arguments, explanations.</td>
</tr>
<tr>
<td></td>
<td>- Increased awareness of the potential of marketing knowledge not in “canned content” but as a service.</td>
<td>- First Smart Content company is bought by Microsoft.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Founding of the Organisation of Knowledge Exporting Countries (OKEC)?</td>
</tr>
</tbody>
</table>

KNOWLEDGE UNITS – A CHALLENGE FOR THE PUBLISHING INDUSTRIES?

There is an increasing awareness that for “traditional” publishers (i.e. newspaper, magazine, book publishers, information service providers, as well as broadcasters) the knowledge economy will bring “somehow different” rules of the game. For example, a recent study by PIRA, a leading commercial consultancy business specialising on the packaging, paper, printing and publishing industries, titled “Publishing in the Knowledge Society” (see [6] and [7]), issues the following warning: “The publishing industry must also make the transition towards servicing a rich, competitive knowledge economy.” However, the study focuses on the use of new technology mainly for productivity gains. It does not address to any great extent, knowledge technologies or new, knowledge-enhanced electronic publishing products or services.

The study suggests that (e-)publishers should invest in “knowledge-enhancements” e.g. to integrate to a higher degree business information systems, publishing systems, fulfilment and order processing; to get into the workflow of customers, i.e. integrate tightly with (business) customer’s IT systems; and to “get closer to the end-customers”, i.e. to make use of advanced customer relationship management techniques to, for example, profile customers more accurately.

However, these are only measures towards improving the “business intelligence”, but what is missing are strategies towards more “intelligent” products or services as one might envisage them in the knowledge economy. PIRA explicitly recommends that the UK Publishing Industry should “take the wrapper off”; “The products of the publishing industry have until recently been closely identified with the way they are packaged, for example, as books, magazines, newspapers, journals and directories, rather than as content which appears in different wrappers. Some products – for example learned journals and directories – have moved much further in the direction of being ‘unwrapped’ than others. Both publishers and users are beginning to evolve towards new concepts of what they can do with content using new technologies, but opportunities are not always grasped.”

More generally and back to the one of the theses of EP2010, the knowledge economy will bring new business opportunities for those players in the publishing sector who are able to address key differences between “traditional” and “knowledge-based” publishing, as indicated in the table below:
TABLE 2 – TRADITIONAL VS. KNOWLEDGE BASED PUBLISHING

<table>
<thead>
<tr>
<th>Traditional publishing</th>
<th>Knowledge based publishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canned one-off content in single proprietary formats, each requiring a specific production process</td>
<td>Non-canned, i.e. structured content in open formats with cross-media production process (i.e. one process for different “renditions” of the content)</td>
</tr>
<tr>
<td>Editions in multi-annual revision cycles</td>
<td>Evolvable content with on-line updates, revision tracing, etc.</td>
</tr>
<tr>
<td>Author “Fanclub”: Authors belong to schools of thought; publishers choose “gurus” as lead authors</td>
<td>Knowledge managers leverage the distributed knowledge of communities of practice, and organise around shared concepts e.g. regular updates, bringing in expert opinion, buying in / trading new related knowledge</td>
</tr>
<tr>
<td>Closed universe of professional discourses, because of the limitations imposed by “canned content”</td>
<td>“Cross section” publishing: While in “canned content” lines of discursive threads, in particular, inter- and trans-disciplinary ones, need to be cut off at certain points, K-managers can mine and offer these intersections as special “knowledge units” (and thereby maybe also act as innovation agents)</td>
</tr>
<tr>
<td>Mode of interaction is READ and make NOTES (if it is your own copy of the book)</td>
<td>Value adding services for the “Reader/User”: e.g. feedback template (e.g. “add your own case study to this book”) – from expert reader to author / publisher; tools for knowledge workers (bibliography service, knowledge based search); annotations/changes/improvements to “texts” (e.g. on a reward basis); and tradable (small) units – e.g. visualisation of a concept, by downloading (plug-in) the publisher’s visualisation tool, using the corporate/brand specific layout scheme</td>
</tr>
</tbody>
</table>

One alley that we envision to move from “traditional” electronic publishing to knowledge-based publishing is through the concept of “smart content”. We will use this concept to support the discussion of novel, knowledge-based electronic publishing technologies, products and services.

SMART CONTENT
We will take a closer look at the concept of smart content, beginning with a user perspective of the requirements and the ideal properties of this new type of content from this perspective. We then will move to a supply-oriented view on the smart content concept by building on a digital content value chain, pointing out the implications for publishers and other content providers in terms of supporting technologies and work processes.

