MANAGEMENT OF INFORMATION FLOWS DURING CONSTRUCTION PROJECTS

Benoît Otjacques, Pierre Post and Fernand Feltz
Cellule de Recherche, d'Etude et de Développement en Informatique
Centre de Recherche Public –Gabriel Lippmann, Luxembourg
otjacque@crpgl.lu

SUMMARY

The right management of information is crucial in construction projects and has been studied by many researchers. This paper begins by defining and discussing some information management concepts. In particular, it proposes a reflection about the pull-push model for transferring information and links the resulting ideas to some generic issues of information management, namely information overload, information retrieval and information asymmetry. On the basis of these theoretical elaborations, it presents a prototype tool for construction project management that is based on original technologies. Instead of setting up a project specific web site, the approach consists of mixing pull and push transfer modes on the basis of an enhanced messaging system. This prototype may thus be considered as a proof-of-concept approach that new technological solutions may be explored if the conceptual analysis of generic problems takes precedence over default choices in terms of technology.

INTRODUCTION

Nowadays, information management is known to be one of the most important competitive advantages of the enterprises. In the construction industry, the design and computing tasks have been the first ones to intensively use the new information technologies (IT). Using IT to support the project management and the company organization has emerged as new trend a few years ago. It has been widely promoted since the Internet widespread all over the world. Therefore, the time has come to identify the new challenges raised by this paradigm shift.

This paper is organized as follows. First, the pull and push modes of information dissemination are defined and discussed. Second, some generic information management issues are identified and put in relation with the dissemination modes. Then, the BBeLink2 prototype is described and confronted with the previously mentioned issues. The last part proposes some brief conclusions.

MODES OF INFORMATION DISSEMINATION

Most of the time, the discussion about the solutions for exchanging information focuses on technology. But, in fact, technology isn't the main point. What really imports is functionality from the user perspective. Therefore, the basic discussion concerns providing the right information to the right person at the right moment rather than opposing web site and e-mail systems.

In this paper, the issue of information dissemination is defined as the set of all elements related to providing a given person with some pieces of information. According to this definition, the information consumer must be considered as the central element of the reflection and it appears essential to study the instantiation of the information flow. This naturally brings us to discuss the well-known 'push-pull' model of information transfer.

First of all, it is worth reminding some definitions proposed in the literature. The push mode is defined as follows:

- '*Data items are sent from the server to the clients without requiring a specific request from the clients.*' (Acharya et al., 1997)
- '*The technical definition of push is any automatic mechanism for getting information off the web from the users perspective.*' (Cerami, 1998)

- '*Push simply means that new information is delivered or retrieved automatically from a remote computer to the PC. Information does not need to be updated manually on a regular basis.*' (Basiel, 1999)
- '*Push (or "server-push") is the delivery of information on the Web that is initiated by the information server rather than by the information user or client, as it usually is.*' (Whatis.com)
- '*Push technology (Webcasting) is the prearranged updating of news, weather, or other selected information on a computer user's desktop interface through periodic and generally unobtrusive transmission over the World Wide Web*' (SearchWebServices.com)

Some of these definitions lack of precision concerning at least one fundamental element: the layer to which they apply. Some relate to the user and others focus on the computer. However, the need for accurate definitions has been identified for many years. The 7-layers OSI model, underlying almost all technologies concerned by information transfer, is probably the most important result in this domain. It permits to clearly identify which nodes are exchanging information and which services are linked to the transfer.

Similar definitions have been proposed for the pull mode. Nevertheless, for the purpose of this paper, it is sufficient to mention one of them. '*Using request-response, clients explicitly request data items by sending messages to a server. When a data request is received at a server, the server locates the information of interest and returns it to the client. Pull-based access has the advantage of allowing clients to play a more active role in obtaining the data they need, rather than relying solely on the schedule of a push based server.*' (Acharya et al., 1997)

This paper discusses the information management in construction projects. This isn't a telecommunication oriented paper. Therefore, the user level is considered as the reference point for the proposed definitions.

- The push mode concerns any information flow that occurs without explicit request of the person that receives it. In this case, the receiver plays a passive role.
- The pull mode concerns any information flow that occurs as a consequence of an explicit request of a person. In this case, the receiver plays an active role.

