The Construction Information Gateway

Stephen R Lockley, Construction Informatics, Newcastle University
Robert Amor, Building Research Establishment

Affiliations: Keith Montague, Construction Industry Research and Information Association

Marshall Crawford, RIBA Companies Ltd.

Graham James, Thomas Telford Ltd.

A recent U.K. study set out the objectives for an Industry Knowledge Base, which would create and exploit information and knowledge to improve the efficiency and competitiveness of the whole U.K. construction industry by:

- improving the quality and efficiency of buildings and building projects by sharing standards and best practice
- improving the efficiency of the construction market by improving communications
- reducing the cost and quality of building design by sharing design knowledge
- providing relevant, value added network services to support communications on construction projects
- providing news and information that will enable the construction industry to compete more effectively

A Construction Information Gateway (CIG) is a way of providing on-line access, using internet and World Wide Web technology, to the industry's knowledge base. The development of an industry specific gateway would give users easy access to a wide range of information, reduce the costs of information delivery mechanisms for both suppliers and users, and would enable standards of both content and context to be more easily unified.

The range of information would include product information; cost information; project information; site information; industry 'knowledge' such as codes of practice, standards and research; software for analysis, calculation and expert systems; and construction news. Users of the gateway would have a single point of access to information appropriate to their needs.

This paper describes the first phase towards the realisation of the CIG being carried out jointly by Construction Industry Research and Information Association, Building Research Establishment, Royal Institute of British Architects Companies Ltd, Construction Informatics at Newcastle University and Thomas Telford Ltd. This phase aims to demonstrate the viability of a CIG – in technical, financial, commercial and organisational terms.

The paper outlines the information flow process model that is proposed for the CIG together with a methodology for managing the exchange of ISO 10303 STEP file based product models.

This research builds upon and provides a review of previous work in the area of construction information process and product modelling. Based on this past work a more detailed investigation of construction industry user types and needs has been carried out together with a study of the requirements of those who supply information to this industry.

1. BACKGROUND TO THE CIG DEVELOPMENT

Construct IT---Online Knowledge base---CIG

The development of a CIG has been of interest to the UK Government and the UK industry since at least 1995. All concerned concede that it would be of great benefit to the industry, though how it would actually work in reality (both commercially and technically) have never been clear. Initial work in this area was funded by the UK Government through a Partners in Technology project[1] which developed the first working demonstrator of a CIG. However, this system was very sparsely populated, and could only show the potential benefits without making them a reality.

At the same time as the demonstrator the UK Government was launching their Construct IT initiative, which looked at the way forward for the construction industry and the impact of IT. One of the major areas considered was the idea of a Construction Industry On-line Knowledge Base (CIOKB) and a scoping study[2] was published. This knowledge base is envisioned as a single point of entry to all information required by the construction industry, from news, journals, standards, codes of practice, practitioners and through to actual product information.

In the same year another Partners in Technology project[3] was let which examined major contractors views of a CIG and the benefits it would provide. The main objectives of this project were to investigate the information already available in electronic form to see how far this met the construction industry's requirements and to specify any changes or improvements the project team felt were necessary. Their main conclusions were that major contractors were committed to the

concept of a CIOKB, but there were concerns over how the information would be structured, presented, and delivered. They also delivered a strong message to agree standards with existing information suppliers.

This paper reports on the next phase of work on the CIG idea in the UK.

Trends in information usage

As part of the CIG definition phase a review of user requirements was carried out. The Mackinder & Marvin studies in the early 1980's were used as the base reference. These studies monitored the use of paper based information in a number of architectural practices over a period of several months. The CIG study examined how current Information Technology and Communications (ITC) had effected the way users perceived and accessed information in the construction industry and if these usage patterns had changed since the studies in the eighties.

The CIG study also built upon previous research findings from the CIOKB which set out to develop an information road map based on the usage patterns of ten large UK contractors. The basic methodology of this study was followed but it was focussed on small (1-4 man) architectural practices and larger consulting organisations. The objective was to determine whether the CIOKB findings were applicable generally to the industry.

In addition "The Barbour Report 1997" presented a current and comprehensive study of the Information Technology trends in the construction industry allowing the CIG study to focus on Information use.

