A design proposal for an internet based Simulation Quality Control Tool

Michael Donn, Robert Amor and David Harrison
Centre for Building Performance Research
Victoria University of Wellington
PO Box 600
Wellington, New Zealand

ABSTRACT

The lack of widespread use of building environmental design decision support tools in architecture appears to be because the tools are often too simplistic. Designers find it difficult to trust the output of a tool that apparently trivialises the design issues. Even, regular users of simulation have a difficult time developing procedures to ensure that they can trust the output of the simulation. A web site has been developed which addresses this need directly. This paper describes the design principles on which the web site is based. Its aim is to develop a means by users can test the "reality" of the performance figures which simulation software predicts for the thermal, visual and acoustic environment of a building

INTRODUCTION

It is common for researchers to observe that designers won't use design decision support tools[1]. It is often assumed that this is because the tools answer the wrong questions. The goal of this research was to identify the types of questions that architect and client wish to have answered where design decision support tools[2] specifically for environmental analysis could be of assistance. In the process it examined the question of architects' interest in creating environments of thermal, visual and acoustic quality despite suggestions to the contrary: *It is a telling commentary on the current situation that architects must now be convinced that it is no mean achievement to design buildings that function well..*[3].

Sometimes, performance assessment tools are used to derive simplified guidelines as design decision support 'for architects'. For many years design tools for building professionals have been developed from the basic equations describing the building physics and psychophysics[4] into simplified charts, nomograms and simple calculator programs. This effort continues today. Developers of computer software for digital simulation of building performance are continually on the search for improved Graphic User Interfaces (GUI's) - the modern day equivalent of the nomogram.

THE RESEARCH 'PROBLEM'

The lack of widespread use of the current crop of environmental design tools in architecture appears to have several root causes:

i) the tools available are often too simplistic. Researchers simplify a rigorous performance prediction equation or set of equations to the point where they judge they will be acceptable to architects who do little in design to predict or systematically document building performance[5]. This need for simplification is often suggested not only by individual architects expressing a desire to be told a "rule of thumb" for a particular situation, but also by their professional associations[6]. Too often this simplification trivialises the issues and makes the performance model so remote from the complexity of reality that the designer sees the tool as irrelevant.

ii) even where a project is of sufficient size to have an expert design team, the environmental design experts in the design team find it difficult to relate environmental design issues to the interests and concerns of the architect. The causes of these problems are many. They include the oft-quoted lack of reading by architects of anything more complicated than manufacturers' brochures[7]; individual design analyst's inability to focus on the whole design rather than their one area of expertise; and the difficulty of establishing a good working relationship in a design team where the professional and financial rewards for team members may well conflict.

iii) experts agree[8,9] that the design decisions made very early in the conceptualising phase of a project determine how well it is going to perform. But, design decision support tools are most accurate when the design is complete and a detailed performance simulation can be undertaken. Unfortunately architects are trained[10] to look in this instance for the "rule of thumb" which directs them towards the successful solution without requiring a great deal of thought. Nils Antoni, architect and then head of the National Swedish Institute of Building Research writing in the CIB journal in 1986 summed up the problems with this approach succinctly in writing about *information which actually reaches the profession and is assimilated ... is highly selective and carefully pre-digested* [into rules of thumb and guide books etc..]: *I am suspicious of selected, processed information. It is a last resort... One never knows what criteria lie behind the choice made and how competent those doing the processing are..*

The research reported in this paper studied the practice of design tool use. It examined use of environmental design tools in real design situations in order to draw general conclusions about:

the types of questions users want environmental design decision support tools to answer;

the nature of the input and output to these tools that is acceptable; (drawing lines on graphs; entering numbers in spreadsheets; automatically transferring data from the CAD drawing to the environmental calculation program?...)

the types of quality control procedures adopted by the current small numbers of regular users of simulation based design decision support tools that provide some guarantee of the reliability of their analyses. (These procedures need to be codified and incorporated into the design tools themselves to ensure that the "black box" design tool yields information the designer can trust.)

The vast majority of designers interviewed on the topic of building performance simulation expressed a strong desire to be directly involved in the quantitative assessment of building performance. They don't want this assessment to take longer than five minutes because they want to know what to do in the design not to take a long time sorting out the calculations. They also don't wish to know what the difference is between a deciBel and an R value.