SMART CONTENT – THE USER PERSPECTIVE
For the purpose of EP2010, we used two main sources to identify the future user requirements with regard to electronic publishing. The first one were the implications of the emerging knowledge economy we introduced and discussed in the previous section. Secondly we considered the objectives of KMCC who in their mission statement focus on "... technologies to support the design, creation, management and publishing of multimedia content, across fixed and mobile networks and devices, with the ability to self-adapt to user expectations. The aim is to stimulate the creation of rich interactive content for personalised broadcasting and advanced trusted media and entertainment applications.”

These objectives indicate a move towards embracing context sensitive adaptation of content, together with increasing levels of interactivity. Interactive content means "content + behaviour", where behaviour - computationally - means software programs. Adaptive content (as specified in the mission) means "content + dynamic, polymorphic, on-the-fly customisation".

The assessment of these objectives lead to the definition of a set of initial user requirements. The figure below outlines how the resulting “smart content properties” were classified in five groups: basic content-related properties; properties related to interfaces and interaction; knowledge-related properties; delivery-related properties; and properties related to the personal user environment.

169
FIGURE 1 – SMART CONTENT USER REQUIREMENTS AND PROPERTIES

EP2010 takes a closer look at these properties by identifying examples for the related technologies and establishes a ranking of the requirements. The aim is to get a better handle on the technical implications and to find out the weaknesses resulting from not having a particular property respectively the strengths expected from introducing this property. Further investigations should lead to a functional grouping of the requirements and finally to a “smart content service architecture”.

DIGITAL CONTENT VALUE CHAIN

From a user-centred view we move on to a supply-oriented view on the smart content concept and will now point out the implications for publishers and other content providers in terms of supporting technologies and work processes. We think that this approach should identify and strongly consider the value chains involved. For the purpose of EP2010, we used two different views on the digital content value chain: A technology-oriented view allows for identifying those technologies that support each of the production and delivery stages of smart content. And secondly, a process oriented-view reflects the workflow and the life-cycle of knowledge-enhanced objects in the publishing process.

TECHNOLOGY-ORIENTED VIEW

From a technology-oriented view, we can distinguish three major stages in the digital content value chain: content creation (raw content creation), content management (value adding, digital content management, transaction and customer management) and content delivery (distribution and delivery, content use). Associated with this value chain are various subsystems of the digital value chain management system and workflow support, e.g. content capturing tools, authoring and re-purposing tools, digital asset repositories, B2B and B2C systems for digital content, digital content delivery infrastructures (QoS networks, streaming media, etc.) and interactive digital content presentation environments.

The following figure outlines this technology-focused breakdown of the digital content value chain:
FIGURE 2 – INITIAL CONTENT VALUE CHAIN AND ATTENDANT TOOLS FOR DIGITAL CONTENT

The value chain and the technology clusters allow for the description of more intricate relationships (e.g. business models that generate profit within a single stage of the value chain, such as application service providers (ASP) for content management).

PROCESS-ORIENTED VIEW

Alternatively, the figure below provides a work flow-oriented perspective of the smart content concept and reflects the requirements of a knowledge-enhanced publishing process:

FIGURE 3 – COMMUNITY AND KNOWLEDGE ECONOMY ENABLES CONTENT LIFE-CYCLE

In our model, the initial step of the workflow, product planning (1), creates a template object, that includes a predefined knowledge structure (e.g. an ontology that was either existing somewhere on the market or developed especially for the purpose of the product), which is enriched by (multi-)media objects (2) and knowledge objects (3), e.g. attributed interrelations between media objects or portions of these objects.
Our knowledge-enhanced publishing object now is ready for editing (4) and aspects related to media production (5) and enters the cataloguing process (6). Following the creation of an publishing object that is ready for trading are steps that are close to the marketing and business scenarios: demos, previews and documentation are added (7) and the purchasing and transaction management adds trading records to our publishing object (8). When the product is delivered, whether personalised or pre-packaged, static or dynamic, adapted to the information space in which it was requested or not, the our publishing workflow proposes to add a further step, user community management (9), allows consumers to interact and give feedback and thereby further enhance the value of the publishing object, which then becomes either a “household” good or a commodity shared by a consumer community under the direction of publishers and authors. This workflow also allows for dissolving the traditional barriers between authors, publishers and consumers in the sense that these roles become deliberately and comprehensible exchangeable.

Consequently, we assume that each process step in the value chain adds a specific layer of information and knowledge to the final “publishing object”, however fugitive, dynamic and adaptable the final product will be.

**RESEARCH POLICY OPTIONS AND RECOMMENDATIONS**

For adopting the smart content concept as a future vision for “knowledge-based electronic publishing”, EP2010 proposes that a research map of knowledge-enhanced electronic publishing that contains three major research strands: foundational, application-oriented, and integrative research.

