

AUTOMATED CAPTURE OF DESIGN KNOWLEDGE

**H.J. REA, G.ROBINSON, R.SUNG, C.SALAMON, J.M.RITCHIE AND
J.R.CORNEY**

kim13rep001hjr10.doc

ACCESS LEVEL: 3

ISSUE DATE: 1 February 2007

DISTRIBUTION LIST: *GameWare Developments Ltd, CBBC*

Today's CAD systems are excellent tools for recording the end result of a design process. But they are poor at capturing the knowledge developed during a design process. Because of this when changes in a component's form or function are required the knowledge about dependences within a design frequently need to be rediscovered because are not explicitly recorded when first encountered. Of course an engineer could manually log all the implicit dependencies and constraints that have been discovered during the design but this would inevitably would reduce short term productivity and be dependent on individual engineers being conscious of them.

However, if CAD systems could automatically capture the constraints learnt during the development of a design then, even years after the original design, engineers could better understand the consequences of potential upgrades and modifications. The hypothesis of the work reported here is that if all users actions were recorded in a structured fashion; then automated systems could deduce useful information about a design's development process. However there is an obvious problem with this scenario: a traditional CAD system contains only part of the overall process, frequently the physical performance testing and assessment of a potential design is conducted over long time spans outside the CAD environment.

So rather than use a mechanical CAD system this paper reports a methodology developed to investigate the potential of design logging using the purely virtual design and test environment found in the BAMZOOKi game. This BBC software allows users to design the geometry and movement of multi-legged robots and test their performance in a virtual environment with simulated physics (e.g. force, momentum). The paper describes how the standard BBC software is being modified to support the investigation of automated knowledge capture.

1. INTRODUCTION

An engineering drawing (or 3D CAD model) only defines one instantiation of a solution to a design process. Although many industries (eg aerospace and defence) have established systems for documenting the analysis done to verify the functionality of a given component, much of the knowledge developed during a design process is frequently lost. In other words, although the solution to a problem is recorded, the process by which this solution was derived and the constraints discovered during the process goes largely un-archived.

There are many reasons for this:

- the design process is frequently distributed over long periods;
- there is a reluctance to document mistakes and failures inherent in the discovery of constraints ;
- the method of documentation required is unclear;
- the additional time required to create documentation detracts from the process of producing the design.

This is unfortunate as a concise and accessible a summary of a component's design process would aid engineers who, sometimes, years later, have to revisit products to upgrade or modify them. For many years these problems have appeared intractable but the last decade has seen the migration of all aspects of the design process from paper onto networked computers. Consequently the opportunity now exists for a previously unimaginable amounts of data related to product development to be archived. However the challenge is to provide a structured record of a design process that allows the details to be effectively abstracted, searched and browsed.

Motivated by these needs the authors are undertaking research into the automated capture (and re-use) of design, design activity and design process data. *The goal is to automatically capture more the process by which a particular design was reached (and the lessons learned by the designer during this process).* However a practical difficulty faced by all efforts to study and record the design process is that "real" engineering design activities are spread both temporally and physically with testing frequently occurring in the field (and so is remote from desk bound observers of the design process).

To allow progress towards the goal of automated design rational capture, the authors looked for a non-trivial design environment that had the following properties:

- supported all stages of the design process from conceptualization to functional detailing;
- incorporate an intuitive graphical interface (allowing non-expert users to participate in the experiment);
- a short design cycle that incorporates integrated in-line testing;
- a feasible, competitive and engaging design goal;
- measurable design success;
- innate design criteria; and
- accessible data logging.

The CBBC BAMZOOKi application was identified as providing a self-contained 'design-build-test' environment; the structure, operation and interface of which make it an ideal platform for this research. The BAMZOOKi application allows users to construct virtual creatures (zooks) from building block elements and modify the properties of these elements to change either the form or motion of the zook. The environment also includes a test section where the performance of the zook can be measured against a set of standard tasks.