But, when asked about design options affecting the environmental performance of their buildings, they stated they want performance 'numbers' that help them rank one option against another. Numerical reports of the performance, like numerical petrol consumption data for a vehicle are needed to comprehend how big an environmental design issue really is.

It is becoming deceptively simple to produce output that represents reality. When the conjunction occurs between the desires of architects to be involved and the availability of simulation software, then as simulationists we have a potentially huge problem. This is most obvious in the field of lighting simulation because the "sexy" pictures produced by the lighting simulators are useful for so many more traditional presentation purposes that there is already huge demand for them to be available in simpler and simpler to use packages[11]..

This situation is well-known and documented in previous papers to IBPSA conferences. Senior designers in a firm using simulation software are aware that they have no way of checking that their simulations are "real". This goes some way to explaining why they might be less confident about using simulation in the office. It is more surprising to discover that simulationists rely on gut feeling rather than formal checklists or other similar procedures to determine that their simulation results do represent reality. Senior engineers in a simulation office have very few options to check the simulation work of a junior in the office.[12]

QUALITY ASSURANCE

Users of simulation software need to be able to judge the "truth" – the "reality" of their simulations The principal research problem having identified this need is to establish a system by which one might calibrate the output of a simulation program in such a way as to ensure that its predictions represent the reality the user understands. What is needed is a test for the output from a simulation program[13] like the *Turing Test*[14] for the 'existence' of computer-based (so-called artificial) intelligence.

The concept proposed is to build on the most common of ways that simulation professionals ensure reliable simulation models: re-use a previously reliable model. At its simplest, this type of re-use takes the simulation model, the 'virtual' building, adjusts the parameters describing it and trusts that it will work as well describing the new situation. At a more complex level, the simulation of a building that is to be refurbished is calibrated against the current performance of that real building. The proposed QA system is based at a minimum on permitting users of simulation software to find instances of building descriptions that have been created to match the building performance scenario they are investigating. At the high reliability end of the scale, it will provide real data against which the new simulation model can be calibrated.

Finally, the proposed QA system is to be a framework for QA processes in a simulation office, not a fully proscribed database and set of performance appraisal tools.

QUALITY ASSURANCE - REALITY TEST

The following statement is intended to function as the same type of truism in digital simulation that the Turing test is in artificial intelligence. Its careful application to digital simulation processes should generate Quality Control tests that convince the sceptics interviewed in the Case Studies in this thesis that any of their design decisions supported by these simulation processes are dependable.

> *Changes in the predictions of a simulation program with changes in building design should always be of the same scale and nature as those perturbations in performance observed in reality.*

The following section examines the role of a Quality test in passive solar house design.

In passive solar house design the number of alternative design tools that might be applied is large. Approaches vary from consulting a list of good ideas in case studies of existing solar houses to full digital simulation of house thermal performance. If one applies a comprehensive digital thermal simulation program to the design, multiple simulations are made of a systematically varied series of digital models. The results of these simulations are summarised in graphs, tables and simplified correlation formulae. If we apply the reality test to this we require:

- the producer of the solar house design tool must demonstrate that for their simulation: *changes in building design should always lead to the same scale and nature of perturbations in performance observed in reality.*
- the user of the solar house design tool must be convinced that their uses of the simulation tool are always *of the same scale and nature as those perturbations in performance observed in reality.*
- the client being advised by the user of the design tool must be able to rely on the fact that its predictions are always *of the same scale and nature as those perturbations in performance observed in reality.*

To achieve the goals highlighted by the reality test, a QA system built into a solar house digital simulation program must therefore provide the following:

- a means of confirming that the mathematical operation of the software installed in a new situation is still accurate - the role played by sample files now.
- a description of the sample e-building and its input file in simple construction terminology.
- a simple set of automated tests that demonstrate the performance response of that e-building to systematic changes in its design.

To provide each of these features, the design tool must contain an automated set of routines. These will apply a standard set of changes to the parameters describing the ‘sample’ buildings and compare these simulated buildings’ performance to a library of corresponding building performance responses. It is essential that this set of routines be automated so that the user is not required to invent test routines but rather is reassured by the output from the QA button or icon in the program. Using a button like this will teach them how to compare systematically the e-building description with its predicted performance using standards that the software independently verifies. Once this process is successfully implemented, it should influence consultancy use of the software so that before making recommendations based on its predictions users would ensure that their e-building ‘behaves’ in a standard manner given the pre-defined standard stimuli.