This can be illustrated by some examples. A phone call is obviously an implementation of the push mode, as well as a warning window on the computer screen. Similarly, surfing on a web site and searching a phone book are considered as pull mode.
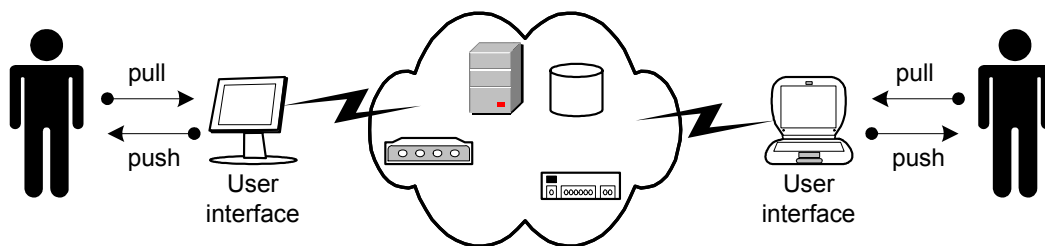
Figure 1 Push-pull modes from the user perspective

In the domain of information technologies, it must be noticed that a push mode at layer (n) may be implemented on a pull mode at the layer (n-1) and vice versa. For instance, while the e-mail technology is often considered as a push technology, it lays on a connection to a mail server instantiated by the client application at a lower level (pull mode). It is thus important to specify which level is concerned. According to the definitions adopted in this paper, the human-computer interface must be considered to evaluate whether the functionality implements the pull or the push mode, independently of the underlying technology (cf. Figure 1). It may be useful to illustrate this discussion with an example. The problems caused by the size of the data to be exchanged are often presented as intrinsic features of the push or pull mode but, in fact, they rely on the technology used to implement the chosen mode instead of the mode itself. One shouldn't confuse the assessment of some communications technologies such as SMTP, HTTP or ISDN with some conceptual reflection about the information dissemination.

This distinction between the pull and push modes is not only a theoretical debate but underpins some critical issues from a management viewpoint that are discussed in the next section.

THE INFORMATION MANAGEMENT ISSUES

The effective management of information is known to have a critical influence on the performance of a company. Previous research focusing on the construction industry have shown that this assertion is also valid in this domain. During a construction project, hundreds of information flows occur. The high number of involved companies makes the effective management of these exchanges crucial. Indeed, the steps of the production process are highly interrelated (Koskela and Vrijhoef, 2000), which leads to a strong dependency among the companies engaged in the project. Suppliers commonly contribute 75-80% of the value of a construction contract, making their effective management and coordination essential to cost, quality and time objectives (Clark et al., 1999). The problems related to the transmission of information have been identified as very important factors of inefficiency in a construction project. For instance, according to O'Connor and Tucker (1999), 22,1% of the problems occurring on a construction site relate to the communication of design information. It has also been estimated that '*up to 30% of the cost of a building project is due to the fractured processes and communication of the AEC/FM industry*' (I.A.I. - International Alliance for Interoperability, 2002). To sum up, a lack of information flow management often leads to delays, additional costs and consequently bad relationships with the customers. It is thus founded to discuss in more detail some generic issues that face information management in the construction sector.

Information overload

The first problem to discuss is the information overload. It is often cited as a major problem for the managers in general (Farhoomand and Drury, 2002) and the construction professionals in particular (Thorpe and Mead, 2001). This issue arises in two main contexts: '*when the user are given more information than they can absorb*' and '*when the information processing demand on an individual's time for performing interactions and internal calculations exceeds the supply or capacity of time available for such processing*' (Farhoomand and Drury, 2002). The former case relates to the amount of information to be processed and the latter to the time available to information processing.

Information retrieval

A second point worth being studied is the issue of information retrieval. This is linked to the information structuring and ease of access. When some information pieces are stored in a repository, the extent to which they are structured and indexed gains a fundamental importance as it underpins the performance of the search process. Moreover, the structure widely influences the degree to which some information processing tasks may be automated. Finally, ergonomics also plays a major role in this domain by influencing the cognitive effort required by the user to get access to the information.

Information asymmetry

Information asymmetry arises when all the persons dealing with a given problem have different levels of information about the same object. This difference may take several forms:
- distortion, due to a loss of information integrity during the communication,
- incompleteness, due to a loss of information pieces during the communication.