If (real-time) design and process data could be captured while people use the BAMZOOKi application to design (or modify) zooks to achieve a target performance in these tests then it should be possible to post-process this captured data to extend the capability of current design practice either through the automatic extraction of higher level information about the design itself, or by providing an accurate description of the problem solving approaches used. The version of the BAMZOOKi software distributed by the BBC does not support any continuous user logging, but discussion with the software's developers [1] suggested that this would be possible if a detailed

specification was provided. Such a specification was developed and is reported in this paper.

The rest of this paper is structured as follows: Section 2 briefly outlines the literature and academic context of the work. Section 3 describes the BAMZOOKi program. Section 4 outlines the proposed experimental methodology. The experimental procedure is defined in Section 5, and a discussion and some conclusions are drawn in Section 6.

2. LITERATURE REVIEW

The engineering design process has been extensively studied by academic researchers who (based largely on manual observation) have proposed various models in attempts to characterize and structure the activity. Typical of these is the model, similar to that depicted in Figure 1, proposed by Wächtler reported in [2] that explicitly models the learning process that develops a design:

“The ‘learning system: designer’ takes a problem from, and returns a solution to, the environment. Discursive and intuitive actions produce solutions (ideas) that are held in the short-term store of the learning system. A comparison of the proposed solution with the environmental demands (requirements) may throw up discrepancies calling for new decisions and hence lead to new actions. Once the discrepancies are reduced to a minimum, the optimum solution is to hand. A cycle in this process of optimisations is called a learning element. The learning system must not be considered in isolation from the environment. In other words, the environment not merely imposes the requirements and receives the solutions, but frequently plays a crucial part in finding the latter.”

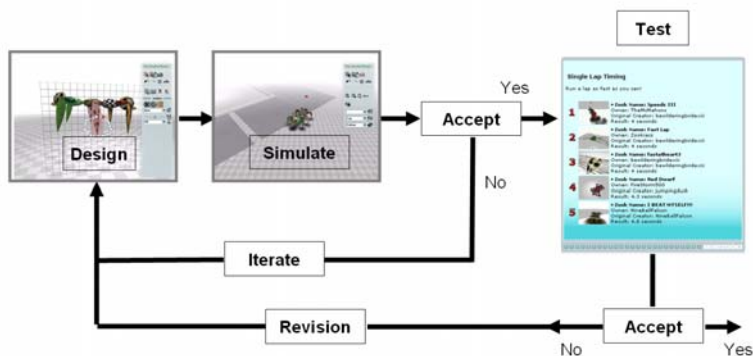


Figure 1 BAMZOOKi Design Process

However the veracity of many academic models of the design process is questionable and it is unclear if this neatly compartmentalised view can be mapped in a meaningful way onto real design activities. In this context several researchers have reported work that seeks to automatically record the details of what designers actually do.

In [3], the authors describe a method of recording design activity on a PC, capturing information such as file names, file creation/destruction times, web pages viewed, screen shots, text input, CAD model construction, meta data of files opened etc. In addition interface information is observed, these are called Profile Elements (PE's).

Monitoring “Information Use PE’s” are good for capturing and recording the order of execution where as “Interface Action PE’s” (user behaviour) enable information “push” such as providing links to relevant documents as a user undertakes a task.

Stumpf et al [4] describe a system which monitors user behaviour in a similar approach to Campbell et al [3], whereby interactions with Microsoft windows are monitored. The actions are tracked, and a prediction is made as to what Task is being carried out, and hence which folder to place any documents produced in. In this case the object is not to collect knowledge, but merely assist knowledge workers to organise task related information. The concept of costs to the user is identified, where the costs can be physical, mental, time, and/or cognitive. There is an initial cost for the user to set up the system and identify the Tasks that they work on. The authors also raise the issue of the users perceived loss of control when predictions are made, and suggest that predictions are only used if there is high confidence in the results.

Chen et al [5] have shown that for 50% of users in their test there was a 0.8 correlation between eye gaze and cursor location. So in many cases it is possible to predict a users focus of attention based on the location of the cursor on a screen.

Varga et al [6] provide a survey of hand motion processing (HMP) technologies with a view to adapting hand motions as an input for 3D conceptualisation. Their analysis is to identify technologies that will provide real-time information extraction and conversion. They compare contact vs non-contact methods. Contact methods include gloves, stylus and non contact are based on 2 cameras or camera and shadow analysis.