The keys to making this process work are: the automation of the process, establishing the reporting process in language that is understood by all users, and most crucially, determining an appropriate set of standard stimuli which reveal the reality of the e-building. As with the Turing test putting this proposed test into operation is the difficult part of the process. The third key requires the most work. There is no known internationally respected library of standard responses of buildings to standard stimuli (such as changes in design) that could be used to test the reality of the response of an e-building. The second key, description of the building in the language and terminology of the building site rather than the mathematics of the algorithms simulating their behaviour, is the subject of much of the interface design work being put in by software vendors internationally. The first key is largely unexplored by vendors and even by users and requires the other two to be complete before it can be attempted.

The closest that any research team has come to defining the standard stimuli required is in the BESTEST[15] system for design tool 'validation'. The research project examined software tools and their application and one of its products was a complex set of validated data based on measurements of real buildings against which the predictions of simulation programs can be compared. An illustration of the complexity of this type of reliability test can be found in the 'simplicity' of the test devised by the BESTEST team: the only measured data they could document well-enough for their purposes is from *test cells* - one room buildings which have been systematically instrumented. Work is progressing on expanding the database to include measurements from buildings with more than one heated interior zone.

A QA instrument produced to be incorporated into a simulation package for designers of solar houses, and of more general application in thermal simulation must contain the following automated package:

- **sample e-buildings that represent the full range of complexity and size of buildings that might be designed by the user of the package** - e.g. a three room dwelling; a five room dwelling with loft and basement; this same five room building with slab-on-ground heat loss; the same building with a sunspace; the same building with a Trombe wall; full disk copies of the output files for these buildings; an on-line tutorial guide instructing the user a) in how to write these input files; and b) in how to make standard changes to them; and finally, an on-line checker that automates the comparison of the output of the user's simulations of these buildings and of standardised changes in them with the expected values.
- **sample e-buildings which are one-room validation files** describing the real data developed for the BESTEST validation programme.
- **a 'validate' button which institutes a standard set of simulations** of the user's building under certain specified standard conditions and compares (graphically) the relative size of the changes in the output with the relative size of changes in the output of the sample buildings. The changes to be tested would be: doubling and halving of all glass areas; making the infiltration rate rise to 5 times and fall to half its established value ; doubling and halving the R-Value of every external surface element in the building; doubling and halving the heat capacity of the floor and wall elements of the building.
- **a standard set of output graphs which contain *base cases*[16]** that allow the output to be measured consistently against well-characterised buildings: these base cases would be described in detailed case notes and would represent relevant situations: they may even be generated by the software based on the user's choices when setting up the model of their building (e.g. it may be a standard building operated as the proposed building is modelled).
- **on-line test or evaluation aids** which graphically compare the fractional changes in the user's own e-building with the changes in the sample and base case e-buildings.

THE BASIC TEST:

What people need to know is whether or not their e-building is real. If it is real, it is expected that the e-building will 'behave' like a real building. The research issue is to develop a test of reality. Merely matching performance to measured data is insufficient. The test must identify 'behaviour'. It was hypothesised that: *if an e-building's changes in performance in response to standard design changes are the same as those noted in reality then the virtual model is 'real'.*

The major limitation of this approach is that real buildings do not have their performance measured, building design changes instituted and then more performance measurements. At best, the services might be changed and before and after performance data might have been collected. A 'real' model was first sought in order to compare to a virtual model. The IEA BESTEST test cell was selected as the 'real' model.

The goal is to follow the BESTEST approach: defining an acceptable range of output changes. Within this range, an e-building performance would be deemed to be 'real'. The absolute value of the output is largely irrelevant. It is the relative sizes of the changes in performance that are to be matched. The performance would be presented in normalised terms (energy use per person, per square meter, per cubic meter, per degree day, per operating hour etc) in order to facilitate the comparison.

In addition to the BESTEST model a real building was sought to test the e-buildng performance against a more complex 'real model'. The 'Moor house' by architect Roger Buck was selected. This house conforms to all standard solar design strategies and a very high proportion of its structure is thermal mass.

Once both spaces were modeled to the highest degree of accuracy, standard changes were performed to both models and the output was compared.

