CONNIE – CONSTRUCTION NEWS AND INFORMATION ELECTRONICALLY

Tomo Cerovsek[1], Gudni Gudnason[2], and Celson Lima[3]

ABSTRACT

The paper presents the current results achieved by the EU e-Content project CONNIE, which aims at delivering pan-European access to building regulations and standards through a network of information services dealing with content coming from 7 European countries. The CONNIE approach is original in several ways: it facilitates the use, the feedback and the exchange of building regulations; it is a network of decentralized portals with flexible conceptual architecture; it provides new B2B and B2C internet business models. The approach adopted to develop the CONNIE systems is based on the "CONNIE commons" (common data models, common APIs, common syntax) and IPR issues. Functionally, the services provided by the CONNIE system are divided into: core services, common added-value services, informative services, and descriptive services. First two support latter two that are regulation specific and provide improved regulation information retrieval and usage.

KEY WORDS

Building regulations, decentralized large scale portals, web services, information services

INTRODUCTION

One of the most important barriers to enter cross-border EU construction market are Building Regulations, which do have a crucial impact on the quality of the built environment. The importance of building regulations is illustrated with well known facts about the construction business environment: it employs around 20% of EU workforce, it consists of 97% of SME companies and most of the activities are project-oriented. Although EU building regulations are being harmonized, provisions can be conflicting, hard to track, and are error prone. Gray (1996) explains "Buildings are bigger, more complex, and contain more people than ever before. They are therefore more expensive and prone to bigger failures and greater loss than ever before. Society depends on their continued functionality and financial stability. Besides, the cost of error is too high." There are several services that provide access to normative documents such as www.iso.org, www.beuth.de, national sites of standardization institutes, or bibliographic services (i.e. www.perinorm.com). Neither they provide value-added complementary services, nor do they provide information on publicly available regulations.

[1] Dept. of Construction IT, University of Ljubljana, Faculty of Civil and Geodetic Engineering, Jamova 2, 1000 Ljubljana, Slovenia, Phone +386/1-4768-521, FAX +386/1-4259-681, Tomo.Cerovsek@fgg.uni-lj.si

[2] Icelandic Building Research Institute, Keldahnolti, 112 Reykjavik, Iceland., Phone +354-570-7332, FAX 354-570-7311, gudni@rabygg.is

[3] CSTB - Centre Scientifique et Technique du Bâtiment, Route des Lucioles BP209, 06904 Sophia Antipolis CEDEX Sophia Antipolis, France, Phone +33-1 93 95 67 22, FAX +33-4-93956733, c.lima@cstb.fr

MOTIVATION AND GOALS

The main three goals of CONNIE are: (1) to extract, organise and index building regulations and normative documents and related contents addressing AEC specific information needs, (2) to promote the deployment of an European Network where the Regulation/Standard-related contents are shared and exploited from a business perspective, (3) to design and implement the software infrastructure to support the abovementioned network.

METHODOLOGY

The iterative software development methodology used on complementary concepts including best practice software requirement specification and business process reengineering. During the development we have used the following modeling techniques: UML, IDEF0 and, in order to cover the transition stage from AS-IS to TO-BE, a Matrix of Change has been used, which was accompanied with SWOT analysis identifying positive and negative aspects of the systems. The approach covers domain specific analysis and as well as platform independent models, which later led to the architecture enabling collaborative parallel developments.

REQUIREMENTS

We provide a brief summary, for detailed requirement analysis see (Cerovsek et al 2005).

CONTENT, CONTEXT AND USERS

User needs are defined by the type of: (1) tasks - carrying out specific tasks with specific information needs, professional roles, (2) context - in which users will interact with the system, and (3) content - which will be available to the end users in a specific context.

Context is firstly defined with the construction business environment – as described in the introduction – was later assessed in more detail in relation to the use of information relevant to the development of CONNIE. An important factor defining the context is the technology environment that was analysed in surveys like (DTI 2005). The following IT factors are important for the CONNIE system: access to the internet in AEC, benefits of IT in the sector, Use of internet, and encouragement for IT, and trends in digital exchange.

