Online Product Libraries: The State-of-the-art

Robert Amor[1], Shailesh Jain[2] and Godfried Augenbroe[3]

ABSTRACT

The movement of product catalogues from paper-based systems to electronic and online systems has proceeded rapidly over the last decade. The commercial product catalogue vendors have established nationally-based online systems of varying sophistication at the same time as classification and standardization producing organisations have developed their national and international support systems. There have also been independent initiatives from major CAD vendors to enable manufactured product data and CAD detailing integration with their design tools. Underpinning much of this work has been a suite of research projects addressing infrastructures which will better support access to product information of all types for all classes of users and their design tools. This paper surveys the state-of-the-art in this domain across all these efforts as a way of assessing where we have got to. It also analyses the gaps which exist between expectations and actuality in order to identify the gaps in research and practice which need to be bridged to enable effective access to, and interaction with, this most important resource in construction activities across all stages of a building's life.

INTRODUCTION

Product information is an integral part of almost every process in the life of a building. It is required for some design decisions, through detailing, simulation, specifying, construction, maintenance and even demolition and recycling. To serve such a wide range of uses the data supplied for each product consists of many types of information, from measured properties, best-practice installation and use guides, code conformance certificates, CAD details, etc. Collating this information for each product is a task for manufacturers and suppliers of particular products and this information is usually available from manufacturers as part of their catalogue of products. To ensure that practitioners know about their products manufacturers and suppliers advertise their products in appropriate trade literature and employ representatives to visit firms and tout their goods. With thousands of manufacturers and suppliers in any single country this is an inefficient method of getting information about products across to practitioners. To increase a practitioner's access to a wide range of products, and to help manufacturers and suppliers be seen by a greater proportion of the market, a range of organizations have developed product catalogues consisting of many thousands of manufactured products indexed and categorized to enable quick identification.

Originally, these catalogues existed in paper form, many of them of prodigious size. Browsing by classification codes allowed users to identify a generic category of product (e.g., aluminium windows) and within this category they could scan the product sheets of a wide variety of manufacturers to identify products potentially usable in a project. This is not a particularly efficient process and requires much human interpretation of specification sheets to determine the suitability of a product. Technical details of products would be transcribed for use in a range of processes, such as simulation, costing, and detailed design. This manual process introduces a pathway for transcription errors which may impact the later working of the building.

[1] R. Amor is an Associate Professor, Department of Computer Science, University of Auckland, Auckland, New Zealand, email: trebor@cs.auckland.ac.nz

[2] S. Jain is a PhD Candidate, College of Architecture, Georgia Institute of Technology, Atlanta, USA, email: gt8015a@mail.gatech.edu

[3] G. Augenbroe is a Professor, College of Architecture, Georgia Institute of Technology, Atlanta, USA, email: fried.augenbroe@arch.gatech.edu

The introduction of electronic versions of these catalogues brought the hope that many of the inefficiencies, inaccuracies and laborious transcription processes would disappear. However, the majority of electronic catalogues, even today, have duplicated the paper paradigm of the original catalogues which requires human interpretation and transcription of information from the catalogue to other tools. A user can more quickly identify products by following the classification hierarchy and in many cases they can access digital documents representing design details, installation guides, certificates, etc which may be copied into a project's information store. Yet, the parameterized information about the products which finally impacts whether they will be selected for a project must still be manually interpreted.

The remainder of this paper looks at research and commercial work which aims to solve the many problems which still exist in collating and accessing manufactured product information for intelligent use in any of the processes which require manufactured product information. This description of projects in this area circumscribes the state-of-play in this area as seen from the many viewpoints of those working on the problem. It also identifies several areas where there are still no good solutions to enable the intelligent use of manufactured product information in our industries.

RESEARCH PROJECTS

In this section we examine research initiatives grouped around the development of electronic product catalogues ranging from the work of individual researchers through to large-scale European-funded research projects. All of these projects are based on the supposition that classifications are not a sufficient access mechanism to product information and hence that product attributes and various associated information resources must also be accessible (initiatives are listed in alphabetic order).