In most of the cases, the asymmetry results from the way that information is managed within the group of involved persons. This paper doesn't consider the asymmetry as the consequence of a deliberate behaviour of some members. The asymmetry may thus be caused by a deficient synchronization of information within the group. In this case, the members have a version of the information that corresponds to different moments of time. The most usual solution to tackle this problem is to set up a reference information repository that is hypothesized to contain the latest version of the information pieces. Consequently, in this context, the refresh rate of the data taken into account by the user or, in others terms, the frequency of the connection to the reference repository, becomes the central element. The asymmetry may also originate from an imperfect transmission path that modifies the integrity and/or the completeness of the content. In this case, the asymmetry is mainly influenced by the choice of the information transfer mode. Indeed, it is well known that the

higher the number of intermediaries in the transmission path the higher the probability of information corruption. In such a situation, the solution consists of providing a mean to certify that the information hasn't been corrupted during the transfer.

RELATION BETWEEN ISSUES AND DISSEMINATION MODES

On one hand, it has previously been shown that two basic modes of information dissemination coexist. On the other hand, three fundamental issues have been identified in the domain of information management. This section explains the relations that link these observations together.
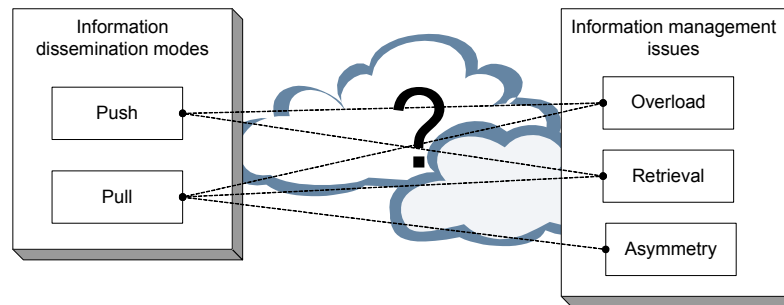
Figure 2 Relation between information management and dissemination modes

The push mode is very useful when the sender demands the receiver to react to the incoming information. The push mode means that some stimuli are sent to the receiver and force him to adopt a given behaviour (read – ignore – answer). This mode is therefore the most adapted to handle information flows that require a rapid answer (e.g. phone call to warn the company that an accident occurred on the construction site). It is also useful to deal with events having some unexpected components (e.g. invitation for a meeting). The push mode also suits the implementation of reminding messages, such as those used in most of the calendars. Finally, it may also be used to tell the user that some elements of the context have changed (e.g. service to alert the users when data has been added or modified in a web site).

The major drawback of the push mode concerns the information overload issue. If a large amount of information is sent to the receiver, it may exceed his cognitive capacity (e.g. spam using e-mail). Therefore, an efficient push implementation must filter the communications in order to keep the number of information flows at an acceptable level. An important remark must be made in this context. In many circumstances, some incoming messages pushed to the user are archived for later use. In this case, according to the definition adopted in this paper, the retrieval of these information pieces must be considered as pull mode (causing the usual problems associated to this configuration, see hereafter). This observation illustrates the fundamental importance to clearly identify the layer considered in the discussion to avoid confusion between technology and functionality. Another drawback of the push mode relates to asymmetry. It is important to notice however that the potential importance of the asymmetry issue depends on the way how the push mode is used in practice. If the information is pushed from point to point within the group, some kinds of asymmetry appear quite irremediably either due to the transmission delay among the nodes (cf. deficient synchronization) or due to content modification by one of them (cf. imperfect transmission). If one trustworthy information provider sends information to several persons via the push mode (i.e. multicasting), the imperfect transmission problem may be avoided and it only remains the synchronization issue.
Finally, if a viewpoint that is consistent to the proposed definitions is adopted, the push mode isn't involved by the retrieval issue because the communication is instantiated by the computer.

The pull mode permits to limit the information overload because the user doesn't have to handle any disturbing incoming information flow. The information is picked up from a repository when needed. The pull mode suppresses the well-known issue of 'just-in-case-they-need-it' push publishing (Thorpe and Mead, 2001) (which must be considered, in fact, as a misuse of the push mode). The greatest advantage of the pull mode is that it doesn't offer the user an easy way to abusively trigger hundreds of information flows. There is however a counter part of this situation: the dependence on the initiative

taken by the user. The pull mode isn't able to resolve a situation where a project participant doesn't instantiate any connection to the repository.

The pull mode presents also an advantage in terms of information symmetry. Because one information source is hypothesized to be up-to-date, it is easy to know where to collect the latest version of any information piece. As already mentioned, the efficiency of the pull mode in this domain primarily depends on the refresh rate of the data. More precisely, in most of the pull mode implementations, a temporary buffer storing previously retrieved information is used to optimize the performance. Due to this feature, the data may be accessed while the synchronization mechanism is deactivated and the risk of consulting old data reaches the same level as in the multicasting solution.

