Grasping Product Pragmatics: A Case with Internet on TV

Aylin Koca
Industrial Design
Business Process Design
Tel: +31 40 247 2357
a.koca@tue.nl

Mathias Funk
Electrical Engineering
Electronic Systems
Tel: +31 40 247 4741
m.funk@tue.nl

Evangelos Karapanos
Industrial Design
User-Centered Engineering
Tel: +31 40 247 4019
e.karapanos@tue.nl

Anne Rozinat
Technology Management
Information Systems
Tel: +31 40 247 5407
a.rozinat@tue.nl

Eindhoven University of Technology
5600 MB, Eindhoven, The Netherlands

Nico van der Gaarden
Philips Consumer Lifestyle
BU Video and Multimedia Applications
5616 LW, Eindhoven, The Netherlands
Tel: +31 40 273 2114
Nico.van.der.Gaarden@philips.com

ABSTRACT

With the emergence of highly interactive products in the domestic space, consumer electronics brands are facing an increasing challenge in predicting the way their products are being used and experienced. Unforeseen experiences relating to functional, emotional and social/contextual aspects of product use lead to a large and increasing share of field complaints that cannot be attributed to a violation of products' specifications. In a project called Soft Reliability, we are trying to develop a product evaluation ecology that enables the anticipation of product use by gathering behavioral and attitudinal data early in the product development process, through longitudinal field studies with working prototypes. This paper introduces a novel framework for behavioral and attitudinal data collection and analysis. The framework enables instrumentation that relies on the event-based experience sampling method, and as such deploys an analysis methodology that includes process mining techniques for the analysis of usage patterns, multivariate techniques for the analysis of longitudinal attitudinal data, and product quality analysis techniques for the analysis of combined information. We illustrate the value and applicability of this framework in practice, through the findings of an ongoing project concerning the conceptualization of an innovative Internet on TV product, which is being conducted in collaboration with Philips, a multi-national consumer electronics company.

Categories and Subject Descriptors

H5.2. User Interfaces: Evaluation/methodology.

General Terms

Management, Design, Reliability, Experimentation, Measurement, Human Factors, Standardization.

Keywords

soft reliability, user experience evaluation, product usage logs, subjective user feedback, product conceptualization

1. INTRODUCTION

The context of product development processes has actively and dramatically been changing ever since the advent of the digitization era. As a consequence, today's businesses are prone to being both empowered by, but also overpowered by, the accessibility of various technological and market opportunities. For instance, the merging of digital technologies allows for the development of very innovative, multifunctional and adaptive products that are capable of reaching wider markets in the global arena. However, especially the lean and agile operations, possibly among distributed business units of a company that develops such products, also yield a pitfall: There are huge uncertainties and little timely indicators about the *actual deployment* of these products in the field, where there are too many new and unknown parameters (e.g., various users with changing expectations, many interconnected devices and technologies). Typically, available indicators regarding product acceptance or deployment come in too late in the process and mostly in the form of increasing numbers of *unidentifiable* field complaints that cannot be acted upon. To avoid this pitfall, special focus should initially be centered on *product conceptualization* activities of especially new products. Proper investment in the early phases of product development is known to be a key differentiator between winners and losers in the increasingly competitive global market [8].

As important as it is, product conceptualization is a truly challenging phase within new product development (a.k.a. NPD). Alexander underlined this a few decades ago [2]:

> "In the case of a real design problem, even our conviction that there is such a thing as fit to be achieved is curiously flimsy and insubstantial. We are searching for some kind of harmony between two intangibles; a form which we have not yet designed, and a context we cannot properly describe."

This account holds even more in the case of designing innovative, multifunctional and adaptive products, in the previously mentioned industrial context. In that respect, it is crucial to shift away from conventional and rigid-in-nature product development processes that are more focused on (an already-established) "form", towards more explorative, flexible, iterative and hence dynamic-in-nature product development processes that are more focused on the "context"[1]. Such a shift entails getting closer to the tangible aspects of the product concept, (i) with prototypes of the intended form; (ii) by acquiring detailed understanding of intended and possible natural use contexts, via field studies; and (iii) by systematically and iteratively evaluating the goodness of the fit between the intended form and context. If this shift is not carried out timely and appropriately, there is a huge likelihood that in practice the concept of good fit (i.e. of the designed form within a context) can and will fail. Failures of good fit, or so-called *soft failures*, i.e. where a user has problems with a technically sound product, are indeed becoming increasingly common [7, 23, 4, 17, 18, 20].

In the context of a large multidisciplinary research project (i.e. Soft Reliability project, [28]), with actively collaborating industrial partners and with support from the Dutch government, we developed a novel adaptable framework for product usage data collection and analysis that addresses the aforementioned issues. While in [20] we outlined the general idea of our framework, in this paper we discuss its usage and report on our experiences when applying it in practice, specifically during product conceptualization activities. Our framework jointly explores and promotes the effective treatment of the following:

- **I. Evaluation mode:** Field studies with real users in natural use environments, instead of premeditated laboratory tests with a limited user group. While laboratory tests are bound to be partial in revealing the many use cases observed in the field; the reality is about actual consumers in whichever country and specific use context, experiencing the product over time.
- **II. Time factor:** Extended, longitudinal and connected studies, instead of short-spanned and isolated studies. This enables monitoring (i) how the affective user perceptions of the product and of the experience with it vary over time [14], and (ii) the different "phases of use" [5] of the product that a user is expected to go through, for ensuring the successful incorporation of the product in users' daily lives.
- **III. Combining behavioral & attitudinal data:** Automatically augmenting *objective* usage logs with complementary *subjective* user feedback, to discover use patterns and to better reason about user experiences, instead of relying only on one kind of data, which may be biased.
- **IV. Providing mechanisms for efficient knowledge transfer:** Accounting for the various interests of different information stakeholders involved in product development (e.g., product manager, quality engineer, interaction designer, marketing specialist), instead of only one, e.g., usability expert's. This involves combining process mining, quality, statistical, and visualization techniques in efficient tools that can easily be used by multiple experts, which in turn removes the gap between design and evaluative data. Such an approach is ultimately aimed to increase the impact of user research on actual design decisions, and also to foster expert collaboration on developing better designed products.