A questionnaire was sent to 100 randomly selected small architectural practices and 20 consulting organisations. The survey revealed that

over 60% of practitioners preferred to use the PC for locating information, the most important information sources to them were Standards, Health and Safety and Specifications. These were consulted on a weekly basis, whereas manufacturer's literature and professional journals were consulted on a daily basis. The most common reason for consulting an information source was to solve a project related problem and there was a clear desire to store these search results in project specific archives. The most common reason for not storing information retrieved was it's perceived shelf life and, if in paper, the lack of physical space.

Practitioners were willing to pay for information on best practice, standards and specifications but not at all for manufactured product data.

When compared with the previous studies on information use the following conclusions could be drawn

- Consultants information usage patterns have change since MacKinder and Marvin Studies
- Dependence on their own and colleagues experience has declined
- Practitioners have now developed more rigorous and advanced searching techniques
- The large contractors and smaller consultants have very similar information usage profiles
- Internet use barriers are coming down, the advantages are seen by most parties
- Electronic media is changing working practice
- Storing information on a project specific basis is essential
- The opportunity to access all information but only pay for that which is required is desirable

2. SCOPE AND OBJECTIVES OF THE CIG

The Latham Report set a challenging target to reduce the real cost of construction by 30% by the year 2000. Information Technology was recognised as a key contributor to this target and the Construct IT vision for the industry has three elements:
- Encouraging improved sharing of information through the use of integrated project databases
- Developing an industry-wide knowledge base to facilitate improved access and sharing of information between industry and project teams
- Using IT to improve specific elements of the construction process

The Construction Information Gateway addresses the second of these elements, the delivery of an Industry Knowledge Base, which would create and exploit information and knowledge to improve the efficiency and competitiveness of the whole construction industry by:
- improving the quality of buildings and efficiency of the building process by sharing standards and best practice
- improving the efficiency of the construction market by improving communications
- providing relevant, value added network services to support communications on construction projects
- providing news and information that will enable the construction industry to compete more effectively

Information services in the industry are fragmented, incomplete and often poorly resourced. Specifiers often use products they know rather than the most appropriate. Product suppliers need better and more cost effective

ways of reaching their markets. Information resellers face problems of accuracy and updating as well as needing to tailor their offerings to meet the needs of particular groups of information users and providers.

The objective of the Construction Information Gateway (CIG) is to provide on-line access, using Internet and World Wide Web technology, to the industry's knowledge base. The development of an industry specific gateway will:

- give users easy access to a wide range of information
- reduce the costs of information delivery mechanisms for both suppliers and users
- enable standards of both content and context to be more easily unified.

The range of information would include:

- *product information*
- *cost information*
- *project information*
- *site information*
- *codes of practice*
- *standards*
- *research information*
- *software for analysis, calculation and expert systems*
- *Construction news.*

The scope of information would range from the provision of standard documents to the downloading of object oriented models of construction components (with parametric performance information) directly into a designer's computer model or an integrated project database.

The CIG will contain a **core** of meta-data (information about information) which will lead a searcher directly to the information required on the web site of the provider. It will also house a dynamic **exchange** enabling users and providers to carry out transactions (buying and selling of information) through the gateway .

The approach being taken is to develop a network based protocol for construction information interchange. This will provide an open standard for meta-data structures for describing documents and building products and a protocol for accessing these structures. These are referred to as Meta-Data Application Protocols (MAPs)

The long-term objective is for Information providers and resellers to have a common standard that will give confidence in inter-organisational information exchange. They will format information to suit their user group constituencies and remain in control of commercial arrangements.

It is envisaged that the CIG will yield three major benefits

- improving the rate of learning of the industry
- improving the market efficiency of the industry
- providing the benefits of increased use of information by the industry

3. RELATIONSHIP TO OTHER PROJECTS/WORK

Over the last few years there have been many projects EC funded and national which have looked at aspects of the CIG. As part of the CIG project a range of CIG like sites were examined to determine both the features currently offered and those being developed in research environments. Some of the most relevant are:

GENIAL[4]

The Global Engineering Network Intelligent Access Libraries project is an EC funded initiative to establish a framework for information access in engineering domains, including construction. Its framework manages personal, project based, company based, national, and international information management and access. It covers product information, standards,

technical information, etc. through an Internet interface and based on the concept of a network of collaborating VASPs (Value Added Service Providers).

BISNET[5]

Building Information System Network (BISNET) is an EC funded project. The BISNET project aims to offer on-line technical information in the field of construction, maintenance, renovation and restoration of buildings.