The information retrieval issue must probably be considered as the major drawback of the pull mode. As the amount of information to be searched increases, the difficulty to rapidly find the required data becomes more important. In the worst case, the required effort may discourage the person to search the repository and consequently to store anything into. Therefore, an efficient pull mode implementation must provide structured information combined with user-friendly retrieval tools.

THE SITUATION IN THE CONSTRUCTION INDUSTRY

In practice, most of the solutions proposed to manage construction projects combine some push and pull functionalities. This can be illustrated with the example of a project specific web site (PSWS), which is nowadays probably the most promoted approach. While this solution is often presented as the perfect example of pull mode, it includes, most of the time, some push-based functionalities (e.g. e-mail technology used for data update notification or invitation to meetings). The push-oriented approaches also often include some pull functions. For instance, an electronic message sent to invite the receiver to visit a website may be qualified in this way. The effective management of a project requires thus an environment able to deal with push and pull communications.

Instead of improving pull-based project management environment with some push functionalities, we have chosen to adopt the symmetric way: designing an enhanced push-based tool – the BBeLink2 prototype – including some pull features. This approach is based on the observation that web sites present widely discussed advantages but also some drawbacks pointed out by quite few authors. Most of the web site systems face the same problems than those encountered for years by integrated project databases (IPDB). For instance, information retrieval becomes more difficult as the amount of data increases, data partitioning and access rights may be very difficult to manage (see Amor & Faraj, 2001, for a comprehensive discussion on misconceptions about IPDB). Moreover, they cause the emergence of new kinds of risks for the companies, especially for the smallest ones:
- dependence on the (potentially rival) company hosting the web site,
- dependence on the usage by all projects participants,
- confidentiality of the data,
- risk of failure of the web site operator.

Finally, the greatest challenge for the PSWS is probably to bring the construction professionals to replace some of the traditional point-to-point communications by the consultation, relying on personal initiative, of a central repository. The effective use of a PSWS requires the project participants to change their way of performing some of their tasks. Unfortunately, some recent empirical experiences have concluded that '*even six years is a short time to change human attitudes to the means of communication used, in spite of impressive new technologies being available.*' (Howard and Petersen, 2001) In conclusion, '*the organisational issues surrounding the use as well as the psychology involved in getting all participants in projects to accept using new technology are now in focus.*' (Björk, 2002)

From this perspective, the BBeLink2 approach is less ambitious. It is based on the strategy '*which does not present technology as a new mode of organisation but only as the way to introduce its necessity*' and this approach '*is not unfounded and has been successfully used within other industries [than the construction]*'. (Brousseau and Rallet, 1993). This approach also takes into account the importance of proposing technologies that may be rapidly accepted, which has been recognized as a fundamental element of acceptation for the small companies (Mundim and Bremer, 2000).

BBeLink2 has been designed to handle both point-to-point communications like classic e-mails and multicasting dissemination, without using a central web site. The multicasting diffusion of information may be instantiated only by the project creator. It is used to diffuse global project properties that present some interest for every participant (e.g. picture of the current state of the building site, project participant list, building description). This feature is intended to increase the confidence of every participant in the data validity. A technological description is provided in the next section.

BBELINK2 TOOL

Technology description

The BBeLink2 prototype is based on the creation of a community. Every member of the system uses a client application that handles the connections to a server. There are only two centralized modules shared by all the participants: a communication server used to transfer the messages and a central repository containing the identification of the people/companies allowed to use the system. There isn't any central repository for project data.

BBeLink2 includes two major types of functionalities: communication modules and storage modules.

The communication issues are handled by implementing two kinds of messages that may be transferred among the client applications: user messages and system messages. The former may be compared to classic e-mails that would be enhanced by some value-added features (e.g. reliable time synchronization for message dating, inclusion of meta data describing some message properties). They are dedicated to be composed and read by human beings. The user messages include some primary and secondary data. The primary data are those that should be read by the receiver in all cases (i.e. sender, sent date, subject, content text). The secondary data regroups those that may be read by the user if he wants to process the message in a comprehensive manner (e.g. for sent messages: reception status for every receiver; for received messages: attached files). The system messages are exchanged between computers for performing some low-level tasks in a transparent manner for the user (e.g. acknowledgement mechanism for every user message, synchronization of global project data via multicasting).