ARROW

The ARROW project (Amor and Newnham 1999, Newnham and Amor 1998) ran for three years funded by the UK PiT program. The main contributions of this project were: an open infrastructure to allow manufacturers to describe the differentiating attributes of their products and to upload any type of electronic documentation associated with the product; a methodology to link structured product information with unstructured textual information from web pages and documents for a more comprehensive search functionality; a demonstration of functionality required to allow CAD systems to automatically incorporate product details into a CAD drawing; and an exploration of simplified methods to allow manufacturers to provide structured information on their products.

CONNET-MPS

The CONNET project (Amor et al 2000) ran for one year funded by the EU. The main contributions of this project were: a scalable data model to allow catalogue producers to migrate from current paper-equivalent online systems to complex attribute driven services; and the incorporation of a services-based infrastructure to supply basic support services to new catalogue producers.

Eindhoven University of Technology

Research at Eindhoven University of Technology (van Leeuwen and Fridqvist 2002) contributes information on: a feature-based modeling approach for product representation and hence the support for dynamic specification of differentiating features for products; and the management of product information across firms, projects and in a more global context.

GEN Projects

PROCAT-GEN (Cook et al 1999, Faux et al 1998) and GENIAL (Debras 2000; Faux et al 1998) ran for three years funded by the EU. The main contributions of these projects were: the development of an open, XML-based, data model for classifications and attributes of products; provision of manufacturer and supplier-side tools to enable searching for products and services; the support for user-based views and profiles of subsets of information within the product

catalogue; and the support for attribute-based control of information modification privileges to particular classes of users.

Georgia Institute of Technology

Research at Georgia Institute of Technology (Jain and Augenbroe 2003, Augenbroe 1998) contributes information on: using the performance concept to facilitate product selection from product catalogues; a methodology for ranking products by substituting them in pre-defined parametric simulations that assess the performance of a selected product in a user defined design context; and a new methodology for ranking products in online searches to augment existing search methodologies based on keywords and product attributes.

International Alliance for Interoperability

Researchers at BRE involved with the IFC development (Nyambayo et al 2000) have contributed information on: the usability of the IFC 2.0 data model for representing manufactured product catalogues.

Loughborough University

Research at Loughborough University (Obonyo et al 2001) contributes information on: an agent-based approach to gathering and querying product data from XML-based repositories; a method to interpret PDF documents for structured data; and an agent-based automated purchase negotiation system.

RINET

The RINET project (RINET 2000) ran for three years funded by the Finnish VERA program. The main contributions of this project were: an open infrastructure to allow manufacturers to describe the attributes of their products and to upload any type of electronic documentation associated with the product; and a methodology to link structured product information with unstructured textual information from web pages and documents for a more comprehensive search functionality.

University of Auckland

Research at the University of Auckland (Amor and Kloep 2003, Cope and Amor 2002) contributes information on: the ability of WWW-based web-service developments to support open access to manufacturer's product information; and what supporting web-services are required to enable parameter-based product identification.

University of Edinburgh

Research at the University of Edinburgh (Ofluoglu 2003, Coyne et al 2001) contributes information on: designer side support services for interacting with online product information (e.g., a virtual sample board); and the management of product information across firms, projects and in a more global context.

COMMERCIAL SYSTEMS

In this section we examine a representative range of commercial systems which have developed online product catalogues beyond the basic duplication of paper-based systems driven by classification codes.

Classification Driven Services

The vast majority of online services operate on a model where access to manufactured product information is driven by classification codes, manufacturer names, word search, and perhaps availability of specific information types (e.g., First Source Exchange http://www.firstsourceonl.com/, Sweets http://sweets.construction.com/, Thomas Register http://www.thomasregister.com/, etc). These services provide information on manufacturers of identified products along with access to CAD details, PDFs of technical information and other

associated information types required within a national context. In some cases there is a basic ability to compare products identified by the classification search on a small set of parameters.