THE QA TEST IN PRACTICE

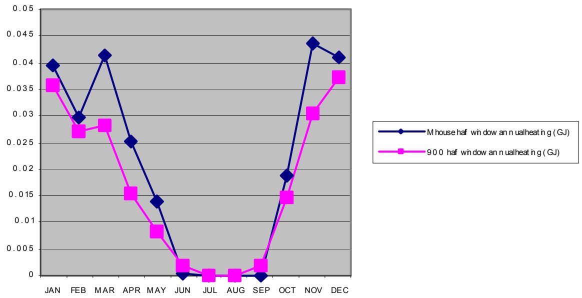
The above two graphs illustrate the types of tests being developed for the Quality Assurance evaluation. They show energy performance calculated by the Sunrel program[20]. The aim is to characterise behavior - changes in performance- rather than total values of energy use or temperature. Thus the values plotted in the charts are all presented as fractions of the January heating figure (January being the month with the highest heating load for the BESTEST weather data used in the calculations).

Further development of the concept will require normalising processes that enable the behavior of the internal temperatures to be compared as well.

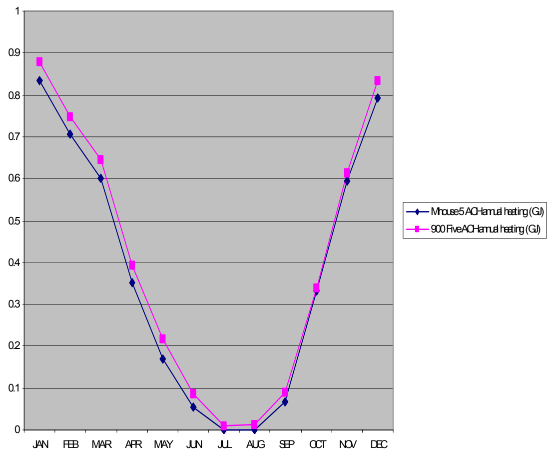
WEB-BASED QA DATA

This QA process is not possible without the establishment of an internet database of building performance benchmarks which could be automatically consulted by the simulation software. These benchmarks would comprise both measured and simulated data. They would be accompanied by the building description outlined above.

It is imperative that any new innovation of this type



being added to the internet works closely into the current and future design of the "world wide web". A single repository of performance data is not



consistent with normal web design principles. What is proposed is a navigation system and a set of standards which can be used to structure web pages describing building performance wherever they are on the web.

For example: at present when one searches the internet as a designer of a daylit school for sites matching keywords such as “school” AND “software” AND “light” one is left sorting out the relevant reports from the 2000+ school physics sites reporting simple software solutions to lab issues. The goal is to design a system where the computers doing the search and the computers storing the data have access to a shared definition of these keywords. They understand their meaning and can thus conduct a much more meaningful search.

EVALUATING THE TEST

The next few paragraphs draw from the ideas presented by Burners-Lee on the W3 consortium web site[17]. They address the issues raised by the idea of a database which is to be accessed more often by automated software than by software operated by people. This is in the spirit of *the semantic web*[18]*:*

The Web was designed as an information space, with the goal that it should be useful not only for human-human communication, but also that machines would be able to participate and help.”

EVOLVABILITY

“By "data" as opposed to "documents", I am talking about information on the Web in a form specifically to aid automated processing rather than human browsing. [19]

The central issue here is that the design of the Building Performance Database is such that it can **evolve**. Evolution means:

iv) that any new type of analytical document can be added: in the structure of the QA web site there is no restriction on a new simulation program’s data type being used as the format for a new file stored on the web

v). that the number of Levels in the classification system for finding data can be increased: while it would be difficult to maintain interoperability if some levels were to disappear altogether, adding Levels would not stop the search system from working.

vi) that the system should survive the birth and death of new internet technologies: The data files could be ASCII, binary, movie files, simulation program binary files - anything. All that the system guarantees is that they will be found.

vii) that the system should survive the birth and death of new computer software: Some data files which are input files for particular analysis programs will go out of date as the programs are updated or made obsolete by new developments. However, nothing in the definitions is dependent on a particular analysis program format.

METADATA

The key to maintaining reliable access to data is to develop an understandable, long term and unambiguous naming convention.

The key to all this is **Metadata**[20] - machine-readable data about data. The QA web site’s central repository of information about building performance data is metadata about buildings and their performance.