Content is described through internationally recognized classifications of legislative and normative documents (i.e. Uniclass, and International Classification of Standards). Content was analyzed according to six major parameters: (1) Ownership and Access, where conventional copyright plays major problem; (2) Format – XML should play a more important role in the future; (3) Structure; (4) Metadata; (5) Volume, which is growing exponentially; and (6) Dynamic nature of building regulations.

Actors were analyzed and divided into seven categories: (1) Legislative powers; (2) Standardizing bodies,; (3) Building Construction Entities; (4) Support entities; (5) Resource suppliers; (6) Institutional entities; and (7) Educational entities. Through a detailed analysis of actors' roles, key needs are presented with corresponding activities related to building standards and regulations. In general, actors are divided into two main groups - producers and consumers with sometimes conflicting interests. Based on surveys, we define user needs that do not follow some common patterns and often use social networks. We describe the systems envisaged through user stories.

EXISTING AND TARGET PRACTICES

As a representative case study we illustrate regulations' life cycle of the ISO organization that has one of the most well established and effective procedures for the development of standards. The procedure is illustrated with following IDEF0 diagram presented in figure 1. The process of the development of standards is crucial for the success of a standard, especially in an international environment. The process starts with identified need, and goes through various changes to become an international standard.

The influence of the CONNIE system to the distribution and use of regulations is presented with bold arrows represent affecte ICOMs (ie. feedback, identification of needs, notification). Consequently, there is an evident improvement of all three major activities in the life-cycle of standards: development, distribution as well as professional use of standards.