Parameterised and Classified Services

A very small number of online services extend the functionality seen in classification driven systems, as described above, by providing the infrastructure to identify products through a wider set of parameters. For example, Building Product Expert (http://bpe.barbour-index.co.uk/bpe/) offers a large number of refining parameters (e.g., 'Fire protection', 'Material: Aluminium', etc) to be assessed alongside the product classification. While this does not allow numeric assessment of the parameters it does provide comparison beyond what is available by general classifications. Edibatec (http://www.edibatec.org/) also defines a large number of refining parameters for its products and allows numeric values to be specified against these parameters.

CAD SYSTEM APPROACHES

In this section we examine the approaches taken by the major CAD vendors to support access to product information within a CAD environment (CAD systems are listed in alphabetic order).

ArchiCAD

ArchiCAD (Graphisoft 2003) focuses its efforts for access to product information on the GDL scripting language provided with the CAD system. GDL allows parameterised objects to be created for use within the CAD tool. These parameterised objects can be published online and, with the appropriate add-ins, be accessed from a website repository, effectively enabling catalogues of products to be created and delivered. At GDLCentral (http://www.gdlcentral.com/) a number of manufacturers have made available GDL driven representations of their products for inclusion within ArchiCAD or Autocad projects as an example of how users may incorporate such manufactured products into their designs. Selection of products is through whatever navigation is supported on the aggregating website, but not on attributes of the parameterised objects.

Architectural Desktop

Architectural Desktop (Autodesk 2003) focuses its efforts for access to product information on the block mechanism within AutoCAD as a way to aggregate the graphical representation of a product into a single entity. With their internal i-drop® technology (http://www.autodesk.com/) it is possible to publish product information to a web repository and make the resultant product descriptions available in the form of manufactured product details which can be drag-and-dropped into an AutoCAD drawing from a website. Selection of products is through whatever navigation is supported on the aggregating website.

CLASSIFICATION SYSTEMS

Classification systems have been used as the major access and filing method for project information for decades. For example, CI/SfB (Ray-Jones and Clegg 1991) was commonly used in English speaking countries as a standard to classify manufactured product information from manufacturers as well as for catalogues. More recently replacements for older style classification systems have been developed to address restrictions inherent in the older systems. Systems such as EPIC (1999) in Europe, Uniclass (1997) in the UK, CSI MasterFormat (2003) in the USA have been developed to provide multi-table hierarchical classifications of products as well as addressing the anomalies which occurred with previous systems such as CI/SfB. These classification systems tend to focus on the needs of a single country and their usage tends to be confined to that country. In contrast UNSPSC (2003), the United Nations Standard Product and Services Classification, was developed as a globally applicable system allowing suppliers, electronic catalogue aggregators and e-procurement solutions providers worldwide to adopt a universal product classification scheme for publishing product information into electronic catalogues. However, there has been little take-up of UNSPSC for building product catalogues or by manufacturers to date.

An alternate approach to enabling interoperability between classification systems is the LexiCon project (Woestenenk 2002) which aims to codify all commonly used classification systems and describe the relationships between terms within those systems to build a taxonomy of construction terms. If achieved this would allow automated translation between terms in any classification systems, and automated language translation for words used for those terms.

STANDARDS

In this section we examine the major standardisation efforts which allow for the semantic description of manufactured products and the transfer of product information between systems (standards are listed in alphabetic order).

IAI-IFC

IFC (2003) is an international standard for the representation of construction information suitable for transfer between applications commonly used in the construction industry. The most recent version of the IFC standard, IFC 2.x.2 comprises over five hundred classes of objects commonly utilised within construction projects. Specified classes are structured into several levels from common resources used throughout a project to domain specific resources used for a particular process within a construction project. Classes are not defined for every conceivable product type within a constructed entity, but represent generic categories of element (e.g., wall, beam, space) which can be categorised to provide a unique specification of the actual type of element. To this extent there are very few attributes associated with a class which could be used to transfer information relevant to a manufacturer.