MODULAR DESIGN

The key to modular design is the simplicity of the interconnection of the modules. Unless each individual module can be understood and worked on independently there is no advantage to dividing a system into pieces. The advantage of the modularity is that each individual module can be upgraded without affecting the others, allowing for incremental improvement in design and for smaller design or maintenance teams. The modules of the Building Performance Database are: i) specification of the local data storage protocols for the recording of each building’s performance; ii) specification of each of the levels in the central repository classification of the building performance data; iii) the use of internet technologies for the communication between the database, each building’s data and the users of the information.

TOLERANCE

Be liberal in what you require but conservative in what you do[21]

As shown by the proliferation of non-standard HTTP which has lead to web pages that can be read by one web browser and not by another, this principle has an inherent weakness: it can encourage a too liberal attitude on the part of the creators of building performance datasets. It still provides an essential guideline. Unless the system is tolerant of various ways in which files can be stored and delivered it will not work. It must be possible, for example, to store data as ASCII or a Binary DOE2 file. Or, for a 3D CAD file of a building’s geometry to be in one of the many different proprietary formats for Computer Aided Design programs.

DECENTRALISATION

The proposed system is highly decentralised. It is possible, though not likely that each building would be described on a different computer. Once the system was running well, even the "Central Repository" of Metadata could be mirrored in a number of locations around the world to improve responsiveness. There is nothing inherently against the principle of decentralisation in having a single standard for classifying and locating data. This is merely the principle of the URI - the "Central Repository" is the Universal Resource that one uses to access Building Performance Data. It has a single unique Identifier on the web.

TEST OF INDEPENDENT INVENTION

If someone else had already invented your system, would theirs work with yours?[22]

What is important and indeed, essential about the proposed database of building performance is that all its pieces could be re-used by other better or different systems:

1. All the individual web locations with building performance data on them will be able to be used by any number of analytical systems.
2. The data in the proposed "Central Repository" is also accessible. It could be used either as a key to the translation between the proposed system and an alternate. As Metadata, and hence data it could also be used as just another data reference in an alternate system.

PRINCIPLE OF LEAST POWER

The rationale for this principle in web design is that the less powerful you make the language in which data is stored, the more each individual can do with the data stored in the language. As this system will be using the languages of the web, and as it seems unlikely that building performance analysts will be devising many new languages for the storing of data, this point needs only to be mentioned for completeness. However, if a building product model[23,24] was to be used as the language to store the performance data, then it would be necessary to re-examine this issue.

The web site that has been constructed for this QA system contains an illustration of how elegantly the XML system separates the **content** of the SuNREL thermal simulation program input file from its **presentation** with the use of a **data model** expressed as metadata in XML syntax in a DTD file.

CONCLUSION

This research and the associated web site (www.aecsimqa.net) defines a development path for the next generation of design decision support tools. It assumes this next generation of design tool will be more detailed computer programs.

A considerable advantage arises from the XML/RDF split in the presentation of data - on the web or anywhere else. This is the reasoning: the rules that define the relationships between parts of a building are explicitly removed from the simulation program revealing the reasoning behind the analysis very clearly. This separation has several benefits when seeking to apply a QA process in simulation.

First, an aspect of simulation that the new analyst often finds puzzling is determination of the appropriate external environment to "apply" in a simulation. What analysts debate is how to characterise the 'typical' external environment. Is it an average day/week/year? What might the risk to the building owner or operator be if the normally expected variations around the average occur from year to year?

Stochastically valid risk analysis is essential in all Quality Assurance procedures related to building performance simulation. In an XML system the weather data for a thermal or lighting simulation would contain the RDF definition of the meaning of its terms. This would enable an XML-aware simulation to translate the columns of weather information to a format compatible with its own views of the world. It would also mean that each weather file would contain synoptic information on how typical it was which could then be used by the simulation package to construct atypical weather scenarios.

A second and often-overlooked aspect of the external environment is the operational environment. The designer needs to know just how vulnerable the simulated performance will be to variations in the way we occupy or operate the building. XML format data on the energy performance of real or simulated buildings would contain Metadata. This would describe the context for the measurements and hence permit the XML front end of the simulation package to *infer* how "typical" the usage patterns.

Finally, the increased complexity of modern computer-based building performance simulation tools has not rid the design profession of its traditional problem with these tools: that they evaluate completed designs. Guidance about how to move forward in improving a design typically only comes only from the informed user looking backwards at how the existing design performs. An XML front end to a design process such as modelling a building in CAD could look up Post Occupancy Evaluation (POE) contributions to the Internet database. It might even generate initial design ideas based on successful precedents.