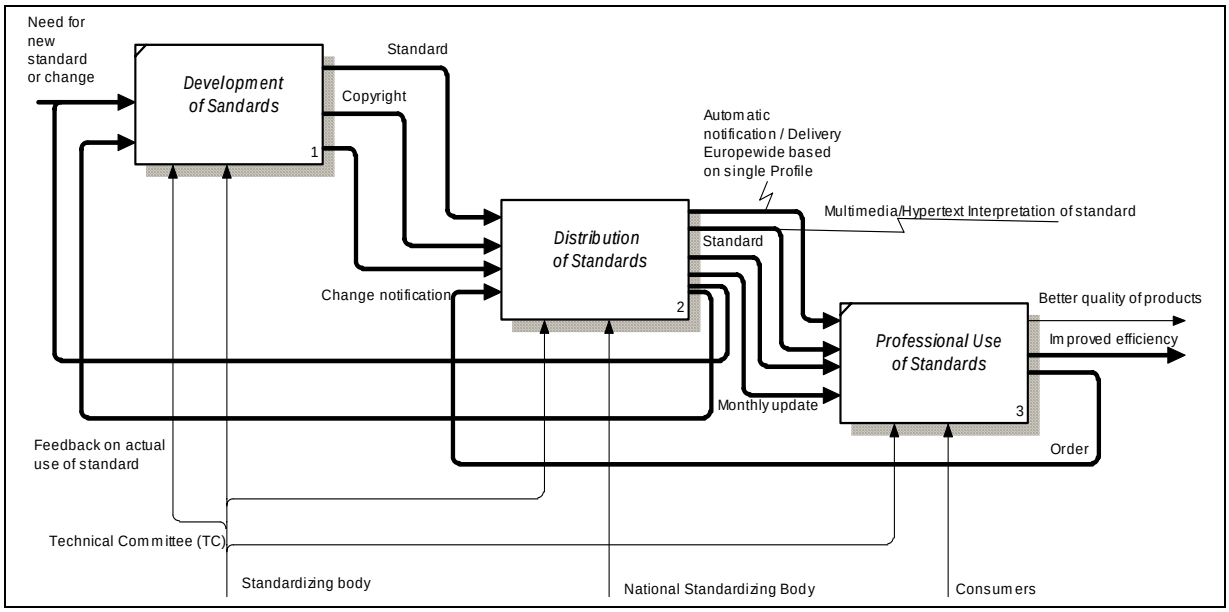

Figure 1: IDEF0 TO-BE model of CONNIE effects on the use of standards

REQUIREMENT SPECIFICATION

Software requirements are described with a set of functionalities resulting in 3 main types of services: basic services, informative services, descriptive services. Functionalities of use cases are classified according to effort, risk and priority for the development. The last section provides a description of representative services as well as detailed explanations of possible internet business models. Amongst the strengths and opportunities identified is the strong position of participating partners and the clear need identified by targeted end users participating in surveys. Based on identified requirements we have divided basic descriptions: (1) Core services; (2) Informative services; and (3) Descriptive services.

The following types of content will be available within the CONNIE network: directory of content sources, bibliographic information, indexed content through crawler and search engine and interfaces to existing data.

CONTENT MAPPING

Content mapping is divided into four layers:

- Content sources: points to where information comes from;
- Content: is primary and secondary legislation, directives, rules and regulations, and normative documents. ;
- Structure: metadata, classifications, taxonomies needed for further development; and
- Organization: addresses how the content is presented to the end user.

THE CONNIE APPROACH

Above described framework for content mapping in the planned system must be supported by appropriate measures, which can be characterized by:

- *CONNIE Commons.* Common types of users and Common context. The solutions are built around common data models common APIs, common syntax, common graphical user interfaces and common IPR (Intellectual Property Rights) approach.
- *Decentralized portals.* Each national node provides information on its own sources of information, users, content itself, application data and services that deal with that information.
- *Flexible conceptual architecture.* Architecture is SOA - Service-Oriented Architecture where all services are defined using a description language and have invokable interfaces used for aggregation and integration.
- *Development and interoperability.* CONNIE is network of EU services that can communicate between each other using common APIs and/or can be accessed through national portals. Networking of nodes is based on services that has three distinctive properties: platform-independence; dynamic invocation; self-sufficiency. Web services provide interface for the use of CONNIE services.

Similar approach was also used for the design of the network of information portals in the CONNET project, which was designed as a set of loosely coupled nodes, operated by the different partners in the project. The idea was to integrate the nodes on both a semantic and technical levels (Turk et al 2001). This idea was taken a step forward as explained below.

CONNNIE COMMONS

CONNIE should provide targeted information through a network of services. Following five main models were adopted as a framework for system development

Common Data Model – Since there are a lot of services that are based on similar data items and deliver similar results, there is a natural need to develop common data models that will be used throughout the services. In the early days of mainframe computer systems, it was common to envision an enterprise-wide "management information system" that mandated a common data model applied to all enterprise information systems. This approach is less stringent than subordinating all systems into some master, all-encompassing system, but it still does require central administration of an abstract and complex model shared by all

interacting systems. The attempt to map this common data model into central portal was tempting, but was not shown in the practice to be successful. Since it does not allow different approach and is not prone to adapt to different needs and capabilities.

Common Applications Programming Interface (API) – The CONNIE Application Programming Interface (API) plays a significant role in operation of the CONNIE network. Firstly, it is common to all the national nodes that operate in the CONNIE network so the same functionality can be expected from each independently operated national node. Secondly, it enables the CONNIE national nodes to share and exchange information and appear seamless to the end-user as a single virtual service. Thirdly, it enables communication with trusted partners and/or third parties to invisibly re-use information provided by the CONNIE services. The CONNIE architecture and API exposes much of the CONNIE solution functionality using a layered approach and as core and optional services enabling flexibility in its implementation.