However, the IFC incorporate a mechanism called Property Sets (PSets) which allow information publishers to dynamically allocate new properties to an object they wish to describe. The assumption is that groups of interested parties will agree a common set of these properties sufficient for the transfer of information within their field. Indeed, a large number of these commonly agreed properties are listed in the IFC specification. PSets have many of the disadvantages of the attributes in PLIB in that there is no specification of the semantics of PSet information outside that published in the IFC distribution.

The IAI have a development team looking at support for electronic catalogues for IFC 3, led by some of the existing electronic catalogue producers.

ISO-PLIB

ISO 13584 (1995) Parts Library (PLIB) is a series of International Standards for the computer-readable representation and exchange of parts library data. The objective of the standard is to provide a mechanism capable of transferring parts library data, independent of any application. The standard specifies the structure of a library system that provides an unambiguous representation and exchange of computer interpretable parts library information.

Attributes in PLIB are used to characterize any kind of property of any kind of entity that belongs within the problem area of the products computer model. Three levels of attributes are defined. Level 1 attributes are characteristics of a part whose value can be inferred from the part identifier. Level 2 attributes are characteristics of a part once the part is chosen, i.e. attributes that are specific to the instance of a part. A Level 2 attribute may itself be a part. Level 3 attributes may also be a structured list, thus allowing higher-level assembled parts. There is no structure or rules to ensure that all manufacturers use the same attributes to identify properties of similar products.

Unlike the IFC that is a data exchange model, PLIB offers a model for creating dictionaries of components. PLIB implementations are typically manufacturer specific; therefore, it makes it difficult to compare products from different manufacturers unless they have the same terminology and provide implementations of the same context parameters. In PLIB, the context knowledge necessary to evaluate performance of products is embedded within the catalogue.

BARRIERS TO THE DEVELOPMENT OF ONLINE PRODUCT LIBRARIES

The preceding survey of work related to electronic product catalogues highlights the wide scope of technology developments which are required to support the successful deployment of electronic product catalogues. From the survey of projects and commercial systems developed it seems clear that there are few technical barriers to the construction of online product libraries which allow the identification of manufactured products through a range of search mechanisms, including product parameters. If this is not an issue in the development of such systems then we must consider where the difficulties currently lie. We posit the following issues as remaining barriers to the uptake and development of online product libraries.

- *Products' data representation:* This is a major issue for the further uptake of product catalogues. Over ten thousand distinct categories of manufactured product are estimated to be available in any one country. In order for these products to be properly identified by search systems there has to be a set of parameters associated with each of these products. There seems to be a tension between the development of nationally or internationally agreed data models for individual products which is likely to take decades to complete and the non-structured approaches which allow for dynamic specification of properties for products. It seems clear that there must be a minimum level of structure associated with each product type in order that applications have some ability to perform automated identification and selection of a particular product, otherwise tools have no starting point for accessing information (or have to be very intelligent to reason over the semantics of a product's definition). The consensus point between fully specified and totally unspecified data models still needs to be decided in this domain. Followed by a huge effort to define properties for the products.
- *Extraction of data from manufacturers:* The manufacturers product specification sheets hold a large amount of information which needs to be structured in order to support parameterised identification of products. The effort of structuring this data, which, in most cases, is in an unstructured textual format is a major cost for manufacturers or product catalogue producers. Current approaches such as parameter recognition from PDFs and provision of templates to manufacturers do not provide an accurate or complete solution to this problem. Issues impinging on this problem include standard representations for products (so manufacturers do not have to describe their products in many different formats for each catalogue producer) and the business case to support the extra effort being made. The ability to incorporate and support differentiating attributes from different manufacturers for particular classes of products is also an imperative from the manufacturers' point of view.
- *Business models for online catalogues:* The openness of the Internet creates a change to the business models which predominated in paper-based catalogues. The ubiquitous nature of search on the Internet reduces some of the need for catalogues, as users can easily identify many products without recourse to online catalogues. The business case to attract users must provide greater benefits to a user (such as automated identification of products based on parameters, associated documentation, etc) and also provide greater benefit to the manufacturers to justify their greater effort in providing the structured information required in these systems.
- *Contextual impact of project requirements:* In the same way that classifications are insufficient for the identification of a specific product, the raw parametric data for a manufactured product is insufficient to determine the applicability of a product within a particular design situation. There needs to be further research into automated methods to enable the design context to be translated into the identification of suitable products, otherwise catalogue systems will still require inordinate human interpretation and decision making for every selection of a product for a building project.
- *Mapping between representations:* The development of open markets and fast transportation between countries means that product selection and purchasing is enacted in an international market. To make product selection through catalogues of greater benefit to users it is clear that interaction with foreign manufacturers and even other national catalogue services will take place. As classification systems tend to be nationally based and product data structures specific to individual catalogues or manufacturers it is