The storage modules are dedicated to store the information related to the projects. This data describes two kinds of objects: the communication instances and the project properties. The former concerns the description of all the information flows (IFs) occurring among the project stakeholders, independently of the communication media used. BBeLink2 includes a repository for storing all IFs. It distinguishes a structured and an unstructured part in an IF description. The structured part is implemented as a set of standardized meta-data to assign to any IF. Some are very usual, such as the identification of the sender, and others are more original, such as the content type (technological, administrative, commercial…) or the communication media used (phone, fax, meeting, e-mail…). The meta-data set has been instantiated from a three-step process. First, a literature review has provided a list of potential values (e.g. list of project stages). Second, a limited list of values has been built, which appeared as an acceptable compromise between usability and utility. Third, the resulting list was revised by a partner enterprise active for several years in the AEC industry. The unstructured part allows to describe the IF with free text and to attach any kind of file (picture, drawing, contract…). The IFs linked to the BBeLink2 user messages form a specific subset. Because the user messages are instantiated inside the BBeLink2 environment, the corresponding IFs are automatically created in the IFs repository. The latter concerns the project itself and is implemented by project properties (e.g. building description, construction site picture, participants list). BBeLink2 distinguishes global and local project properties. The former are defined by the project creator. They are diffused to the project participants and synchronized by multicasting. The latter are also synchronized but they may be modified by the project participants.

BBeLink2 and the information management issues

The BBeLink2 concept must then be confronted with the three basic issues related to information management that have been previously mentioned. In order to facilitate the understanding of the different discussed elements, the Figure 3 provides a global view of the BBeLink2 features, in terms of information dissemination modes.

The first issue to analyze is the information overload. BBeLink2 includes several features intended to handle this issue.

- The server handles only messages of the members of the community. Moreover, it may be configured to block the messages of any member showing an inappropriate behaviour. This mechanism may thus limit the risk of information overload due to spam.
- The BBeLink2 clients integrate a filtering mechanism. A filter is a logic expression defined using primary (preferably) or secondary data of the user messages. A filter may also be a combination of existing filters. The filters are evaluated for each incoming message. If the expression is true, a specified action (e.g. forward, save in specified directory…) is automatically triggered. This mechanism may limit the amount of incoming flows to handle.
- The distinction between primary and secondary data in the user messages also limits the information to handle during the first step of the cognitive processing. Moreover, the user may choose which (primary) meta data of the messages should be displayed in the graphical interface.
- It is possible to generate summary reports about the information flows that have occurred during the project. This may help the professionals to rapidly get an overview of the current state of a project.
- The user may deactivate the push-based warning messages sent when some global project properties are updated.
- The user interfaces used to handle user messages and information flows are very similar, in order to reduce the learning effort for the user.
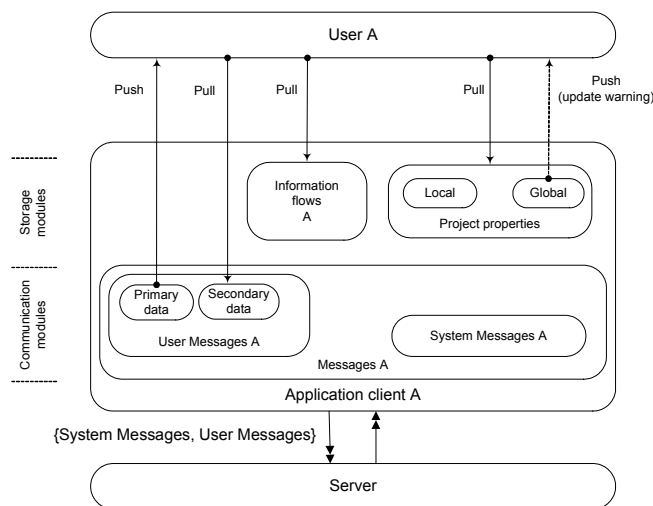
Figure 3 BBeLink2 conceptual model (push-pull viewpoint)

The second issue concerns the retrieval of information. The design of BBeLink2 has tried to structure the data as far as possible, in order to ease the information access.

- The whole BBeLink2 environment is organized by project. Every message and every information flow is linked to a specified project.
- Every user message and information flow is characterized by a (similar) set of meta data. This allows an efficient sorting and retrieval of any instance of communication.
- The filtering mechanism may be used to automatically organize the incoming user messages in a given way, specific to the company.