Common syntax/user interface. The CONNIE system is oriented towards improving information retrieval (IR) of building regulations as well as in the ways the content is presented to the end user. Interfaces represent the bases for the human-computer-interaction and therefore represent an important aspect of the successful system. The goal was to define both: (1) specialized syntax for regulations retrieval, and (2) a set of reusable interfaces that could be used as templates for typical content. An important aspect of the CONNIE systems is searching, which must address two types of interfaces for SQL and IR queries. Different approaches are needed since the first is oriented towards database querying resulting in tables matching specific criteria, and IR querying of large collections delivering ranked results. The latter is usually used format for all major public and desktop search engines. The goal of the syntax is to cover specific needs that are related to building regulations and the context in which they are used.

Common IPR model. CONNIE content IPR is based on CreativeCommons.org, which enables copyright holders to grant some of their rights to the public while retaining others through a variety of licensing and contract schemes including dedication to the public domain or open content licensing terms. The intention is to avoid the problems current copyright laws create for the information sharing. The project provides several free licenses that copyright holders can use when releasing their works on the web, which can also be applied to building regulations related material. They also provide RDF/XML license metadata, that makes it easier to automatically process and locate licensed works.

NETWORKING INFORMATON PORTALS

Decentralized Portals for the CONNIE network where seen as the only feasible solution since the idea of centralized portal for all EU regulations is unrealistic due to the following facts: applicability of laws and standards is geographically dependent, interfaces to several sources are required, partner institutions already have their own user accounts, specific content, and have established business and legal arrangements with sources of regulative information, as well as have invested into technical solutions for content delivery. These are only some indicative requirements for the interoperability of nodes forming the network enabling effective communication among different national sources where each node can communicate with other nodes and can establish flexible B2B business arrangements.

CONCEPTUAL ARCHITECTURE

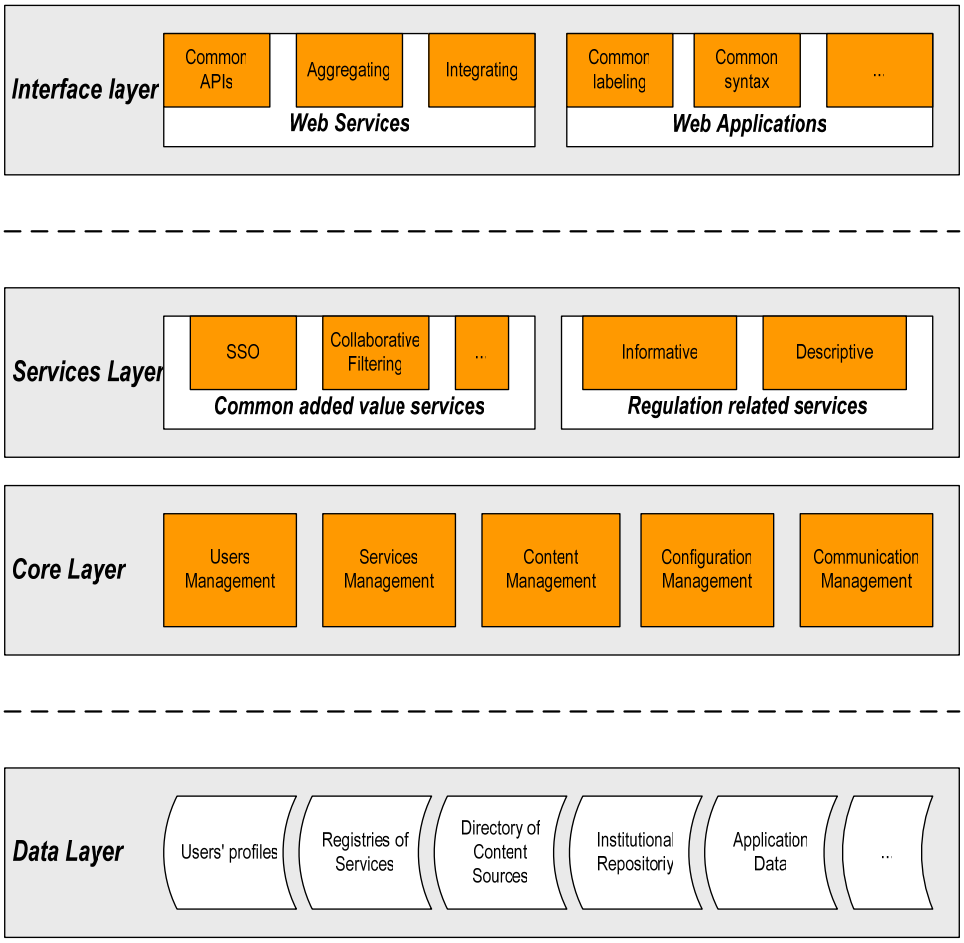

Figure 2: CONNIE conceptual architecture

CONNIE conceptual architecture provides a flexible framework. The architecture is divided into four layers grouping corresponding functional units: data layer addressing data storage, core layer establishing CONNIE network, services layer dealing with regulation handling and interface layer providing common interfaces:

- **Data layer:** contains databases, which are based on CONNIE common data models.
- **Core layer:** contains basic services related to general use of the CONNIE system and are not explicitly dedicated to the use with building regulations (they enable management of users, data, services, content, etc.).
- **Services Layer:** is related to the use of regulations where each national portal is independent but through common added value services such as single–sign–on (SSO) and collaborative filtering CONNIE enables new business models and enhanced the use of

building regulations. Informative and descriptive services both deal with regulations, but act on the content differently - only the latter are used to interpret the content.

- **Informative services:** provide information about regulations and include IR techniques (search mechanisms, meta-search, news and change notification, clustering, thesauri search, etc.).
- **Descriptive services:** describe/extract/interpret the content embedded in regulations (online learning, multimedia presentations, process models interpretation, etc.).
- **Interface layer:** Users are able to access information through personalized portals, specialized client applications or can use push technologies. Once subscribed at any national node, users are able to access services at any CONNIE node using SSO. CONNIE will provide templates, unified labelling and syntax for its Web applications.

DEVELOPMENT AND INTEROPERABILITY ISSUES

The CONNIE system enables quick adoption of new solutions. In order to achieve that the following criteria was the bases for the selection of a programming environment:

- Operating system independence: Supporting more Web Servers (IIS, apache) and Scripting Languages that will be run at server side. PHP was selected as preferred scripting language due to platform independence, support for databases and other factors.
- Database engine: support for open source and commercial SQL compliant engines.
- Application server: Strong support from the development and end user community - if possible from open source community.
- Open APIs for syndication of content from nodes. This criteria together with previous three addressing platform independence, SQL and application lead to selection of mambo CMS open source system as platform for content management..

The technical implementation development process is divided into following phases:

- Rapid prototyping: We have used several tools that enable rapid prototyping of specific network services and components in order to demonstrate possible functionality and population with data for testing purposes we used langueages such as perl, jsp and php.
- Refinement of data models, APIs and interface functionality: Through rapid prototyping and testing we refine data models, APIs and functionality.
- Development of final production components: After we have agreed on common data models and prototypes we will develop final production components.
- Implementation at local nodes: Stable components are deployed at individual sites.

The work was carried in two phases: (1^{st} phase) establishing infrastructure for CONNIE network with core services and the development of the most important services, (2^{nd} phase) developing and extending common and national CONNIE node' services.

CONNIE SERVICES AND COMPONENTS

CONNIE nodes can communicate amongst themselves through web services (fig. 4). A local CONNIE node integrates local services as well as is communicating with remote services.