Its main objective is to overcome all technical language barriers by means of classification plans and identification codes that thoroughly define standard characteristics and performance of materials, components and measured works throughout the E.U. In particular, BISNET is to be developed according to a logical classification system which, through the classification of products, components and works of the building and renovation industry, will provide access to a series of national databases (initially located in Italy and Germany) connected with one another to form a network. The network will be structured according to different groups of data typology as listed: prices; archives of commodities; technical standards; items of standard specifications.

WONDA [6]

WONDA is an EC project that will enable industry to exploit the vast information and software resources located both on WWW and enterprise databases throughout Europe. As such WONDA aims toward 'putting information in the hands of industry'.

- European industry can use the WWW for competitive advantage while maintaining the quality

assurance associated with internal database systems.
- Diversity of databases will no longer be a barrier European firms who wish to exploit the Internet for co-operative working
- Rapid expansion of Europe's information infrastructure through providing an easy migration path for firms to make their enterprise databases available through WWW

NEWCON[7]

This EC project is to provide a fully interactive electronic news service for the construction industry operating on the public service network to be accessible from any personal computer via the Alpha DIDO system, developed under CEC IMPACT 1, and the Internet.

The service will have modules dealing with technical, commercial, financial, legal and political matters, jobs vacant and required, and items for sale using multi-media where appropriate. The data will be taken from editorial agencies as a direct by-product of their normal publishing processes and it will be organised for comprehensive search over a period of six months after which it will be automatically archived. Facilities will be provided for users to give a profile of their interests, which can be revised or extended as required. This will be used to review the news as it is received to extract items of matching interest and notify users on entering the system. Items with critical dates, e.g. tender invitations, the introduction of regulations and meetings, will be given priority for the user's attention allowing for a warning interval.

Common Features Offered by the Web Sites

This section provides a breakdown of the services and facilities offered by

the 43 web sites examined for the CIG project.

The majority of the web sites included a search system, either of a database, the web site itself or other web sites. The table below shows the different types of searches available.

Keyword search	25
Search by Title/Name	5
Organisation	6
Web Address (URL)	1
Date	1
Location	4
Improve search facility	2

Some of the web sites classified the information contained on the site or in a database. Common classifications included product classifications and internal company specified classifications. The classifications were commonly used to supplement a keyword search.

Classified, with no search	2
Classified, with a search	11
Alphabetical Index of information	4

Many of the web sites included various lists of information, usually providing links to other web pages.

Links/Resources (other web sites)	27
Members/Clients/Partners	7
Jobs	3
Publications (on-line and paper)	12
Events	11

Help pages were available on some sites, giving information about the web site as well as how to use it. A site map laid out the structure of the web site in order to aid navigation around it, and a FAQ provides answer to Frequently Asked Questions.

Help page	8
Site map	8
Frequently Asked Questions (FAQ)	4

There were many other types of pages available, the most common are listed below. Comments/Feedback forms were a method of sending comments to the organisation, News/What's new was usually new pages or information on the web site, but sometimes included news about the organisation or industry as a whole.

Comments/Feedback forms	7
Discussion Forums	7
Software download	5
News/What's new	19
Guestbook	1
Update notification by Email	2
Press releases	2

4. ARCHITECTURE OF THE CIG

Existing frameworks, such as the CPI[8], for the modelling of information flows in the construction industry tended to orient around the project based types of information, drawings, specifications, bills and the information users, engineer, architect, contractor. From the perspective of a gateway, this was found to be inappropriate for the overall architecture; the information types mostly reflect printed output and the distinction between user roles are often blurred. A less "project" oriented and more "global" oriented model of information flow in the industry was required. Five major information repositories were identified

- The (existing) Industry Knowledge Base
- The CIG Core
- The Organisational Knowledge Cache
- The Project Databases
- The CIG Exchange

Figure 1 The CIG Architecture
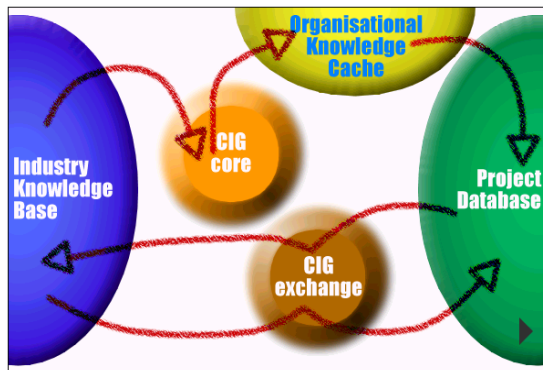
The industry knowledge base comprises everything from Multi-media on-line presentations through paper documents to the experience of an individual. It is unlikely that all of this information will become encapsulated in a single computer database, but essential that every piece can potentially be accessed through the CIG.

The CIG Core is the link to this existing information, it is a repository of meta-data, information about the information, in the Industry Knowledge Base. Each piece of meta-data is uniquely identifiable and provides the key for future access and retrieval. This important feature will be an enabler for future industry information integration. For example, meta-data may begin as an abstract to a British Standard, it is passed to the contractor where it is used to obtain automatically the electronic version of the standard and later any amendments that arise. This integration of different manifestations of the same information object takes place in the Organisational Knowledge Cache (OKC) and the Project Database. These are similar databases differing only in application. The OKC relates to information that is shared by all projects within an organisation rather than specific to one project. In the previous example of the British Standard an organisation would only wish to retrieve and pay for the electronic version once but share it with many projects.

In the CIG context, the Project Database may vary from a simple store of information collected for a given project through to an Integrated Project Database fully describing the design and construction of the building.

Finally, the CIG Exchange deals with the volatile information that arises during the lifetime of a building. This includes many forms of electronic commerce or EDI such as requests for pricing of materials, ordering and payment for documents.

More detailed information about the scope and architecture of the CIG can be found on the CIG Website[9].

5. PROBLEMS IDENTIFIED AND ADDRESSED

Legacy Data

Though only a small proportion of the information required by the industry is currently available on-line, the majority of it has been produced through electronic means. This usually implies electronic documents at the most basic level, but through to existing internal databases at the more advanced organisations. However, the structuring of information in these electronic media is unlikely to exactly match that agreed by the industry for a CIG. Therefore there will need to be methods provided to map existing data through to the form utilised in the CIG.

The CIG is currently adopting two approaches to this problem. The first is for information providers who have a very small amount of information to be made available through the CIG. For these providers there will be a controlled forms interface which allows information to be entered according to the internal CIG structure. This interface will provide checking of the inputs to validate information that

is entered (e.g., against a master list of publishers).

The second interface will be for information providers with a large source of existing data (e.g., major manufacturers). They will be provided with a toolkit to ease the process of mapping their internal database structures into the form required by the CIG system. This toolkit will make public the required structures in the CIG and provide a basic mapping functionality to enable information to converted in bulk and then uploaded directly into the CIG.

Distributed data

There is reluctance for the CIG system to become a publisher of information. It will not have the resources to manage information publication as well as to provide a gateway to all information. To this extent the CIG system will simply index the publicised meta-data provided by all information sources and point through to the existing source of the raw information for an individual to get the full document, etc. This will greatly reduce the information-processing load imposed upon the CIG as well as removing potential conflicts of interest if it were to be a publisher as well. It is expected that electronic publishing services would start-up alongside the CIG when it is operational. The meta-data for all information, however, will be held centrally, rather than gathered from individual sites whenever the CIG is used. This allows the CIG to guarantee to identify all information it knows about for every query, rather than relying upon only those servers which respond quickly enough to each individual query.

Evolution rather than revolution

It is clear that many in the industry will be reluctant to see radical changes to their existing processes. The CIG has the potential to radically change the way in which the construction industry works. There will be a fine line between engendering the changes required in this industry to make the CIG as useful as possible, without changing the industry so quickly that many are alienated in the process. The CIG proposes an evolutionary approach where existing processes are supported alongside newer processes. However, the CIG will always prefer to support beneficial process change rather than supporting current processes that are shown to be hindering the industry.

Product data structures

Defining standards for the product data required by the industry is a major task for the construction industry. Currently the IAI and ISO are working on an international solution, but it will be many years (if ever) before they have standards for all products required by the industry. However, other initiatives, such as CIG, can not wait until these standards are fully developed before launching themselves. The CIG is taking a pragmatic approach to product data standards. In the first instance it will base all product data structures on existing standards, to this extent the 'Building Core Model', which is the version 1.5 release from the IAI, will be the core product data structure in the CIG system. This will be extended to encompass the most commonly references attributes for products as currently seen in product catalogues, and this will be promoted as the initial description for any one product category. However, in recognition of the need for manufacturer's to

differentiate their products, the CIG will allow arbitrary attributes to be defined by any manufacturer along with the core data for that product (technically feasible with the property sets currently supported by the IFC version 1.5).

6. INFORMATION MODELLING OR BUILDING MODELLING

Examination of the current Building models currently under development (IAI and BCCM) shows a clear focus on the final building object, the information that describes the building and the process of instantiating the building model. In the CIG consideration is given to the information that exists independently of any specific building but from which a subset will eventually be incorporated in many building models. Rather we model the information **about** buildings in general some of which will need to be related to specific buildings, as models are instantiated.

A key issue for the CIG is how the models of generic construction information relate to, or are integrated with, specific building instances and what is the information flow process between these two different types of models.

7. OBJECT MODELLING WITH META-DATA

Object Modelling and Meta Data are terms that arise from two different disciplines, however from the CIG perspective they are similar. Object modelling comes from the software engineering and product modelling community and is the process of identifying the essential characteristics of an entity and producing a definition in a formal language for example EXPRESS or C++. Meta Data is a term associated with the Information Science community and is the process of identifying the essential characteristics of a piece of information. It is in essence "information about information" hence the use of the term meta. In the CIG our focus is primarily to model **information** relevant to building, not the **content** of the information. The objects we model fall into two broad categories,

- Information about Documentary objects used in the construction
- Information about building products

It has been found necessary to model these differently.

Documentary Objects

These include traditional publications e.g. books, standards, technical literature, as well as less tangible objects such as conferences, software, news. The STEP generic resources provide some support for this type of entity but it is quite shallow and is oriented towards integration within a product model rather than the CIG type of application. A review of other Internet gateway developments revealed efforts to address this problem and an emergent standard for documentary meta-data; the current standard is the "Warwick framework" which is an extension of the prior "Dublin Core". These standards define the key attributes for meta-data that is to be exchanged between gateways and form the basis of a simple entity definition for documentary objects.

Building Products/Components

There is a wide range of building product types and from a meta-data perspective these various types have different information structures. The Warwick framework has not been found to provide a good basis for these

objects and the CIG has accepted the need to develop new meta-data structures for the construction industry.

Currently there are many methods being promoted to allow the structuring and exchange of product information between systems. The STEP Part 21 and Part 22 formats provide for the exchange of data files which contain either the data for a model, or queries to a data source which has the capabilities of a relational database. The CORBA standard allows for a greater distribution of the data and supports full object based representations. However, these are formats for the exchange of information between tools, rather than the representation of published information. Information providers need a method to describe their products in such a way that the CIG can index the main properties of the products and point an interested query to the source for full information. There are many ways that this can be supported, for example through the ability to embed product information inside a document, or through a reference to a file, or embedded database query, which holds to product information. Current work on XML-STEP links would seem to provide the methodology that will allow embedded product information to be captured into web-based documents. However, whether they will provide enough flexibility to support coded references to files or database queries is not clear.

8. CONCLUSIONS

The current industry wide product and process modelling initiatives will begin to make an impact in the next two years. The models produced in the industry will need to make reference to the existing industry knowledge base. The CIG proposal to create a framework for modelling the meta-data of the industry and supporting the necessary information flow processes is timely and with phased delivery practical.

The fundamental difference between the CIG and other gateway initiatives is that it aims to enable the information copyright holders to set-up and maintain their own meta-data. It is not a proposal for a single search engine rather a methodology to be employed within the UK industry that has compatibility with other world-wide initiatives.

By enabling copyright holders to generate and maintain their own information to an industry accepted format we ensure the long term viability of these information flows. Also since the effort of generating the initial meta-data is distributed throughout the industry the practical problems associated with cost, time and effort, of creating the huge repository required are overcome.

[1] Parand, F. (1996) The Construction Information Gateway Demonstrator, CIBSE Journal (see also http://cig.bre.co.uk/cig/)

[2] CIOKB (1997) Construction Industry On-Line Knowledge Base, A study to establish the on-line information needs of the major UK contractors, draft final report, Ref. CI 38/19/87 (cc 1056), May

[3] DETR (1996) Construct-IT: bridging the Gap, Scoping study for the Construction Industry Knowledge Base, Construction Sponsorship Directorate, Department of the Environment, HMSO, UK

[4] http://www.uni-paderborn.de/GENIAL/)

[5] http://www.bisnet.org/

[6] http://www.bild.ie/wonda/

[7] http://alphadido.com/

[8] Co-ordinated Project Information Committee

[9] http://archhive.ncl.ac.uk/cig