If we are able to unobtrusively monitor designer’s behaviour and progress, hence deriving the context of the users work, there is also the possibility of providing context aware information either on a push or pull basis. Investigations of the advantages of push and pull information systems for a mobile unit have been carried out by [7]. In this case context was derived from the physical location of a mobile information display unit via GPS.

Pioneering work at Heriot-Watt University has demonstrated the use of immersive virtual environments to derive design process knowledge from logs of user activity. Specifically work focussed on cable harness routing [8] and simple assembly tasks [9]. Mueller [10] and Chen [5] have shown that simply tracking a cursor location can provide knowledge about user activities. Jin and Ishino [11] use the combination of user activity logging and data mining to automatically generate design activity knowledge.

The current state of the art in automated knowledge capture at this time appears to rest with the digital enterprise industries. Implementations of automated profile generation in Amazon [12], Google [13] and Myspace [14] show advanced methods in automated capture and analysis of user activities, where there is no cognitive or time costs for the user. For example Myspace provides additional knowledge tools in the form of peer-to-peer networking. In the academic arena parallel work on knowledge capture and user profiling for multiple software packages is being carried out by Campbell [3, 15], and Stumpf [4]. While in the area of Artificial Intelligence research into connectionism and the nature of virtual communities is being carried out by Memmi [16].

3. BAMZOOKi

BAMZOOKi was developed for a BBC TV children's programme, and the toolkit is available free for download and standalone use. Zooks are virtual creatures that users create to perform with specific mechanical behaviours. Users can then send their zooks into battle with other players' creations, and watch to see who wins. Winning depends on players developing a complex understanding of how to create a zook, and what behaviours to design it to perform. Competitions on the website include sprinting and hurdles.

The BAMZOOKi Zook-kit is an intuitive 3D design environment. The interface allows interaction with 3D models based on “blobs”. Initially the model is a single blob representing the Zook’s body as in Figure 2(a). The user can change a number of variables such as: length; width; “squareness”; and “pointiness” to customise the shape of the blobs either moving the sliders on the menu, or dragging key nodes on the model itself. The user can then add blobs, or pre-defined limbs constructed of blobs, to the body as in Figure 2(b). Mirroring tools allow the rapid population of limbs (see Figure 2(c)). In addition to shape variables, blobs have movement variables that can be accessed through the motion tab on the edit menu (Figure 2(d)) and on screen (Figure 2(e)). At any point during the design cycle, the user can test the Zook by running a simulation animation of the Zook using the test tools (Figure 2(f)).

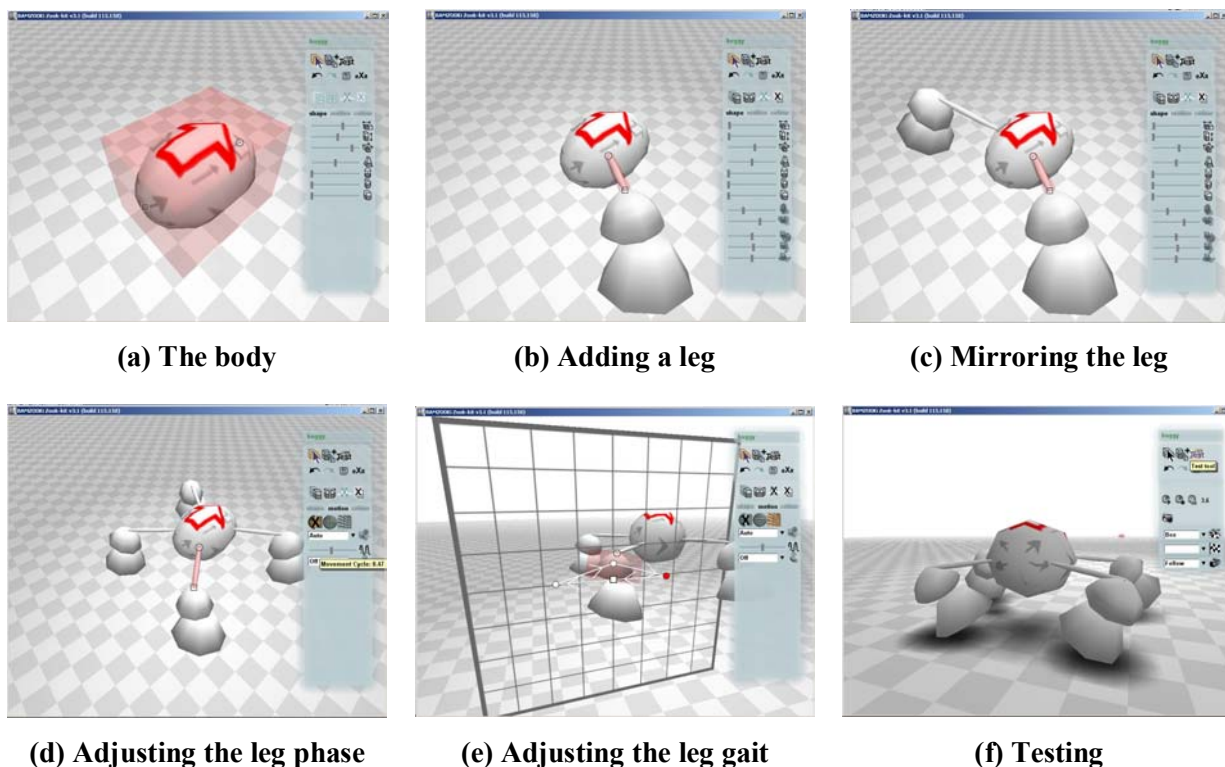


Figure 2 Creating a Zook

Despite its ease of use, BAMZOOKi offers an impressively rich design environment. As one of its developers said:

"BAMZOOKi's bespoke integrated artificial life technology suite combines physics, artificial intelligence and 3D graphics for anyone who can point and click

a mouse. Anyone can play and compete with BAMZOOKi's entertaining and educative system to learn the art of physics, movement, and mechanics with ease through CBBC's engaging online and TV mass media channels. It's simple, fun to use and freely available through CBBC online, yet it's one of the most powerful forms of cyber-plasticine available anywhere." [17]

4. THE EXPERIMENT

4.1 Aims

The objective of logging the evaluation sessions is to investigate the processes by which people “discover” what aspects of their design (ie the Zook) can be best changed to meet a specific performance goal. There are many options open to the designer to improve the Zook's speed: weight reduction; leg lengthening; change of gait etc. Each designer will “discover” which of these parameters has the most effect. The research challenge is to extract this “knowledge” created during the design process from the log files.

4.2 Methodology

Each participant in the experiment will be given a default Zook (known as the “Baseline Zook”) with a performance goal they are required to meet (eg perform a the sprint test at 40cm/s). The participants will then proceed to modify the design of the Zook to meet this specification.

4.3 Data Logging Specification

Figure 3 shows an outline of the expected data flow in this data capture and re-use scenario. The first phase of the process involves capturing a detailed record of the design activity (logfile) and the design output produced (Zook model geometry and performance data). The second phase involves interrogating this complete data set to extract specific data that is relevant to certain research questions and objectives. These extracted (focused) data sets are then reformatted for importing into different downstream applications so that detailed analysis can take place.

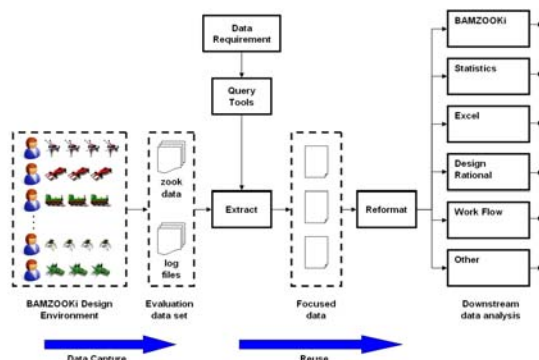


Figure 3 Overview of Data Flow

The objective of the logging capability is to unobtrusively capture a complete record of each BAMZOOKi design session. To be effective during downstream analysis the

recorded files must include sufficient detail that enables the session activity and evolution of the design to be accurately reconstructed solely from the information contained within the files. It must also be possible to identify all of the files that were created during a single session e.g. separate model geometry file, and the corresponding data (e.g. based on time) from the different files.

