From Hands-on Sessions to User Insights on Designing an Interactive System for Data Science

Ana Sofia Almeida
Centre for Informatics and Systems (CISUC),
University of Coimbra, