

Knowledge Management for a Large Service-Oriented Corporation^{*}

Sylvia C. Wong, Richard M Crowder, Nigel R Shadbolt and Gary B Wills

School of Electronics and Computer Science,
University of Southampton, UK
{sw2,rmc,nrs,gbw} at ecs.soton.ac.uk

Abstract. The design and maintenance of complex engineering systems such as a jet engine generates a significant amount of documentation. Increasingly, aerospace manufacturers are shifting their focus from selling products to providing services. As a result, when designing new products, engineers must increasingly consider the engine's complete life-cycle as part of the design process. To identify possible areas of concern, engineers must obtain knowledge gained from the entire life of similar engines. However, because of the size and distributed nature of the company's operation, engineers often do not have access to front-line maintenance data. In addition, the large number of documents accrued makes it impossible for them to be examined thoroughly. This paper presents a prototype knowledge-based document repository for such an application. It searches and analyzes distributed document resources, and provides engineers with a summary view of the underlying knowledge. The aim is to allow engineers to incorporate maintenance issues into the initial design. Unlike existing document repositories and digital libraries, our approach is knowledge-based, where users browse summary reports instead of following suggested links. To test the validity of our proposed architecture, we have developed and deployed a working prototype.

1 Introduction

The design and maintenance of large and complex engineering systems requires a significant amount of documentation, particularly if the system being considered is a turbofan jet engine used on the current generation of aircraft. The jet engine is amongst the most complex machine ever designed, incorporating a wide range technologies of including high temperature materials, complex fluid dynamics and high speed rotating components.

A fundamental shift is currently occurring in the aerospace industry away from selling products to providing services. Companies such as Rolls-Royce aims to make half its engine fleet subject to long-term maintenance service agreements by 2010 [1]. Essential to the success of this market shift is the significant cultural change from *offering a service to support a product* to *designing a service and*

^{*} In Proceedings of the 6th International Conference on Practical Aspects of Knowledge Management (PAKM), Vienna, Austria, 30 Nov - 1 Dec 2006

the product to support it [1]. In other words, new products must be designed to provide lower and more predictable maintenance costs. To minimize maintenance costs throughout the engine's life cycle, engineers must obtain knowledge gained from maintenance histories of similar engines during the design phase of a new engine. This will help engineers identify parts most likely to be problematic throughout the new engine's entire life cycle, possibly eliminating the problem during the design phase. It should be noted that engine design is typically undertaken by a number of teams who are responsible for individual engine modules, e.g. compressor, turbine. Therefore it is impossible for any single member of a design team to access more than a fraction of available documentation. As is widely recognized, information systems usually develop over time into a set of heterogeneous resources. As a result, it becomes difficult for engineers to follow a trail through these resources [2]. The challenge for organizations is therefore to develop an information system that is both comprehensive and will satisfy the increasing demands from industry for up-to-date and easily accessible information.

In response to these challenges, we are implementing an intelligent, knowledge repository to support engineers to design for better performance and lower maintenance costs throughout the engine life-cycle. The knowledge repository searches and analyses relevant maintenance records and design guidelines, and provides design engineers easy access to such information. It is hoped that the summary reports provided by the intelligent document repository will help engineers in creating design documents that incorporate these issues into the design requirements.

This paper is organized as follows. Section 2 explains how knowledge can be re-used within an aeroengine manufacturing company. It also introduces a scenario explaining how maintenance records can be used to help to improve the reliability of both existing and new products. Section 3 gives a brief overview of knowledge management. Section 4 then discusses other works that aim to help user to discover knowledge by integrating heterogeneous documents sources. Section 5 introduces our proposed architecture for an intelligent document repository. This is followed by Sections 5.1 to 5.3, which describes the underlying components of the proposed document repository. A discussion is presented in Section 6 with conclusions in Section 7.

2 Scenario

As is well recognized in engineering design, the use of past experiences and previously acquired knowledge, either from the designer's own experiences or from resources within their organization forms an important part of the design process. It has been estimated that 90% of industrial design activity is based on variant design [3], while during a redesign activity up to 70% of the information is taken from previous solutions [4]. A cursory consideration of these figures identified two immediate challenges — how to capture knowledge, and how to retrieve it. The purpose of the work reported in this paper is to develop an

information system that can be used within a manufacturing organization for the retrieval of knowledge from across the organization to support the design activities.