The log data will be stored in XML (Extensible Markup Language) which is a W3C-recommended general-purpose markup language currently used in the BAMZOOKi application [18]. The primary purpose of XML is to facilitate the sharing of data across different information systems and XML languages are easy to design and to process. The XML data can be post-processed to conform to a lexicon to become an efficient information model. One useful lexicon may be the Decision Support Problem Technique (DSPT) developed by [19], which “provides a framework to represent typical decisions or tasks in a product realisation system”. A draft sample of an XML log file that will be generated by a BAMZOOKi session is shown in Figure 4.

```

<!-- draft sample logfile for logging Zook-kit-->
<logfile>
<header>
<session_ID text="id text"/>
<session_start>      <time          hour="hh"          minutes="mm"          seconds="ss"
fraction="tt"/></session_start>
<session_control filename=""/>
<baseline_zook filename="test.zook"/>
<control_text> elements copied from under the CopyToLogHeader element in the session
control file </control_text>
</header>

<body>
<event desc="UI Only">
<time hour="hh" minutes="mm" seconds="ss" fraction="tt"/>
<UI_interaction><mode_select name="edit"/></UI_interaction>
</event>

<event desc="UI Only">
<time hour="hh" minutes="mm" seconds="ss" fraction="tt"/>
<UI_interaction><main_panel_select name="shape_tab"/></UI_interaction>
</event>

<event desc="Zook">
<time hour="hh" minutes="mm" seconds="ss" fraction="tt"/>
<UI_interaction element="slider">
<parameter_change body_part="3" yaw="2.34"/>
</event>
<event desc="Zook">
<time hour="hh" minutes="mm" seconds="ss" fraction="tt"/>
<UI_interaction type=zook_body_part="3"/>
<parameter_change body_part="3" pitch="22.55"/>
</event>

<event desc="Zook">
<time hour="hh" minutes="mm" seconds="ss" fraction="tt"/>
<UI_interaction type=zook_body_part="3"/>
<parameter_change body_part="3" yaw="3.34"/>

```

Figure 4 Exemplar of Design log file

5. PROCEDURE

As explained in the methodology section, each participant starts a design with a “Baseline” Zook and a “goal” for its performance. The next section details these:

5.1 Baseline Zook

The BAMZOOKi session starts with a baseline zook (revision 0). The overall design of the zook used may change but in the first instance will be an eight legged ‘spider’, see Table 1.

Table 1 Baseline Zook Properties

Physical (approx. values):	Performance:
Height: 10 cm	Sprint: 40 cm / s
Width: 30 cm	Block Push: 25 cm
Length: 30 cm	Hurdles: 5 cm / s
Weight: 1 kg	
Components: 22	

The baseline zook has a ‘below-average’ performance in the three test tasks (sprint, hurdles, block push) and its design includes specific ‘poor’ elements that can be identified and modified to improve this performance.

Several variations of the baseline zook may be used to provide different levels or complexities of starting ‘problem’. Example zook problems and solutions are detailed in Table 2.

Table 2 Possible Zook Problems and Solutions

Problem:	Potential Solution:
<p>Sprint: The zook walks rather than runs. The zook takes small steps. Legs are too close to increase stride length. Legs are too short to have a big stride length. Zook appears to ‘limp’ or is ‘jerky’.</p>	<p>Increase the gate cycle speed. Increase the stride length. Increase body length and increase leg spacing. Increase leg length. Adjust phase of legs within the gait cycle.</p>
<p>Block Push: Zook is too small to push the blocks. Zook legs contact the blocks rather than the body / head. The head makes contact but zook skids on the blocks.</p>	<p>Make the zook bigger. Make head larger or change angle of front legs, (etc). The front of the zook has a point contact, make the jaws larger and wider to produce a more stable contact.</p>
<p>Hurdles: Zook gets stuck on the larger hurdles. Zook cannot get onto the larger hurdles. The zook is unstable and sometimes falls over.</p>	<p>Increase the zook’s ground clearance. Change gait cycle to increase step height. Increase width of the zook.</p>

5.2 Goal

The goal of the participants is to make modifications to the zook that will lead to it achieving a higher ‘Gold’ level of performance. This ‘Gold’ performance must be set at a realistic level (average) that can be achieved by a typical user within the timescale of a test session. (This is unlikely to be a best in class performance and further improvements may be expected from a longer session or with more experienced users). An example of a set of ‘Gold’ performance parameters is given in Table 3.