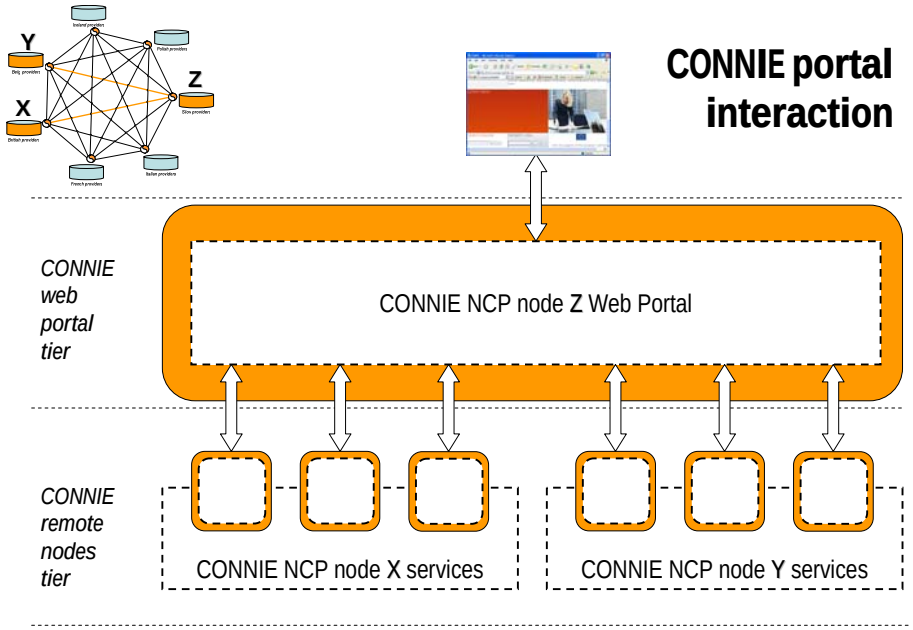

Figure 3: CONNIE portal to portal communication

CORE SERVICES

A conceptual and technical descriptions are provided for all type of services. The following Core services have been developed in the first phase for set up of CONNIE network:

- CMS: Content management system – (CMS) was selected and set up covering the CONNIE needs supporting the creation, management, distribution, publishing and retrieval of information. The role of CMS is also to serve as a platform for distribution and development as well as reuse of components such as FAQs, records management, subscription, etc. that each can configure into the portal as needed. It is also supported with Web Services technology to integrate with external applications.
- CNC: Connie network connector which supports the creation, configuration, and operation of the CONNIE network. The CNC handles the configuration of nodes and their respective CONNIE services. It also checks who is alive within the network, it registers itself into the network (send a status messages to the other CONNIE services provider), and offers a mechanism to trigger the need of updating amongst nodes.
- DCS: Directory of content sources provides multilingual Dublin Core and Connie specific meta (about) descriptions of sources of content relevant to the CONNIE project It is a web database application that enables maintenance, exchange and propagation of Content Sources' Metadata, but not content itself. It is used as a source for the crawler and other CONNIE services.

For the prototype development we have selected the Mambo CMS (www.mamboserver.com) version 4.5.2. In the project the Mambo CMS has been implemented under different platforms and web servers e.g. Apache on Debian Linux and Windows IIS and Apache on MS Server 2003. Several prototypes have been developed in PERL as well as in PHP and JSP. The CONNIE Network connector was implemented in php, and has an open API. Network nodes are connected through XML-RPC calls. The component was developed as a standalone application as well as built and deployed as a standard CMS component. DCS has a web interface as well as an open XML access with API described by a WSDL.

COMMON VALUE ADDED SERVICES

These services provide added value in relation to the use of all services from different nodes in the network that deal with building regulations.

- Single-Sign-On (SSO): Users will be able to sign to only one node, and will be able to access all nodes from the network. SSO is not only used for inter-node communication, but also enables individual nodes to make agreements with organizations, and allow them to access a CONNIE node using theirs existing active directory user accounts.
- Profiling and Collaborative filtering: Profiling provides a mechanism that enables gathering of information about users' preferences, professional background as well as stores information on personalized GUI settings. Information stored in profiles, together with the systems' usage analysis we will also enable cross-border collaboration filtering.