Our ongoing collaboration with Philips, concerning the conceptualization activities of an Internet on TV product, has already led to the discovery of key insights influencing the formulation of the value proposition of the product, and hence to reliably (i.e. based on concrete and complementary data) identifying its competitive edge, while identifying fallible use cases and calling attention to unexpected consumer behavior. In short, our framework provides grounds to devise good fit between the product and its use context.

The paper is organized as follows: Section 2 presents the basic motivation for our framework by introducing the idea of product pragmatics and its significance in the industrial context. Section 3 provides the infrastructure of our adaptable framework for product usage data collection and analysis. Section 4 outlines the industrial case to which we applied our framework. Section 5 concludes the paper.

2. PRODUCT PRAGMATICS

In Section 1, we described the changing context of NPD processes in today's competitive global market; identified the inherent problems embedded in these processes; hinted at ways to overcome such problems; and thereby emphasized the significance of the proposed work. In this section, we elaborate on the problem definition by utilizing long-existing theory in language studies to draw on similarities between linguistics and design.

Although both the study of languages and the study of design are relatively old, the notion of "design as product language" does not date back too long ago. This notion originally relates to semiotics (i.e. sign systems studies, which embody linguistics), a view according to which products and their design aspects are regarded as representations. More specifically, a product can be considered in its totality as a sign that states a value or benefit, while carrying several messages regarding the product's purpose, properties, functions, producer, and ultimately the profile of its owner [22]. According to Steffen [25], many symbolic functions, various complex cultural, social, technological, economic, and ecological meanings and representations might be traced in the design of a product. Consequently, while language can be seen as a symbolic instrument for communication between a message *sender* and a

[1] The focus on "context" is due to the industrial aim to get a competitive edge by creating value to users, in terms of getting the product to provide customized and adaptable services, and hence (i) either smoothly fit in the common practices of its users, or (ii) suggest (radical) changes to common practices in favor of more convenient and pleasurable experiences.

receiver, a product can be seen as a symbolic instrument for communication between a product's *designer* (or developer/manufacturer) and a *user*. In that respect, what is intended by the designer while designing a product, may not be the same as how it is interpreted by the user in a specific use context.

The parallelism between language studies and product design becomes clearer when considering the three branches that linguistics is studied within: Syntax, semantics, and pragmatics. In linguistics, syntax refers to the structure and form of language; semantics refers to its literal meaning; and pragmatics further refers to its implied and inferred meaning *during use*. When projected onto product design (Figure 1), we can derive that product syntax refers to the design materials, structure/form and technology used in the product; semantics refers to the symbolic literal meaning of the product in the world in terms of its (designer-) defined "value proposition"; and pragmatics refers to the specific "use experience" of a user with the product in a certain physical, psychological, social, cultural, etc. context.

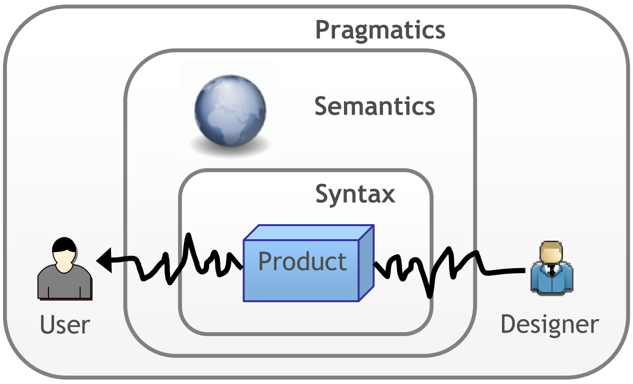

Figure 1. The subsumption relationships between syntax, semantics, and pragmatics.

Although the utility of applying the three linguistics branches to product design has more recently been recognized to some extent in the design literature [26, 22, 27, 15, 24], a clear mapping between the two disciplines has not yet been proposed. Specifically, the distinction between product semantics and product pragmatics is not explicit from previous work. This is also apparent from the work of Krippendorff and Butter [21], who pioneered the use of the term "product semantics" within the design community. They defined product semantics as "the study of symbolic qualities of man-made shapes, in the cognitive and social context of their use and application of the knowledge gained to object of industrial design." To highlight the distinctive role of the use experience, we distinguish between product semantics and product pragmatics as stated in the following: Semantics refers to the literal meanings of symbols in the world in general, disregarding an actual and specific use experience (i.e. no involvement or interaction of the sender/designer with the receiver/user); whereas pragmatics refers to the implied (by the sender/designer) and inferred (by the receiver/user) meanings of symbols *during an actual and specific use experience* (Figure 1).

A major challenge regarding early product conceptualization as defined in Section 1, is the recent necessity to start tracing Figure 1 from outside-in during NPD projects. This is in contrast with the conventional approach of tracing it inside-out (or in most cases inside-inside only), where product evaluation tests are conducted on an almost ready-for-release product. The obvious shortcomings of the conventional approach are twofold: *firstly*, not sparing time to experiment with the product on the semantic or the pragmatic level; and *secondly*, the infeasibility to do major revisions of the product regarding any failure of good fit identified at a late development phase. Since the highly uncertain context of innovative product development cannot afford risking the consequences of such shortcomings, more and more NPD processes start considering the outside-in scheme; i.e. starting first with grasping product pragmatics by testing a number of working prototypes in the field. To facilitate this shift away from the conventional approach, we developed our adaptable framework for product usage data collection and analysis, which can be tailored to the needs of various stakeholders in full-fledged product development teams.

3. PRODUCT USAGE ANALYSIS FRAMEWORK

As motivated heretofore, relevant user, usage and context information is of topmost importance in designing easy- and pleasurable-to-use products; especially, where innovative breakthroughs are involved that often yield unanticipated product use. Therefore, it is crucial to first probe with working prototypes in the field and get feedback about product pragmatics (i.e. how users are using and experiencing products). To incorporate early feedback into the product design, it needs to be captured appropriately and processed in a timely manner.

Our adaptable framework for product usage data collection and analysis (Figure 2) automates (longitudinal) behavioral and attitudinal data collection in order to promptly deliver this feedback to the different stakeholders of product information. This is achieved by the integrated collection of *objective* (i.e. relating to the actual usage behavior) and *subjective* (i.e. relating to perceptions and expectations of the user) data.

In the following subsections, the three consecutive main phases that comprise our framework (cf. white, bold print in Figure 2) are described. These are: (i) specification of the required information by various information stakeholders; (ii) behavioral and attitudinal data generation during use, and collection of these data via relevant filters; and finally (iii) information visualization and analysis via a combination of techniques, by the stakeholders. It should be noted that these phases are dynamically linked and hence can be executed iteratively, and with run-time reconfigurations. To better illustrate the way-of-working with this framework, each phase is traced with a running example, from the viewpoint of a product manager, as an information stakeholder.

3.1 Information Specification

Initially, the stakeholders of product related information involved in the product conceptualization phase, specify the kind of information they would like to obtain from the field (cf. "Information Specification" in Figure 2). There are various options to do this, based on the (*objective* versus *subjective*) nature of the required information.

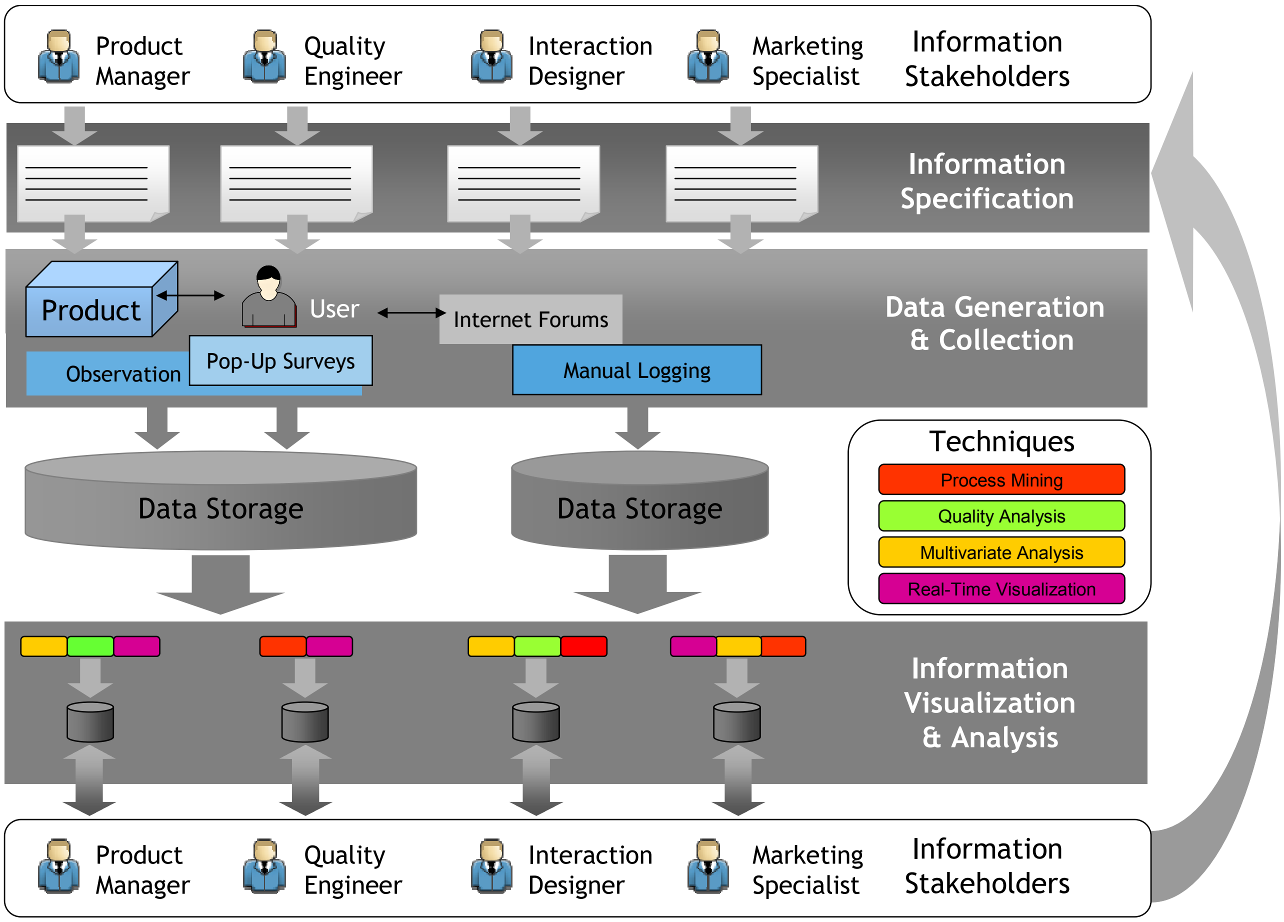

Figure 2. Adaptable Framework for Product Usage Data Collection and Analysis.

In the case of requiring *objective* usage information directly from the product, the product has to be equipped with an observation software module. In order to get relevant and rich data, this module has to be configured in a way that accurately reflects the stakeholder's information needs. For these purposes, we developed D'PUIS (Dynamic Product Usage Information System, [10, 20]), which is a set of tools and components connected via the Internet. D'PUIS also enables information stakeholders to conveniently define the specification of the desired information by utilizing a graphical editor. This is a domain-specific way to formalize what should be "observed" during product use, and how the collected data should be processed. Furthermore, this graphical editor can be used by various stakeholders, irrespective of the depth of their technical knowledge about the product architecture, or their technical programming skills. While defining the objective usage information specifications, also a priori semantic links between collected data items can be established by means of ontologies. Ontologies capture concepts and the relations between them. With use of relevant ontologies (e.g., product-specific ontologies about the product's functional features or structural parts), it is possible to model categories of information that can be used by the automatic observation module to process raw data. Consequently, the objective usage information specifications, optionally augmented with embedded ontologies for richer semantics, get automatically distributed to prototype products in the field.

In the case of requiring *subjective* usage information from users, event-based Experience Sampling [12] is used. Users' experiences are captured through *qualitative* and *quantitative* data gathering, which can be initiated (i) by the user, when s/he wants to manually report a question via, e.g., an Internet forum, or give feedback to the design team via, e.g., filling in a corresponding D'PUIS pop-up survey; or (ii) by D'PUIS pop-up surveys, which are triggered by an automatic detection of pre-defined interaction patterns; or (iii) again by D'PUIS pop-up surveys, which are automatically triggered at pre-defined time intervals (i.e. time-based Experience Sampling). Based on stakeholder requirements, *qualitative* data may provide information regarding the failures and the successes of the product-concept based on the match between product capabilities and user expectations (i.e. negative and positive disconfirmation of user expectations) [16], and further based on how the consecutive "phases of use" of the product are being experienced by the user [5]. The model we developed (a.k.a. designer-centered feedback analysis and classification model) and use for these purposes are described in [18, 19]. Additionally, *quantitative* data may provide information regarding users' momentary affect [6, 9], as well as about users' longer term perceptual and evaluative judgments [14]. For the evaluation of personal consumer electronics products, we specifically use the Attrakdiff2 questionnaire [11] that aims at assessing not only products' practical qualities (i.e. utility and ease-of-use), but also their hedonic qualities such as *stimulation* (i.e. the product’s ability to address the human need of stimulation, novelty and challenge) and *identification* (i.e. the product’s ability to address the need of expressing one’s self through the objects one owns).

D’PUIS pop-up surveys for *subjective* data gathering can be authored and sent to the prototype products in advance of the field study (i.e. during "Information Specification" in Figure 2), as well as during the study as a reaction to preliminary results (i.e. during "Information Visualization and Analysis" in Figure 2). We developed a simple markup language that enables the easy creation of such surveys, and that links the embedded ontology concepts with the surveys' textual input fields, multiple choice fields, and check box fields. As the ontologies are automatically being used for both *objective* and *subjective* data categorization, it is possible to obtain highly structured data that integrates a priori semantics in a holistic way.

Example. A product manager may be interested in a per-country overview of the visited local sports sites, and how the quality of the content provided is appreciated by users. Therefore, the product manager would decide on parameters like a list of Internet sports sites, and the frequency, time, duration of visits to these sites as a content quality measure. In order to complement this kind of *objective* data with users' *subjective* feedbacks, the product manager would arrange for D’PUIS pop-up surveys about content quality to find out how the users actually experience the browsing process (e.g., their awareness of what is possible or not with the product/services, motivation to explore, initial-use and sustained-use impressions), and how their attitudes towards this experience develop (i.e. their perceptions about products’ practical and hedonic qualities) over time. Furthermore, to do this objective-subjective data combination in an automatically integrated way, the product manager would ensure that both kinds of data are linked with corresponding semantic concepts from ontologies: E.g., a user input to the pop-up survey about a positive initial-use remark on the browsing process would need to be linked to the related concepts on the respective product-specific ontologies that triggered this positive remark, such as the "connected peripherals" or the "display resolution" concept from the product's structural parts ontology.

3.2 Data Generation and Collection

Based on the specified information requirements from different stakeholders, this phase is dedicated to generate data when the user starts interacting with the product, and to collect the relevant part of the generated data in the required format (cf. "Data Generation & Collection" in Figure 2). The collected data is then stored in a central data repository: Both the *objective* data captured via the observation modules within prototype products, and the *subjective* data captured via the user- or D’PUIS-initiated pop-up surveys, are stored in the same repository as shown in the left-middle part of Figure 2. In addition, our framework also supports incorporation of other (external) repositories used to store, e.g., manually registered *subjective* data from users who post questions or feedback on public/private Internet forums as shown in the center-middle part of Figure 2.

In the case of *objective* usage data, each prototype product in the field, already having received a specification of what to observe and how to process the collected raw data items, collects data accordingly and feeds that to a common repository, e.g., a web server. There, data are stored in a structured and machine-readable way (including their semantic links to concepts of relevant ontologies that they represent) for later analysis.

In the case of *subjective* usage data, the user-initiated or D’PUIS-initiated pop-up surveys as designated by the stakeholders' specifications, enable relevant data generation in the field, in a semantically structured way. Then those data are collected and fed to the common repository, where they are kept ready for (qualitative and quantitative) analysis. In addition, subjective data that are generated via other (external) means (e.g., via collection of feedback posted manually by users on public/private Internet forums) may be incorporated at this phase, for inclusion in the same qualitative analyses later on.

Example. During this data generation and collection phase, the users would be presented with two short D’PUIS pop-up surveys (as designated by the product manager's information requirements), when the observation module in the product detects that they left one of the specified sports sites. One survey would be targeted at capturing their impressions about the browsing process by identifying their likes, dislikes, problems and suggestions (e.g., expected but missing features), based on phases-of-use. The other survey would be targeted at measuring their momentary affect about the specific experience with the site just visited, by asking them to rate the product/application based on attributes such as confusing-clear, standard-creative, amateurish-professional, etc. Meanwhile, the objective usage data generated (e.g., the time and duration of the visit) would also be collected and logged as specified.

3.3 Information Visualization and Analysis

The stored data from the previous phase can now be visualized and analyzed in various ways by utilizing several techniques (cf. "Information Visualization & Analysis" and "Techniques" in Figure 2).

Objective usage information can be quickly reviewed using real-time visualization in the form of charts and data tables, but to be able to address advanced questions, our framework provides a connection to process mining techniques, as we describe in [20]. Real-time visualization helps to instantly grasp the structure and content of the incoming information, and thereby enables real-time monitoring to ensure that precisely the intended data is being collected in the field. Process mining tools like ProM [1] allow for richer analysis. They can be used to, for example, discover models of typical usage behavior, answer specific questions (e.g., "Do users who do *X*, typically also do *Y*?"), analyze the timing behavior and discover frequent patterns. Furthermore, the semantic links can be leveraged to include domain-knowledge in the analysis, such as to analyze the usage process on different levels of abstraction.

Subjective usage information can be visualized and analyzed using multivariate analysis, quality analysis and semantic process mining techniques. Users’ perceptions regarding the product can be analyzed with multivariate analysis techniques such as Factor Analysis (FA) or Multi-Dimensional Scaling (MDS). These techniques typically aim at modeling relations between stimuli (e.g., products), attributes (i.e. users’ perceptions such as professional–unprofessional) and overall judgments (e.g., preference), as we show in [13]. One may thus visualize how the respective product compares to competition, how different generations of the same product compare, or even how users' perceptions of the same product develop over time (i.e. by depicting perceptions of the product at different use phases as different instances). In addition, quality analysis can be conducted on the related qualitative survey results, by viewing the product's fault-failure[2] versus success distribution proportions over the categories of our designer-centered feedback classification model [18, 19]. Such analysis enables to systematically view the user-perceived strong and weak points of the current product concept, and as such enables (i) visualizing how those compare to other products', (ii) analyzing the relative importance of certain failures or successes, and hence prioritizing, and (iii) suggesting relevant actionable items for diverse information stakeholders to work on. Semantic process mining is yet a more sophisticated technique than process mining, which our framework supports, that is used to discover and visualize the graphical models of how users typically experience products based on their subjective judgments, which may be represented by domain-specific ontologies [3]. By aggregating the different levels of information provided by the survey results that are linked with ontologies capturing e.g., complaints, or praises, or use purposes, with the information provided by the usage log data that are linked with ontologies capturing e.g., usage statistics (with semantic links), it is possible to get a very rich view on the all-encompassing user experience.

Example. The product manager would review the initial incoming data by means of real-time visualization. This helps to remotely monitor the field study for its entire duration. If he then realizes new sports sites in the most visited sites chart, he would like to include them also in his information specification, while the test is still running. Using more detailed analysis, i.e. by means of process mining, he would for example visualize visit timing statistics. The embedded semantic links that distinguish between various categories of web sites, e.g., sports vs. shopping sites, would enable displaying the usage on different relative levels of abstraction: E.g., frequently visited sites ranking per topic, at-length visited sites ranking per country, day-time vs. night-time preferences per household, etc. When combined with results of pop-up surveys (also embedded with semantic links), he would identify if the reported failures or successes regarding a product use-phase can be verified by the usage behavior of that user: E.g., if the user reports that she finds the new product interface good looking but too complicated, and if this can be verified by the product logs, which reveal that she seldom used the interface. For a field study that spans a couple of weeks, the product manager can also visualize and analyze the effects of time on the product's incorporation into its users' daily lives, based on the results of periodic pop-up surveys.

4. INDUSTRIAL CASE

In summer 2007 an industrial collaboration between Eindhoven University of Technology (TU/e) and Philips, a multi-national consumer electronics company, was started. The plan was to do two consecutive studies centered on a conceptual prototype of a new Internet on TV product. For confidentiality reasons we will not go into the details of the business case, but focus on the studies and report on our experiences. While the first study had an explorative nature and is already finished, the second one focuses on details and is currently being conducted. The setting of both studies is deliberately not a dedicated testing lab, but the home environment of potential users.

The cooperation between the two organizations combined complementary assets. The Philips team provided a basis of representative key users around the world, supplied them with working prototypes, and offered dedicated facilitation whenever they requested. The TU/e team provided surveys and questionnaires for collecting relevant data, theoretical categorization schemes for structuring the data, and post-processing support over the collected data. Also, the TU/e team equipped all prototypes with special observation capabilities that tracked relevant user-system interactions and transmitted pre-processed information to a central server, as described in Sections 3.1 and 3.2.

Figure 3 depicts the timeline of the collaboration between TU/e and Philips. While Study 1 and Study 2 show the two consecutive studies on different versions of an Internet on TV prototype (current status on the timeline is marked with "now" in Figure 3), the right half of the figure shows some of the activities carried out in cooperation between the two organizations. These activities are referred to in Sections 4.1 and 4.2 in slightly more detail[3]. It should be noted that, our framework in its current form, as presented in Section 3, has been to some extent shaped by our experiences gained during Study 1.

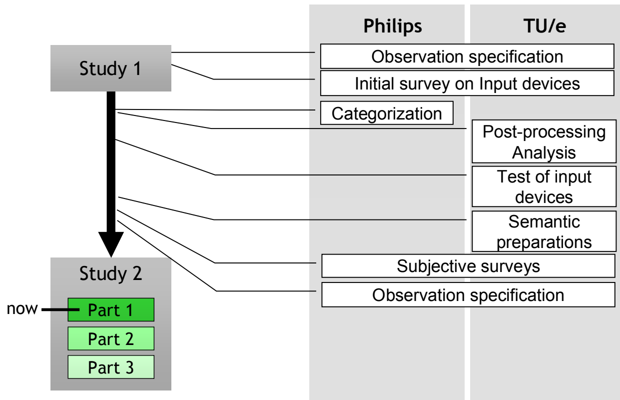

Figure 3. Industrial case timeline, showing collaboration between Philips and Eindhoven University of Technology

In the remainder of this section, we report on the objectives, the setup and the results of the first study. Furthermore, we describe how our experiences with the first study helped in shaping the structure of the second study, which is currently being performed.

4.1 Study 1

4.1.1 Objectives

The first exploratory study was conducted with the aim to test users' attitudes towards gaining convenient and high-quality access to the Internet from their TVs in the living room. In view

[2] N.B. The use of failures in this text refers only to *soft failures*, i.e. where a user has problems with a technically sound product.

[3] N.B. We are restricted to discuss the collaboration in greater detail, due to confidentiality reasons.

of that, several questions aroused about how users would react to such an application, would they welcome more active interaction in a usually relaxed environment, if so, would they "lean-forward" and actively pursue information or entertainment, or would they "lean-backward" and go for a rather passive experience, or would they do something in-between and change their mode of use often. The objective was not only to find out about the general acceptance and relevance of product features for its users, but also to explore the rather new domain of Internet on TV products: E.g., what kind of content would be searched for, played with, and enjoyed, even over a longer period of time? Furthermore, there were also unknowns about the peripherals provided with the prototype: What kind of pointing device would be more appropriate for such an application, and what type of textual input device would be necessary?

4.1.2 Setup

This study was carried out with 20 users from different countries around the world for 5 weeks. To provide *objective* usage data, the D'PUIS framework that we developed at TU/e was customized to the Philips case. This involved the insertion of an observation module into all 20 prototypes that were sent to the users' homes. This module provided objective usage information via an Internet connection and hence could be reconfigured remotely, whenever the required information specifications by Philips were updated (cf. mutual activity "Observation specification" in Figure 3).

As sources for *subjective* usage data, we utilized (i) a private forum on the Internet, which was moderated by Philips facilitators and used by all prototype users, and (ii) surveys sent to the users via the Internet forum (i.e. mutual activity "Initial survey on input devices" in Figure 3). While the first source collected rather direct, user-initiated opinions and questions, the second one collected focused answers to specific questions of the Philips product development team. Moreover, since the Internet forum was actively supported by facilitators, rich insights could be obtained regarding how users experienced setting up the device.

4.1.3 Results and Discussion

During this study over 5 weeks, roughly 800,000 (objective) data items were collected by all prototypes and the users submitted over 400 opinions and detailed remarks (in freetext form) to the Internet forum. The mass of disconnected and unstructured data called for extensive post-processing, in order to gain usable and comprehensive information. Therefore, objective data was analyzed by means of process mining, and subjective freetext data was manually categorized and filtered for product quality analysis purposes (cf. TU/e activity "Post-processing analysis" in Figure3).

The collection of objective information helped to answer many open questions, especially in the first exploratory study, e.g., *how* and *how often* people were using the device. This was useful information complementary to the more qualitative feedback that is typically acquired in user studies that take place in the early stages of product development.

The diversity of kinds of user interactions with the prototypes was a major problem in the analysis phase: The Philips team helped to manually break down 5000 categories, which resulted from the automated analysis, to approximately 160 categories (cf. Philips activity "Categorization" in Figure 3). Thus, the huge amount of initial data became easily manageable such that usable information could be derived from it.

As the prototypes were distributed to users in various countries, the results could be analyzed per country. It was observed that users from the same country used similar features and the results show that they were hindered by the same problems regarding infrastructure and services (e.g. content availability). This gave valuable insights for Study 2.

Runtime logging in the prototypes proved to be useful especially for validating the subjective feedback from the Internet forums. Subjective information regarding users' attitudes was triangulated with objective information regarding actual product use. For instance, the opinion from a user who did not use the device could be handled differently in terms of experimental invalidation or augmentation (e.g., justification of dislike, etc.), than the opinion from a user who extensively interacted with the device. This proved to be a direct indicator of the user feedback reliability.

Furthermore, subjective data analysis of the Internet forum data, based on the designer-centered feedback classification model we developed at TU/e, showed what each user's likes, dislikes, problems and suggestions were, and how they evolved over the 5 weeks of use, with respect to their effects on different phases-of-use of the product. It was also possible to derive detailed product-specific strengths and weaknesses, as experienced by users. This was useful input for further development actions.

4.2 Study 2

This second study comprises three complementary parts as depicted in Figure 3. The first part has already started, whereas the latter two are planned. The goal of all these follow-up studies is to further leverage the points that worked well in the first study, while benefiting from the lessons learned from it, and introducing more a priori structure (i.e. at the "Information Specification" phase, cf. Section 3.1) to fully apply our framework. Therefore, in preparing the second study, most effort has been spent on the development of methods. Especially the immense post-processing, for both objective and subjective data, that had to be performed after the first study hindered prompt feedback to the development team, which could otherwise have been quicker and even richer. Accordingly, the main improvements in our approach resulting from the first study have been (cf. TU/e activity "Semantic preparations" in Figure 3) (i) to establish a priori semantic links between data items in order to speed up general post-processing and enable a more detailed analysis, and (ii) to combine objective and subjective data analysis. For this purpose, it has been investigated how several ontologies capturing different aspects of the prototype, objective and subjective data collection can be linked together and automatically be attached to data items that are captured throughout the study. It has hence also been part of the improvement activities to embed surveys with relevant (e.g., product failure, success, based on phases-of-use) ontologies, to eliminate manual post-processing of the subjective data (cf. mutual activity "Subjective surveys" in Figure 3). Consequently, during the second "Observation specification" mutual activity in Figure 3, all these improvements have been made use of.

The first part, Part 1 of Study 2, is designed to be more in-depth and focused to answer certain questions for the business case. It is conducted in one country only, with 20 users, for 2 weeks. The objective log data in this case is planned to be utilized for verifying the subjective opinions of users as recorded during detailed interviews, and vice versa.

The second part, Part 2 of Study 2, is designed to be conducted in a high-contrast country, geographically and culturally distant from the previous (i.e. Part 1).

The third part, Part 3 of Study 2, is designed to be conducted in the Netherlands, over an extended period of time, specifically to measure how users' usage behavior and attitude evolve over a longer period of time.

5. CONCLUSIONS

Successful development of innovative products is somewhat paradoxical: Rich insights are needed into how new technology transforms the design space early in the development of a product. However, the impact of new technology cannot be foreseen until it is fully incorporated in the specific ecology of values and activities. We argue for a radical approach to user centered design, that of intervening into specific cultures with technology probes. Only by observing users' changing habits and attitudes in response to radical interventions can one foresee the impact of new technology in socio-cultural settings.

In this paper, we formulated the design space in terms of product syntax, semantics, and pragmatics. We specifically highlighted the importance of (i) grasping "product pragmatics", which is about understanding the use experience of a user with a product in a certain physical, psychological, social, cultural, etc. context, and (ii) doing it early during product conceptualization. To address these two points, and as a means of resolving the paradoxical design problem, we proposed a novel adaptable framework for product usage data collection and analysis. Using this framework, we illustrated a product evaluation ecology that enables the anticipation of product use by gathering behavioral and attitudinal data early in the product development process, through longitudinal field studies with working prototypes. We discussed how the framework enables instrumentation by relying on the event-based experience sampling method, and as such, how it deploys an analysis methodology that includes process mining techniques for the analysis of usage patterns, multivariate techniques for the analysis of longitudinal attitudinal data, and product quality analysis techniques for the analysis of combined information. Finally, we exemplified the applicability of our approach by reporting on our ongoing collaboration with Philips, about the conceptualization activities of an Internet on TV product. Our efforts focus on five main principles:

I. **Design interventions** with fully-working prototypes provide rich insight for the development of innovative and successful concepts.

II. **In-situ observations with field studies** yield reliable information about how various users are using and experiencing products.

III. **Longitudinal data analysis** reveals if a particular usage behavior is the result of a "wow factor", or if it sustains over time and use. Furthermore, longitudinal data also reveals how users appropriate new technology and use it in their own ways, which shows new opportunities for innovation (i.e. transforming the design space).

IV. **Objective and subjective data**, when analyzed jointly, give richer and more reliable insights into how various users are using and experiencing products. This opens up new possibilities for combining the knowledge of various domain experts.

V. **Bridging evaluation to design** enables designers to have an active role in evaluation (i.e. increases the impact of user research on actual design decisions). Moreover, accounting for the various interests of different information stakeholders involved in product development during (field) studies is ultimately aimed to foster expert collaboration on developing better designed products.

We envision that our approach can help academics as well as practitioners to gain a competitive edge in establishing successful innovations in value-rich contexts.

6. ACKNOWLEDGMENTS

This work is being carried out as part of the "Managing Soft Reliability in Strongly Innovative Product Creation Processes" project [28], sponsored by the Dutch Ministry of Economic Affairs under the IOP-IPCR program. The authors would like to thank Serge Stevenart, Thomas P. John, Nuno Ferreira, Ronald Geerlings, Michel Jacobi, and Frank Berben from Philips, for their kind support. Furthermore, the authors also gratefully acknowledge Prof. Aarnout C. Brombacher, the project leader.

7. REFERENCES

[1] Aalst, W.M.P.v.d., Dongen, B.F.v., Günther, C.W., Mans, R.S., Alves de Medeiros, A.K., Rozinat, A., Rubin, V., Song, M., Verbeek, H.M.V., Weijters, A.J.M.M. 2007. ProM 4.0: Comprehensive Support for Real Process Analysis. In Proceedings of the 28th International Conference on Applications and Theory of Petri Nets and Other Models of Concurrency (Siedlce, Poland, June 25-29, 2007). ICATPN '07. Springer, 484-494.

[2] Alexander, C. 1964. Notes on the Synthesis of Form. Harvard University Press, Cambridge, Mass., USA.

[3] Alves de Medeiros, A.K., Pedrinaci, C., Aalst, W.M.P.v.d., Domingue, J., Song, M., Rozinat, A., Norton, B., and Cabral, L. 2007. An Outlook on Semantic Business Process Mining and Monitoring. OTM Workshops (Vilamoura, Portugal, November 25-30, 2007). Springer, 1244-1255.