Table 3 'Gold' Performance Parameters

Physical (approx. values):		Performance:	
Height:	no limit cm	Sprint:	100 cm / s
Width:	no limit cm	Block Push:	35 cm
Length:	no limit cm	Hurdles:	75 cm / s
Weight:	2 kg		
Components:	22		

5.3 The BAMZOOKi Design Session

Each design session has three distinct phases: start, design and exit.

5.3.1 Start

The participant starts the BAMZOOKi session through a ‘start’ dialog box (1) with a baseline zook automatically loaded. Each session has a unique logfile, this is created and the ‘header’ data completed at the start of the session.

5.3.2 Design

The participant looks at the ‘My Goal’ performance targets for the zook. ‘My Goal’ is a new ‘pop-up’ dialogue box that is activated from the ‘general zook menu’. ‘My Goal’ lists the target ‘Sprint’, ‘Block Push’, ‘Hurdle’ goals for the session. Then the sequence of action is typically as follows.

1. The participant tests the baseline zook to see what its current performance is.
2. The participant modifies some aspect of the zook model geometry.
3. The participant tests the zook to see what effect this has had on performance.
4. The participant continues to modify and test the zook until the ‘Gold’ zook performance target is reached. (Design Cycles).

All of this activity is recorded in the log file as it happens with a time stamp of when it happened.

5.3.3 Exit

The participant exits the session. This causes the log file footer data to be added to the log file and the log file to be terminated.

6. DISCUSSION AND CONCLUSIONS

The software modification required to generate XML logs of user activity with the BAMZOOKi software are being implemented by Gameware during February 2007. The following activities are planned with the bespoke software.

- 10-11 March 2007 Initial experiments at the National Engineering and Science Week exhibition, Stratosphere, Aberdeen
- 2-15 April 2007 Main experiment at Edinburgh International Science Festival
- October 2007 Combined experiment with the University of Strathclyde to include social interaction data capture

The design log files generated by these activities will form a unique data set that will be freely available to design researchers. Post processing of the captured data files will involve using ‘query tools’ to allow a researcher to define parameters that identify specific data to be extracted from all the files that make up the evaluation data set. In other words tools will be developed that allow a “trail and error” design process to be mined for information about effective performance enhancement strategies.

This data extraction will produce focused data sets only containing the data of interest to the researcher at that time. For example Figure 5 illustrates the process to extract all the “hurdle” performance data (cm/s) and the model data for all of the zooks that had a “hurdle” test result (including its design cycle number). But any part of the data set could be extracted depending on the current goals of the research. This focused data set is then reformatted so that it is compatible with the importing requirements of whichever package will be subsequently used to undertake the analysis of this extracted data.

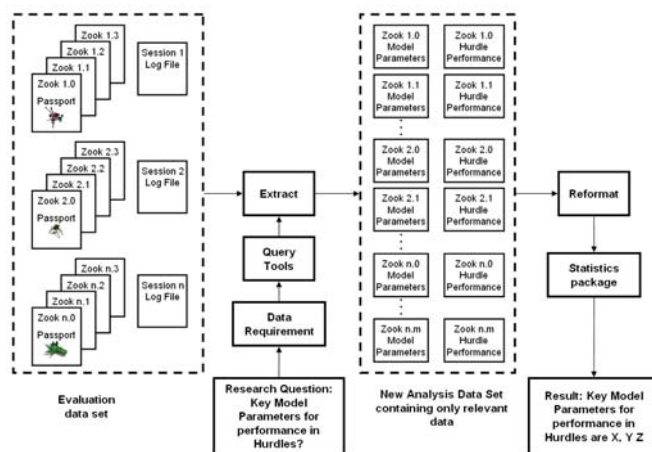


Figure 5 Example of Data Extraction Process

7. ACKNOWLEDGEMENTS

We wish to thank the BBC and CBBC for giving us permission to use BAMZOOKi for our research, and Dylan Barnarse of Gameware Development Ltd for his assistance in creating the specifications for the bespoke BAMZOOKi program.