The following scenario illustrates the potential use and benefit from such an information system. The scenario involves three separate and different groups of users that are involved in the life of a jet engine. Front-line maintenance engineers are involved in the day to day servicing of the engines, and thus responsible for populating the document repository with maintenance reports and other similar documents.

During the regular pre-flight checks, a flight crew reported a problem caused by a leak from an engine's bleed air system. Subsequent inspection which required the removal of the engine revealed that a duct had failed at a joint due to vibration. After repair, the engine was returned to service, and a full maintenance event report submitted to a document repository.

Traditionally, after production and sales, responsibility of maintenance of the engine passed from the manufacturer, to the airlines, who own the engines. This maintenance activity is supported by the manufacturers technical support and operations team. Technical support and operation engineers are responsible for improving the performance of existing engines. They can use information collected in the repository to monitor trends that develop over a fleet of engines. Modification can then be designed to mitigate any problems found:

Following a review of the maintenance events relating to a specific engine fleet, a trend was noticed in the high than expected number of failure of a bleed air duct joint due to vibration. To maintain the reliability of the engine fleet, a modification was developed and implemented.

Design engineers are currently remote from the problems experienced in the field by operations. Due to the importance of increasing operational reliability and minimizing maintenance costs in the new market paradigm of product support, information gained in the operation of a fleet of engines needs to be fed back to the designers of subsequent engines. However, the current information infrastructure makes this difficult as design engineers do not have access to maintenance knowledge. Our work aims to strengthen and help formalized the information flow between aftermarket operations and the design teams. It allows knowledge gained during the design, production and operation of an engine to advice the design of the next variant:

The design team for the next variant of this engine reviews the performance of the air bleed system across the fleet to learn from previous design rationale and operational history. The designers noticed a higher than predicted failure rate on a particular joint. Subsequent finite element analysis showed that a joint failure could occur due to vibration if certain operational conditions were met. It was therefore decided that the future variant of the engine would both eliminate the joint and reroute the duct

work. The revised design costs 50% more than the original. However, the saving over the life of the engine will be substantial due to lower likelihood of in-service failure.

The goal of our work can thus be summarized as follows: To feedback and harvest knowledge gained from the aftermarket operations documents to help (a) operations engineers in designing modifications to existing engines, and (b) design engineers in designing the next variant engine for the aftermarket.

3 Knowledge Engineering

Organizations have become increasingly concerned with knowledge management [5], amassing large amount of information into their corporate memory [6]. The aim is to use this repository of information to inform future discussions, decisions and activities. Figure 1 shows the four key activities in knowledge management – knowledge creation, knowledge mapping, knowledge retrieval and knowledge use.

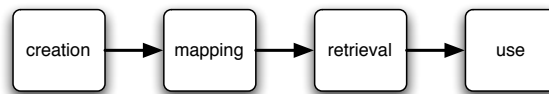


Fig. 1. The four key activities in Knowledge Management.

The first step in creating a knowledge system is knowledge creation. In our scenario, this is the documentation written by engineers throughout the life time of an engine, from design to maintenance. In practice it can take many forms, including formal design reports, e-mails submitted by airlines and field service personnel, detailed inspection reports and the engine performance over individual flights recorded by engine monitoring systems.

Knowledge mapping involves creating an *ontology*, which is a *specification of a conceptualization* [7]. Gruber explains that a common ontology defines the vocabulary with which queries and assertions are exchanged among agents (people or software). The ontology sets out all the entities (objects or concepts) that we are interested in and the relationships that connect these entities together. This is intended to be a *pragmatic* definition, i.e. it defines the vocabulary that is actually *in use*, and the concepts that are *useful* in problem-solving. It does not give the deep underlying philosophical vision of the fundamental entities in the field. Hence, in knowledge management, an ontology is a tool, whose quality is entirely dependent on its usefulness.

Retrieval is the step that transform mere *information* into *knowledge*. We believe that knowledge is relevant information delivered at the right time and

context [8]. To deliver this knowledge, information needs to be semantically enriched so that it can be better reused. When a knowledge system has a shared ontology for its disparate information resources, software agents can handle the semantically enriched resources consistently. Thus semantics and ontology together help deliver the right information at the right time, hence generating knowledge.

In our scenario, knowledge use occurs when engineers apply the knowledge gained from the repository in their work. For example, a design engineer applies the knowledge gained to create a better engine for the aftermarket. Or an operations engineer discovers a recurring, but minor problem with an existing engine, which would indicate a larger problem than each individual incident suggests.

4 Related Work

The work described in this paper is an extension on our previous work with Rolls-Royce. In [9], we presented a future vision for the working practices of designers within a manufacturing organization. We have found that engineering design environments are highly distributed in nature and are characterized by a large number of information sources. It is concluded that a range of knowledge management tools would be required to support this future vision of engineering design environments [10]. Therefore one of the objectives of our current work is to define a future engineering design environment, with particular emphasis on the social and technical systems that will support designers in their day-to-day activities.

In [2], we created a document repository from distributed and heterogeneous engineering document resources. When an engineer searches for documents within the repository, the system generates a list of documents ordered according to the engineer's role and its related concepts. Thus, the document retrieval process is *intelligent* and adapts to the user's role within the company. However, we found that engineers actually want the knowledge that is buried within the documents, instead of the actual documents themselves. This is due to the large volume of documents available within the repository which makes it very time consuming to peruse thoroughly. In other words, engineers prefer to see *summary reports* of documents archived. For example, when searching for engine part deterioration, engineers want to see how many times the deterioration of a particular part leads to engine removal, but not the list of original maintenance documents. Thus, for our new information system, we aim to include the ability to provide analysis of information stored, in addition to simple document search.

Our work in [2] can be seen as a digital library, with the extension where information presented is adapted to the role of the user. Digital libraries concentrate on the problem of searching for documents distributed over multiple repositories. For example, Priebe and Pernul [11] developed a portal over multiple document repositories by using an integrated metadata store. In contrast, search functionality does not form part of our proposed infrastructure. How-

ever, global document searches can be provided to our knowledge repository as *services* that implement document indexes and metadata indexes.

Another area digital libraries concentrate on is dynamic links generation to relevant document resources [12, 13]. Dynamic links are injected into documents automatically during presentation time, and does not alter the original documents. These dynamic links can point to related documents, or even services such as searching annotation and peer reviews [13]. In comparison, our proposed system does not perform dynamic link injection on existing documents. However, it can provide a list of suggested documents as a *service*.

Finally, there are also projects working on creating new semantically marked up data for large knowledge repositories, such as [14] and [15]. Creating new documents is outside the scope of our project, as we concentrate on the problem of delivering knowledge from existing documents. However, employing techniques to generate semantic information automatically for new documents to be deposited into the knowledge repository will improve document analysis and searching inside our framework.

5 Architecture

Two key technologies are used in integrating the distributed and heterogeneous data sources available from the document repository – Web Services and the Semantic Web. Web Services are loosely coupled standards-based components that perform specific tasks. Using Web Services, applications can be developed and deployed incrementally. In addition, new features can be easily added after the system is deployed. This modularity and extensibility make it especially suitable as a platform for an integrated knowledge repository within large engineering organizations. The Semantic Web is an application of the World Wide Web aimed at computational agents, so that *programs*, and not just humans, can interpret the meaning of documents on the Web (or an intranet). This allows the Web to be used for more than a human-browseable repository of information. The basis of this interpretation is an ontology, a structure which forms the backbone of the knowledge interpretation for an application.

The system architecture diagram of our knowledge repository is shown in Figure 2. Existing web standards are used wherever possible, to maximize tool reuse, compatibility and portability. A web portal provides an integrated user interface, and handles authentication and workflows. The user interface is formed by a series of portlets — web applications that run within a portal. In general, portlets appear to end users as rectangular sections of a web page offering a limited set of information. The portlets access one or more web services. We envisaged that these web services are provided by different departments within the company, and can be distributed across multiple sites. These web services will perform more than document searches that are already available in all document repositories. For example, it can provide an interface to existing manufacturing systems such as PDM (Product Data Management), or allow access to analysis tools such as Life-Cycle Cost Models and Finite Element Analysis. Documents

in the repository use the Semantic Web data format RDF [16], with its schema in OWL [17].

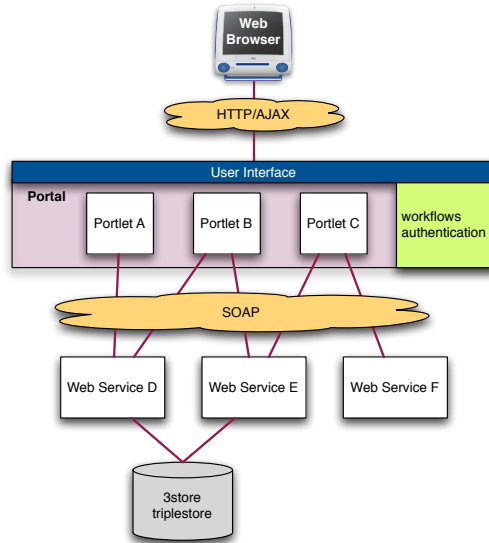


Fig. 2. System Architecture of the proposed knowledge system.

The remainder of this section discusses each component within the architecture in greater details.

5.1 Documents and Ontology

For each maintenance event, maintenance engineers document information surrounding the event for later references. A maintenance event occurs whenever actions are performed on an engine. Usually, each maintenance event involves multiple actions. Information documented includes the circumstances of the event, actions taken, parts installed and removed and any other findings observed. The engineers who created these documents are located in numerous sites around the world, in the manufacture's or third party repair bases or at airports. As a result, the maintenance documents are in a large variety of formats. Some of these reports are in the form of unstructured Word documents, with different sites using different templates. Other engineers record maintenance information using a centralized Service Data Manager (SDM).

To enable machines to interpret meanings stored within the documents, we need an ontology that captures all the terms and concepts used. Moreover, since the document repository is to be used by both design and service engineers, the

ontology should capture concepts from engineers working in both areas. The ontology is created by analyzing existing documents and conducting knowledge acquisition interviews with engineers [2]. The engineers interviewed are carefully selected and are domain experts from several different specializations. During these interviews, the *card sort* technique is used to help the engineer show how they used different document types and the relationships between these documents. The result of these interviews enabled us to identify, by specialization, the main concepts and the associated keyword for these concepts used by the particular type of engineer when searching for information.

The resulting ontology contains concepts ranging from engine deterioration mechanisms, engine models and parts to airport locations. Using this ontology, we have populated the triplestore with maintenance records from both the SDM database and reports provided as Word documents. Currently, our triplestore contains approximately 3250 maintenance events, with around 31,000 actions. This information equates to 389,000 rows in a SQL database table. The space usage in the table is 12MB in data and 25MB in index (ie total of 37MB).

5.2 Web Portal

Users access our knowledge framework via the web portal. The portal uses username/password authentication, and role-based access control. Using role-based access control, we can customize the content of the portal according to the engineer's specialization. The roles defined in the system can reflect job functions of the engineers. Thus, engineers with different specializations can be served a different set of portal pages. For example, when studying deterioration mechanisms, maintenance engineers can be presented with information from individual engine parts, while design engineers will be presented with overall information for the entire fleet. This difference arises from the different tasks the engineers have to perform. Maintenance engineers are interested in locating and fixing problems of existing engines in service. On the other hand, design engineers are interested in overall performance to aid the design of future variants. Another possible customization can be used for navigation/drill down over the engine taxonomy. Since an engine contains a large number of parts, a simplified taxonomy for navigation will greatly speed up information browsing. The navigation taxonomy can be a subset of the entire engine taxonomy, based on the engineer's specialization.

A questionnaire based study was carried out to better understand what kind of knowledge engineers would like to discover from the knowledge web [18]. In the questionnaire, engineers are presented a list of questions relating to maintenance experience with a product. They are asked how often they might ask them when designing a new product. They are also asked what other questions they might want to ask. The result of this questionnaire tells us what are the most important and most common life cycle information design engineers seek from maintenance documents.

Based on the results of the user study, we developed two portlets that help engineers harvest knowledge from the data sources. The first portlet is a main-

tenance event browser that allows engineers to follow an engine taxonomy to retrieve lists of maintenance events, view timelines of events and obtain detailed information about a particular event. The second portlet allows engineers to browse a list of parts that dominate maintenance statistics. Figure 3 shows screenshots of the two portlets.