The information asymmetry is the last issue to be discussed.

- The multicasting system guarantees that the global project properties are identical for every project participant. Indeed, only the project creator is granted to set and modify them. The local replication that is stored on the computer of each participant is encrypted and may not be accessed in write mode.
- The information synchronization depends on the connection frequency to the Internet. As soon as a new connection is instantiated, every new modification of the global project data is automatically updated on the local instance. This update is transparent to the user and can't

be deactivated.
- BBeLink2 also provides a reliable dating system. Each time value set in relationship to a message is collected from an external time server. In conjunction with the low level, permanently activated acknowledgment mechanism, this permits to be really informed on the reception status of any message, for any receiver. The sender may thus also be aware of a lack of synchronization.
- By integrating a digital signature implementation, BBeLink2 prevents the corruption of the transferred data.

CONCLUSIONS

In many discussions, the pull and push mode are presented as rival modes of information dissemination while they are in fact complementary. Moreover, the technological viewpoint often dominates the functional perspective. This paper describes a proof-of-concept tool, called BBeLink2, which mixes push and pull functionalities. It is original in many ways but the absence of a central web site is certainly one of the most worthy to be mentioned. The next step aims to get a sample of architects to evaluate the tool, especially in terms of utility, usability and intend of use.

REFERENCES

Acharya B., M. Franklin and S. Zdonik (1997), *Balancing Push and Pull for Data Broadcast* in Proceedings of ACM SIGMOD Conference, May 1997, Tuscon, Arizona.

Amor R. and Faraj I. (2001), *Misconceptions about integrated project databases*, in Electronic Journal of Information Technology in Construction, http://www.itcon.org/2001/5/, accessed 12 Dec. 2002.

Basiel A. (1999), *Push & Pull technologies using a WIDE approach – Web-based Instructional Designed Environment*, 3rd Human Centred Technologies Postgraduate Workshop 1999, Brighton.

Björk, B.-C. (2002), *The impact of Electronic Document Management on Construction Information Management*, in Proceedings of the CIB w78 Conference, Distributing knowledge in building, June 2002, Aarhus, Denmark.

Brousseau E. and A. Rallet (1993), *Développement des systèmes télématiques et évolution des relations interentreprises dans la construction*, Rapport d'une étude financée par le Plan Construction et le PIRTTEM, Université Paris-Sud, available on demand from http://www.jm.u-psud.fr

Cerami E. (1998), *Delivering push*, Mc Graw Hill, NY.

Clark A. M., B. L. Atkin, M. P. Betts and D. A. Smith (1999), *Benchmarking the use of IT to support supplier management in construction*, in Electronic Journal of Information Technology in Construction, http://www.itcon.org/1999/1/, accessed 12 Dec. 2002.

Farhoomand A. F. and D. H. Drury (2002), *Managerial information overload* in Communications of the ACM, October 2002, Vol. 45, No. 10, pp. 127-131.

Howard R. and E. Petersen (2001), *Monitoring communications in partnering projects* in Electronic Journal of Information Technology in Construction, http://itcon.org/2001/1, accessed 26 Nov. 2002.

I.A.I. (2002), http://cig.bre.co.uk/iai_uk/iai/page2.htm, accessed 11 Dec. 2002.

Koskela L. and R. Vrijhoef (2000), *The prevalent theory of construction is a hindrance for innovation*, in Proceedings of International Group for Lean Construction 8th Annual Conference, http://www.susx.ac.uk/spru/imichair/news/template.cfm?content=iglc-8.cfm, accessed 14 Dec. 2002.

Mundim A. and C. F. Bremer (2000), *Design of a computer-supported cooperative environment for small and medium enterprises*, 2nd IFIP Conference: Managing Business-to-Business Cooperation, Dec. 2000, Florianopolis, Brazil.

O'Conner J.T. and R. L. Tucker (1986), *Improving Industrial Project Constructability*, in Journal of Construction Engineering and Management, Vol. 112, No1, March, pp 69-82.

SearchWebServices.com (2002), http://searchwebservices.techtarget.com/, accessed 14 Dec. 2002.

Thorpe T. and S. Mead (2001), *Project-Specific Web Sites: Friend or Foe?* in Journal of Construction Engineering and Management, September/October 2001.

Whatis.com (2002), http://whatis.techtarget.com/, accessed 14 Dec. 2002.