clear that some level of mapping between various classification systems and product data representations will have to be supported. Projects such as LexiCon (Woestenenk 2002) may provide a solution to the classification problem in the fullness of time, but the issue of mapping between alternate data representations of the same manufactured product type will require further research.

- *Support for preferences:* Analysis of current information processes within firms and projects shows that there are many levels of management of preferences for product types and preferred manufacturers. This ranges from personal preferences, through to a firm's relationship with individual manufacturers, decisions made within particular project teams, and even national constraints. In order that catalogues match the processes undertaken by users within projects there has to be support for the way that these preferences are applied to product retrieval within the tools offering such services. Several research projects have started tackling this problem though it is not clear that a viable solution has yet been found that will work with the major tools used by practitioners.
- *Life-cycle support:* The majority of the research and commercial effort in the development of access to manufactured product information has concentrated on the design and specification processes within a project. While it is clear that this is where purchasing decisions are made, and hence a business case is available for supporting these processes, it is also clear that manufactured product information is required throughout the life of a building and support functions need to be offered to enable interaction with product information many years after the initial selection of a product was made.

CONCLUSIONS

This paper shows the wide ranging efforts which are being made to support online access to manufactured product data from a wide range of interested parties. The research projects and commercial offerings indicate that the selection of manufactured products through classifications and product parameters is a solved problem. Similarly, the ability to associate various information types to a product (e.g., specification sheets, CAD details, compliance certificates, etc) is also widely supported. However, this survey of projects identifies many issues which are still barriers to the widespread uptake and commercialisation of online manufactured product data for use within projects. Reviewing these barriers it appears that many of them are associated with the development of business cases for the flows of information that we expect to make these services viable. Until these barriers are overcome we can expect our access to powerful online identification and incorporation of manufactured product information to be severely constrained.

REFERENCES

Amor, R. and Kloep, W. 2003. E-Product Catalogues, Proceedings of the EIA9 Conference on E-Activities and Intelligent Support in Design and the Built Environment, Istanbul, Turkey, 8-10 October.

Amor, R., Finne, C., Turk, Z. and Hyvarinen, J. 2000. CONNET: CONstruction information service NETwork, Final report, EC-ETTN-Lot 4, Contract 501999, March, 71pp.

Amor, R. and Newnham, L. 1999. CAD Interfaces to the ARROW Manufactured Product Server, Proceedings of CAADfutures'99, Atlanta, USA, 7-8 June, pp. 1-12.

Augenbroe, G. 1998. Building Product Information Technology, Executive white paper, Georgia Institute of Technology, http://www.arch.gatech.edu/crc/ProductInfo/

Autodesk 2003. Architectural Desktop http://www.autodesk.com/

Cook, G., Czubayko, R., Klemme, R., 1999. PROCAT-GEN – Conformance Specification, http://www.procat-gen.org/

Cope, G. and Amor, R.W. 2002. UDDI for a Manufactured Product Brokering Service, Proceedings of the EC-PPM Conference on eWork and eBusiness in AEC, Portoroz, Slovenia, 9-11 September, pp. 603-608.