The web based QA system which went on-line in May 2001 (www.aecsimqa.net) currently focuses on thermal simulation data. It is being expanded during 2001 into lighting. In the future it is intended that it will operate in the fields of thermal, lighting and audio simulation.

REFERENCES

1. Cooper, Ian. **Barriers to the exploitation of daylighting in building design: UK experience** Energy and Buildings, **6** 1984. AND Sebastian Lera, Ian Cooper and James A Powell **Information and designers** Design studies, Vol 5 No 2, April 1984
2. O.O. Ugwu, C. J. Anumba, L. Newnham and A. Thorpe **Agent based decision support for collaborative design and project management**, The International Journal of Construction Information Technology, Special Issue: Information technology for effective project management and integration. 7(2), pp 1-16, 1999 (see http://helios.bre.co.uk/adlib/pubs/ for more information).
3. Jackson, Anthony **Reconstructing Architecture for the 21st Century - an inquiry into the architect's world.** Univ. of Toronto Press, p197, 1995.
4. Psychophysics is used to imply that in buildings responses to heat, light and sound are human responses, not merely the interaction of thermal, light and acoustic energy with building materials. It is a term that has far more proscribed meanings in Psychology.
5. Mackinder, Margaret and Heather Marvin **Design decision-making in architectural practice** BRE Information Paper IP11/82, BRE Garston, 1982:
6. A. Bailey **Discussion on a Diploma in Architecture**, in papers read to the Royal Institute of British Architects, 1856. AND Kaye Barrington in **Development of the architectural profession in Britain** Allen and Unwin, 1960.
7. Marvin, Heather **Using experience and publications in building design**. BRE Information paper IP13/85, BRE, Garston 1985. AND: **Meeting building designers' needs for trade information** BRE Information paper IP14/85, BRE, Garston 1985.
8. Burt Hill Kosar Rittelman Assocs. and Min Kantrowitz Assocs. **Commercial Building Design: Integrating climate, comfort and cost.** Van Nostrand Reinhold Co., New York, p 12, 1987.
9. Burt Hill Kosar Rittelman Assocs. Op Cit. p 15.
10. Purcell, A.T. **Ritualistic, rhetorical, reactionary**, Architecture Australia, July 1985
11. Mischler, Georg. **Rayfront, interface to Radiance** http://www.schorsch.com , 2001.
12. Donn Michael **Tools for quality control in simulation** Building and Environment, V36, No 6, pp673-680, July, 2001

14. Turing A. M. **Computing machinery and intelligence** Mind, Vol. 59, No. 236, pp. 433-460 1950.
15. Judkoff, Ron Op. Cit.
16. Gordon, Harry, Min Kantrowitz and Justin Estoque. **Commercial Building Design. Integrating Climate, Comfort and Cost** van Nostrand Reinhold, Co., Inc. New York, pp. 50-51, 1987.
20. SUNREL Building Energy Simulator 1.5, National Renewable Energy Laboratories, Golden Colorado (2001).
17. Burners-Lee, Tim *Principles of Design* http://www.w3.org:80/DesignIssues/Principles September 1999.
18. Burners-Lee, Tim *The Semantic Web* http://www.w3.org:80/DesignIssues/Semantic September 1998
19. Burners-Lee, Tim *Evolvability* http://www.w3.org:80/DesignIssues/Evolution March 1998.
20. *Metadata is machine understandable information about web resources or other things* in **Metadata Architecture** http://www.w3.org:80/DesignIssues/Metadata January 1997.
21. Burners-Lee, Tim **Principles of Design: Axioms of Web Architecture** http://www.w3.org:80/DesignIssues/Principles September 1999.

22. Burners-Lee, Tim **Principles of Design: Axioms of Web Architecture** http://www.w3.org:80/DesignIssues/Principles September 1999.
23. Lassila, Ora and Ralph R. Swick **Resource Description Framework (RDF) Model and Syntax Specification Specificationhttp://www.w3.org/TR/WD-rdf-syntax/** January 1999.
24. ***aecXML™ - A Framework for Electronic Communications for the AEC Industries;*** White Paper available at http://www.aecxml.com supported by Bentley Systems (http://www.bentley.com)