![Research Map Diagram]

**FIGURE 4 – RESEARCH MAP OF KNOWLEDGE MANAGEMENT AND CONTENT CREATION**

**FOUNDATIONAL RESEARCH**

The first group - foundational research – deals, for example, with methodologies for creating valid ontologies or with merging multiple overlapping ontologies. Foundational research will play a key role towards common understanding of the issues surrounding the relation of knowledge and content in the publishing sector: there is knowledge in content, but knowledge per se is not yet content. Themes to be covered in foundational research are likely to be: filtering and publishing of information according to knowledge profiles; descriptive query languages, that prevent the user from having to learn the underlying knowledge structures; the inclusion of natural language into interface techniques for publishing environments; etc.

**APPLICATION-ORIENTED RESEARCH**

The second group - application-oriented research - provides us with the key enabling technologies for rich knowledge media production, maintenance and business. This comprises of knowledge creation and maintenance, workflow support, and advanced media production techniques. It focuses on the integration of information-, content- and knowledge repositories as a key issue towards a knowledge economy. In particular, this research strand focuses on integrating and streamlining the following areas: (1) Hypertext – the emphasis is on associative clustering of content for optimised navigation. Standards in this area include HyTime and Topic Maps. (2) Databases – the emphasis is on the separation between conceptual structure (intension) and information items (extension). Standards include SQL for relational databases, ODMG-93 for object oriented databases. (3) Knowledge bases – the emphasis is on logic statements that can be proven true according to some model, expressed by inferencing rules. Standards include the knowledge interchange format (KIF) and conceptual graphs, but using these is the exception whereas the non-standardisation of knowledge based models remains the rule. (4) Document based information retrieval – the emphasis is on searching text corpora for the
purpose of relevance ranking of documents, with respect to a user query. Standards for document encoding and annotation include the text encoding initiative (TEI), Dublin Core (DC) for meta data, and other library standards. (5) Media repositories – the emphasis is on effective digital storage of archived audio-visual material, including cataloguing mechanisms using meta data standards. The major set of standards for the media industries are the JPEG and MPEG standards. The latter set of standards ranges from compression standards to workflow and business process definition (MPEG-21). (6) The World Wide Web – the emphasis is on access to distributed, federated, autonomous information sources that used to be tagged documents, but increasingly, are data aggregated to appear as tagged documents. Standards include XML, RDF + RDFS, and the Web Ontology Language OWL, for the representation of knowledge bases.

INTEGRATIVE RESEARCH

The third group - integrative research - deals with communicative functionality that users can expect from an ambient intelligence/smart content environment (at a knowledge level and interaction level!), and with semantic interoperability in the sense of intelligent multimedia middleware.

The smart content environment will have to provide sophisticated communicative capacities to address themes like the following one: Publishers may have masses of information and knowledge stored in huge repositories. The question is how are they informed of what is in the repository?

Regarding semantic interoperability, a clean architecture should be devised that allows “agents” to engage in meaningful interaction. There is nothing to be gained from translation between language layers that have similar semantics. Firm action on the part of the funding body may assist in getting to a consensus soon.

SUMMARY AND REFLECTION

Smart content as the unifying vision appears to be a good “condensation point” for several ideas and directions proposed at various consultation meetings and internal meetings of IST staff. For example, the Internal Reflection Group on Knowledge Technologies took as one of its guidelines, the following statement of ISTAG: “EU research should take a holistic approach to RTD in this area, putting innovative technology development and generic research issues to the fore but taking account of user needs in downstream application areas like education, culture or e-business.” (see [9] and [10]).

Our theses on the implications and challenges of the knowledge economy for the electronic publishing sector should at no point be misunderstood as the only way towards the future and prosperity of the publishing industries. There are business and revenue models that are well established and work well in the “traditional” publishing sector. We put forward the thesis that the smart content concept opens new avenues for publishers who decide to develop into knowledge providers, while the others – with good reason – remain what they are: magazine, book, etc. publishers. Knowledge based-publishing will provide a complementary approach to the established ways of electronic publishing.

It becomes evident, however, that the vision holds enough research potential in the area of foundational, application-oriented and integrative research and still requires a continuous dialogue between the industry and research community. The role of EP2010 will be to provide key orientations for policy makers and input to concrete implementation measures for the needs of the publishing industries towards 2010.

ACKNOWLEDGEMENTS

At this point the authors would like to thank the members of the steering committee as well as the members of the expert groups for the fruitful comments that brought forward the discussion process in defining the future perspectives for the electronic publishing sector.

We further would like to thank Ms. Andrea Mlarenin and Mr. Siegfried Reich from Salzburg Research for their input and comments on the paper.

Please note that EP2010 was commissioned by the Directorate-General Information Society and that the interim results presented don’t necessarily reflect the views of the European Commission.

NOTES AND REFERENCES