[4] Bly, S., Schilit, B., McDonald, D.W., Rosario, B., and Saint-Hilarie, Y. 2006. Broken Expectations in the Digital Home. Extended Abstracts on Human Factors in Computing Systems (Quebec, Canada, April 22-27, 2006). CHI '06. ACM Press, 568-573.

[5] Bouwmeester, den K., and Bosma, E. 2006. Phases of Use: A Means to Identify Factors that Influence Product Utilization. Extended Abstracts on Human Factors in Computing Systems (Quebec, Canada, April 22-27, 2006). CHI '06. ACM Press, 117-122.

[6] Bradley, M. M., and Lang, P. J. 1994. Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential.

Journal of Behavior Therapy and Experimental Psychiatry, 25, 49-59.

[7] Brombacher, A.C., Sander, P.C., Sonnemans, P.J.M., and Rouvroye, J.L. 2005. Managing Product Reliability in Business Processes 'Under Pressure'. Reliability Engineering & System Safety. Elsevier, 88 (2), 137-146.

[8] Cooper, R. G. 2001. Winning at New Products: Accelerating the Process from Idea to Launch. 3rd edition. Perseus Publishing.

[9] Desmet, P. M. A., A. 2003. Multilayered Model of Product Emotions, The Design Journal, 6, 4-13.

[10] Funk, M., Putten, P. v.d., and Corporaal, H. 2008. Specification for User Modeling with Self-Observing Systems. In Proceedings of the 1st International Conference on Advances in Computer-Human Interaction (Sainte Luce, Martinique, February 10-15, 2008). ACHI '08. IEEE Press, 243-248.

[11] Hassenzahl, M. 2004. The Interplay of Beauty, Goodness, and Usability in Interactive Products. Human-Computer Interaction. Lawrence Erlbaum Associates, 19 (4), 319-349.

[12] Hektner, J. M., Schmidt, J. A., and Csikszentmihalyi, M. 2007. Experience Sampling Method: Measuring the Quality of Everyday Life. Sage Publications Inc.

[13] Karapanos, E., and Martens, J.-B. 2007. Characterizing the Diversity in Users' Perceptions. In Proceedings of the 11th IFIP TC13 International Conference on Human-Computer Interaction (Rio de Janeiro, Brazil, September 10-14, 2007). INTERACT '07. Springer, 515-518.

[14] Karapanos, E., Hassenzahl, M., and Martens, J.-B. 2008. User Experience over Time. Extended Abstracts on Human Factors in Computing Systems (Florence, Italy, April 5-10, 2008). CHI '08. ACM Press, 3561-3566.

[15] Karjalainen, T.M. 2004. Semantic Transformation in Design. Communicating Strategic Brand Identity from Product Design References. Ilmari Publications. University of Art and Design Helsinki.

[16] Koca, A., Lu, Y., Brombacher, A.C., and Hartmann, J.H. 2006. Towards Establishing Foundations for New Classes of Reliability Problems Concerning Strongly Innovative Products. In Proceedings of the 3rd International Conference on Management of Innovation and Technology (Singapore, June 21-23, 2006). ICMIT '06. IEEE Press, 581-584.

[17] Koca, A., Schouwenaar, A.J.M., and Brombacher, A.C. 2007. Field-Feedback in Innovative Product Development: A Comparison of Two Industrial Approaches. In Proceedings of the 14th International Product Development Management Conference (Porto, Portugal, June 10-12, 2007). IPDM '07. EIASM, 657-668.

[18] Koca, A., and Brombacher, A.C. 2008a. Extracting "Broken Expectations" from Call Center Records: *Why* and *How*. Extended Abstracts on Human Factors in Computing Systems (Florence, Italy, April 5-10, 2008). CHI '08. ACM Press, 2985-2990.

[19] Koca, A., and Brombacher, A.C. 2008b. User-Centered Analysis of Feedback Operations for Quality Improvement in New Product Development. In Proceedings of the 15th International Annual EurOMA Conference (Groningen, The Netherlands, June 15-18, 2008). EurOMA '08. EIASM.

[20] Koca, A., Funk, M., Karapanos, E., Rozinat, A., Aalst, W.M.P.v.d, Corporaal, H., Martens, J.B.O.S., Putten, P.H.A.v.d., Weijters, A.J.M.M., and Brombacher, A.C. Soft Reliability: An Interdisciplinary Approach with a User-System Focus. Quality and Reliability Engineering International Journal. Wiley, DOI: 10.1002/qre.937.

[21] Krippendorff, K., and Butter, R. 1984. Product Semantics: Exploring the Symbolic Qualities of Form. Innovations. MIT Press, 3 (2), 4-9.

[22] Monö, R. 1997. Design for Product Understanding. Stockholm: Liber AB.

[23] Ouden, P.H.d., Lu, Y., Sonnemans, P.J.M., and Brombacher, A.C. 2006. Quality and Reliability Problems from a Consumer's Perspective: an Increasing Problem Overlooked by Businesses? Quality and Reliability Engineering International. Wiley, 22(7), 821-838.

[24] Souza, de C. 2005. The Semiotic Engineering of Human-Computer Interaction. MIT Press.

[25] Steffen, D. 2000. Design als Produktsprache: Der "Offenbacher Ansatz" in Theorie und Praxis, Verlag Form GmbH, Frankfurt am Main.

[26] Vihma, S. 1995. Products as Representations. A Semiotic and Aesthetic Study of Design Products. Helsinki: University of Art and Design Helsinki UIAH.

[27] Warell, A. 2001. Design Syntactics: A Functional Approach to Visual Product Form. Theory, Models, and Methods. Doctoral Thesis. Chalmers University of Technology, Göteborg, Sweden.

[28] Soft Reliability. URL: http://softreliability.org/