SSO was based on Shibboleth, which is and open source project of Internet2/MACE providing robust architectures, policy structures, practical technologies, and implementation supporting inter-institutional access to restricted resources (http://shibboleth.internet2.edu/). It is based on standards like SAML (Security Assertion Markup Lanugage), and provides mechanism for federated attribute exchange. We have also established a conceptual framework for personalization that will be further enhanced with collaborative filtering.

INFORMATIVE SERVICES

These services provide meta-information on bregulations, reprepresentative examples are:

- Search engine: consiting of several tools including specialized: (1) Crawler with online administration scheduler, deep searching, synchronization with directory of content sourcs; (2) Indexer with specific "regulation-code" indexing; and (3) Interfaces with specific syntax and query extensions such as clustering, summarization, ontology and thesaurus-based query expansion.
- News service: news will be extracted from domain-specific relevant international, European and national culling mechanisms for indexing daily news from carefully selected sources with browsing, searching, alerting capabilities, as well as aggregation.

Connie search engine was developed based on the Nutch open source search engine with a specialized crawler as well as a complete pluggable php port of Nutch search engine has been developed to be used within the CMS environment.

DESCRIPTIVE SERVICES

Descriptive services provide following major types of applications:

- Worked examples: these are case based studies - online learning practical material and would provide worked examples that are based on specific building codes and standards.
- Rule based interpretations: Semi-automatic extraction of rules – detection of conflicts, Rule based interpretation of standards different, Ontology building, which provide Ontology enhanced information retrieval, Semiautomatic checking of IFC models.
- Process model interpreted standards: Interpretation of standard with process model depends on type and size of standard. It can be carried out using different modelling techniques such as IDEF0, flowcharts or workflow models.
- Hypertext and Multimedia interpretation of standard Provide multimedia interpretation of standard Automatic linking of content – interpretation of contents with thesaurus enriched content, Linkage to related material, and Multimedia representation of content.

LOM will be used to describe learning material, including multimedia, process models, hypertext representations of standards with features like automatic linking mechanisms (Turk & Cerovsek 1998), rule based interpretations (Lau & Law 2004), other based on RuleML, OWL, or checking methods as demonstrated on IFC model (Pascual et al 2002).

CONCLUSION AND OUTLOOK

The results presented in this paper cover the first phase of the project. The approach was proven as very effective, and shows a great potential for informative and descriptive services.

ACKNOWLEDGMENTS

The CONNIE Consortium would like to acknowledge the financial support of the European Commission under the eContent programme. We would like to acknowledge the contribution from partners and valuable suggestions from interviewees in participating countries.

REFERENCES

Cerovsek, T, Gudnason, G., Storer, G., Bloomfield, D., and Padula, M. (2005). CONNIE D12-D14: User requirements and software specification.

Gray, C. (1996). Value for Money. Copyright ©2001 - 2002: Reading Construction Forum.

DTI (2005). Survey of IT in construction use, intentions and aspirations. The IT Construction Forum. Constructing Excellence. www.itconstructionforum.org.uk

Lau, G. & Law, K. (2004). An Information Infrastructure for Comparing Accessibility Regulations and Related Information from Multiple Sources. ICCBE Weimar, 2004.

Pascual, E., Vankeisbelck, R., Zarli, A. (2002). Knowledge based checking of building designs, based on digital regulations. In: Turk, Z. & Scherer, R.J. (eds.). Proceedings Of Fourth ECPPM '02 , Portorož/Slovenia/, 9-11 Sep. 2002, pg. 467-464.

Turk, Z., Cerovsek, T. and Amor, R. (2000) CONNET - Design Issues, Proceedings of INCITE 2000, Hong Kong, 17-18 January, pp. 416-428.

Turk, Z., Cerovsek, T., Duhovnik, J. Regulation Processing In Tocee. In: Amor, R. (ed.): Proceedings Of Second ECPPM '98; BRE, UK, 19-21 October 1998. Garston: Building Research Establishment, Cop. 1998, pg. 523-532.