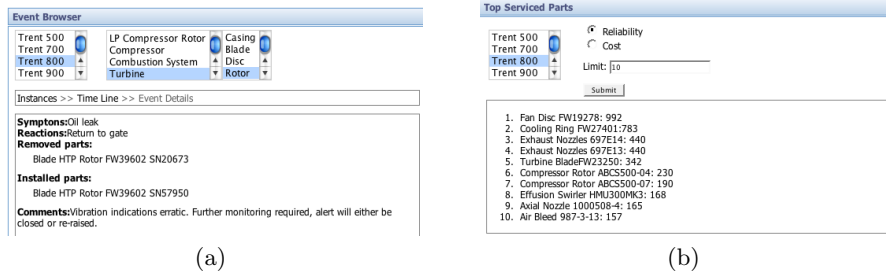


Fig. 3. (a) Maintenance event browser portlet, (b) Most serviced part portlet.

5.3 Web Services

Web services provide processing functionalities that can be accessed from anywhere on the network. The web services are decoupled from the portlets. In other words, multiple portlets can execute the same web service, and a single portlet can execute multiple web services. We have developed several web services that perform common requests on the maintenance documents. For example, obtaining a list of parts that are involved in the highest number of maintenance events, tracing the maintenance record of an engine or a part, or retrieving details of a maintenance event. We have also developed a graphing service that returns a histogram or a bar chart for a set of given data. We have also created a document index for the maintenance reports that are in Word format.

6 Discussion

In this paper, we have presented an intelligent document repository for an aero-engine manufacturer. However, the system architecture is generic and can be applied to organizations outside of aero-engineering. Most large organizations generate a substantial amount of documents everyday. This is regardless of the sector the organization operates in, be it healthcare or finance. Our proposed architecture is applicable to document repositories in any industry, where the number of documents are too large to peruse, and that users are interested in knowledge extracted from documents deposited.

Heterogeneous document sources are integrated by a shared vocabulary – an ontology. To answer some of the questions design engineers want to know about maintenance, we need to combine information from multiple sources. For example, to know which parts dominate the cost drivers in the engine, we need to combine reliability data from maintenance event reports with the engine taxonomy and costing information. These data are contained in heterogeneous data sources with different schemas within the company. Integration is made possible with the use of the ontology, which allows software agents to reason over the different resources.

Users access the document repository via a standard web browser. System components are hosted on distributed servers. As a result, the software can be deployed and updated centrally, without changing the configurations of thousands of desktop computers. Also, users can access the document repository without special software installed on their computers.

The two portlets developed demonstrate the concept of ‘summary reports’ within a document repository. These summary reports correlate information that are present in the underlying heterogeneous documents, but cannot be easily discovered through simple document browsing. For example, the event browser portlet allows operations engineers to follow through the sequence of events that leads to a part’s failure. The serviced parts frequency portlet allows design engineers to discover which parts are dominating maintenance statistics and thus require further attention when designing the next variant engine.

7 Conclusions

In this paper, we discussed the development of a knowledge-based document repository. The system allows users from across a company to access knowledge that they need to undertake their activities. In particular, we have developed a prototype for a large aero-engine manufacturer. A fundamental shift is occurring in the aerospace industry away from selling products to providing services. This shift in market focus means that new products must be designed to provide lower and more predictable life cycle costs. To achieve this, engineers must obtain knowledge gained from the entire life of an engine.

However, because of the large size and distributed nature of aerospace manufacturers, engineers often do not have access to front-line maintenance data. In addition, the large number of documents accrued during the life of an engine makes it impossible to examine all information thoroughly. Therefore, unlike existing document repository or digital libraries, our system is knowledge, and not information, based. Our proposed system searches and analyzes relevant maintenance records and design guidelines, and provides engineers with a ‘summary report’ to such information. It is hoped that the summary reports provided will help engineers in creating design documents that incorporate aftermarket issues into the design requirements.

Two key technologies are used in integrating the distributed and heterogeneous data sources — the Semantic Web and Service-Oriented Architecture.

The Semantic Web provides the framework allowing computer programs to interpret and reason over the heterogeneous document sources. The documents are integrated using an ontology, which captures the terms and concepts used in aerospace engineering. Service-Oriented Architecture allows knowledge extraction and analysis functionalities to be added to the system as modules called web services.

To test our proposed architecture, we have developed a working prototype. The prototype consists of a web portal, several web services and a RDF triplestore. Users access the document repository via a web portal, which is customized according to users' roles. The prototype has been demonstrated to Rolls-Royce engineers and has received positive reviews.

8 Acknowledgments

This research was undertaken as part of the IPAS project (DTI Grant TP/2/IC/6/I/10292). The authors would also like to thank the project partners for providing us with data and ontologies. Specifically, we would like to thank Derek Sleeman and David Fowler from Aberdeen AKT for the ontology, and Alymer Johnson and Santosh Jagtap from Cambridge EDC for the user requirement analysis.

References

1. Harrison, A.: Design for service – harmonising product design with a services strategy. In: Proceedings of GT2006, ASME Turbo Expo 2006: Power for Land, Sea and Air, Barcelona, Spain (2006)
2. Wills, G., Fowler, D., Sleeman, D., Crowder, R., Kampa, S., Carr, L., Knott, D.: Issues in moving to a semantic web for a large corporation. In: Proceedings of 5th International Conference on Practical Aspects of Knowledge Management (PAKM). Volume 3336 of Lecture Notes in Artificial Intelligence., Springer (2004) 378–388
3. Gao, Y., Zeid, I., Bardasz, T.: Characteristics of an effective design plan system to support reuse in case-based mechanical design. Knowledge-Based Systems Knowledge-Based Systems **10** (1998) 337–350
4. Khadilkar, D.V., Stauffer, L.A.: An experimental evaluation of design information reuse during conceptual design. Journal of Engineering Design **7** (1996) 331–339
5. Shadbolt, N., Milton, N.: From knowledge engineering to knowledge management. British Journal of Management **10** (1999) 309–322
6. Heath, I., Wills, G., Crowder, R., Hall, W., Ballantyne, J.: Towards a new authoring methodology for large-scale hypermedia applications. Multimedia Tools and Applications **12** (2000) 129–144
7. Gruber, T.R.: A translation approach to portable ontology specifications. Knowledge Acquisition **5** (1993) 199–220
8. Millard, D.E., Tao, F., Doody, K., Woukeu, A., Davis, H.C.: The knowledge life cycle for e-learning. International Journal of Continuing Engineering Education and Lifelong Learning **16** (2006) 110–121

9. Crowder, R., Bracewell, R., Hughes, G., Kerr, M., Knott, D., Moss, M., Clegg, C., Hall, W., Wallace, K., Waterson, P.: A future vision for the engineering design environment: A future sociotechnical scenario. In Folkesson, A., Gralen, K., Norell, M., Sellgren, U., eds.: Proceedings of 14th International Conference on Engineering Design, Stockholm (2003) 249–250
10. Wallace, K.M., Clegg, C., Keane, A.: Visions for engineering design: a multi-disciplinary perspective. In: Proceedings of 13th International Conference on Engineering Design, Glasgow, Scotland (2001) 107–114
11. Priebe, T., Pernul, G.: Towards integrative enterprise knowledge portals. In: CIKM '03: Proceedings of the twelfth international conference on Information and knowledge management, New York, NY, USA, ACM Press (2003) 216–223
12. Carr, L., Hall, W., Bechhofer, S., Goble, C.: Conceptual linking: Ontology-based open hypermedia. In: Proceedings of 10th International World Wide Web Conference (WWW), Hong Kong (2001) 334–342
13. Nnadi, N., Bieber, M.: Lightweight integration of documents and services. In: DocEng '04: Proceedings of the 2004 ACM symposium on Document engineering, New York, NY, USA, ACM Press (2004) 51–53
14. Carr, L., Miles-Board, T., Woukeu, A., Wills, G., Hall, W.: The case for explicit knowledge in documents. In: Proceedings of ACM Symposium on Document Engineering, Milwaukee, Wisconsin (2004) 90–98
15. Ciravegna, F., Wilks, Y.: Designing adaptive information extraction for the semantic web in amilcare. In Handschuh, S., Staab, S., eds.: Annotation for the Semantic Web, IOS Press (2003)
16. Manola, F., Miller, E.: RDF primer. Technical report, W3C Recommendation, <http://www.w3.org/TR/rdf-primer> (2004)
17. McGuinness, D.L., van Harmelen, F.: OWL web ontology language overview. Technical report, W3C Recommendation, <http://www.w3.org/TR/owl-features> (2004)
18. Jagtap, S., Johnson, A., Aurisicchio, M., Wallace, K.: Pilot empirical study: Interviews with product designers and service engineers. Technical Report 140 CUED/C-EDC/TR140- March 2006, Engineering Design Centre, University of Cambridge (2006)