Coyne, R., Lee, J., Duncan, D. and Ofluoglu, S. 2001. Applying Web-Based Product Libraries, Automation in Construction, Vol 10, No 5, pp. 549-559.

CSI MasterFormat 2003. Construction Specifications Institute, http://www.csinet.org/

Debras, P. 2000. Construction Application of a GEN-Network: Uniform Access to Standards, Products and Company Information, Proceedings of CIB W78 Construction Information Technology 2000, Reykjavik, Iceland, 28-30 June.

EPIC 1999. Electronic Product Information Co-operation, http://www.epicproducts.org/

Faux, I., Radeke, E., Stewing, F.-J., van der Broek, G., Kesteloot, P. and Sabin, A. 1998. Intelligent Access, Publishing, and Collaboration in the Global Engineering Networking, Computer Networks and ISDN Systems, Vol 30, No 13, August, http://www.procat-gen.org/

Graphisoft 2003. http://www.graphisoft.com/

IFC 2003. International Alliance for Interoperability, http://www.iai-international.org/

ISO 13584 (1995) ISO CD 13584-1, Industrial Automation Systems and Integration, Parts Library, http://www.plib.ensma.fr/

Jain, S. and Augenbroe, G. 2003. A Methodology for Supporting Product Selection from E-Catalogues, ITcon, Vol. 8, http://www.itcon.org/2003/27/, pp. 381-396.

Newnham, L. and Amor, R. 1998. Translation of Manufacturer's Product Data for the ARROW Product Search System, Proceedings of 2nd European Conference on Product and Process Modelling, London, UK, 19-21 October, pp. 405-412.

Nyambayo, J., Amor, R., Faraj, I. and Wix, J. 2000. External Product Library - An Implementation of the Industry Foundation Classes Release 2.0 Model, Proceedings of Product Data Technology Europe 2000, Noordwijk, The Netherlands, 2-5 May, pp. 147-156.

Obonyo, E.A., Anumba C.J., Thorpe, A. and Parkes B. 2001. Specification and Procurement of Construction Products: Potential Role of Intelligent Agents, Proceedings of the International Conference on Intelligent Agents, Web Technologies and Internet Commerce (IAWTIC'2001), Las Vegas, USA, 9-11 July, pp. 268-279.

Ofluoglu, S. 2003. Interactive building product information in the context of e-commerce world, Proceedings of EIA9, Istanbul, Turkey, 8-10 October, pp. 141-150.

Pocsai, Z., Debras, P., Gui, J.-K., Hagemann, D., Seifert, L. and Stumpf, D. 1998. The Common Semantic Model in GENIAL: A Critical Gateway towards an Intelligent Access to Component Libraries and Engineering Knowledge, Proceedings of Product Data Technology Days, 24-26 March, London, UK.

Ray-Jones, A. and Clegg, D. 1991. CI/SfB Construction Indexing Manual, RIBA Publications.

RINET 2000. RINET – Building product database, http://cic.vtt.fi/projects/rinet/rinet.html

Shocklee, M., Burkett, B. and Yang, Y., 1999. Product Data Markup Language (PDML) Specification, http://www.pdit.com/pdml/

Uniclass 1997. Unified Classification for the Construction Industry, RIBA Publications.

UNSPSC 2003. United Nations Standard Products and Services Code, http://www.unspsc.org/

van Leeuwen, J.P. and Fridqvist, S. 2002. On the Management of Sharing Design Knowledge, Proceedings of the CIB W78 conference on Distributing Knowledge in Building, Vol 1, Aarhus, Denmark, 12-14 June, pp. 235-242.

Woestenenk, K. 2002. The LexiCon: structuring semantics, Proceedings of CIB W78 conference on Distributing Knowledge in Building, Vol 2, Aarhus, Denmark, 12-14 June, pp. 241-247.