Gameware Development is a pioneer of cross media artificial life technology, creating exciting games and interactive communications for TV, PC, console, wireless and hand-held games devices. Based in Cambridge in the United Kingdom, Gameware Development comprises of Creature Labs, the A-Life division fuelling Gameware's development projects, and SceneMachines the mobile technology division evolving and inventing cutting edge intellectual properties to entertain users on the move. www.gamewaredevelopment.com

8. REFERENCES

1. Gameware_Development_Ltd, *BAMZOOKi Zook-kit*. 2006, BBC: Cambridge.
2. Pahl, G. and W. Beitz, *Engineering Design, A Systematic Approach*, ed. K. Wallace. 1988, London: The Design Council.
3. Campbell, D.R., et al. *A Methodology for profiling computer based design activities*. in *International Conference on Engineering Design (ICED 05)*. 2005. Melbourne.
4. Stumpf, S., et al. *Predicting user tasks: I know what you're doing!* in *Workshop on Human Comprehensible Machine Learning at the Twentieth National Conference on Artificial Intelligence (AAAI-05)*. 2005. Pittsburgh, PA.
5. Chen, M.-C. and J.R. Anderson, *What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing*, in *CHI '01*. 2001, ACM Press: Seattle, Washington.
6. Varga, E., et al. *Survey and Investigation of Hand Motion Processing Technologies for Compliance with Shape Conceptualization*. in *Design Engineering Technical Conferences and Computers and Information in Engineering Conference, DETC'04 ASME*. 2004. Salt Lake City, Utah: ASME.
7. Cheverst, K., K. Mitchell, and N. Davies. *Investigating context-aware information push vs. information pull to tourists*. in *Third International Workshop on Human Computer Interaction with Mobile Devices, Mobile HCI 01*. 2001. Lille, France.
8. Ng, F.M., *Virtual Reality and Computer-Based Tools for the Routing of Cable Harnesses.*, in *Mechanical and Chemical Engineering*. 1999, Heriot-Watt University: Edinburgh.
9. Ritchie, J.M., R.G. Dewar, and J.E.L. Simmons, *The Generation and Practical Use of Plans for Manual Assembly using Immersive Virtual Reality*. *IMechE, Journal of Engineering Manufacture (Part B)*, 1999. **213**: p. 461 - 474.
10. Mueller, F. and A. Lockerd, *Cheese: tracking mouse movement activity on websites, a tool for user modeling*, in *CHI '01 extended abstracts on Human factors in computing systems*. 2001, ACM Press: Seattle, Washington.
11. Jin, Y. and Y. Ishino, *DAKA: design activity knowledge acquisition through data-mining*. *International Journal of Production Research*, 2006. **44**(14): p. 2813-2837.
12. www.amazon.com. [cited.
13. www.google.com.
14. www.myspace.com. [cited.
15. Campbell, D.R., S.J. Culley, and C.A. McMahon. *Observations on Data Relating to User Browsing Behaviour and Usage of Electronic Information*. in *IDMME 2006*. 2006. Grenoble, France.
16. Memmi, D., *The nature of virtual communities*. *AI Soc.*, 2006. **20**(3): p. 288-300.

17. Saunter, I. http://www.gamewaredevelopment.co.uk/news_index.php. 2005 [cited.
18. <http://en.wikipedia.org/wiki/XML>. 2007 [cited.
19. Kulkarni, R., et al. *An Information Model for Finding and Integrating Distributed Resources for Engineering Design-Manufacturing Processes*. in *Proceedings of 2002 ASME Design Engineering Technical Conferences, DETC2002/CIE-34464*. 2002. Montreal, Canada: ASME.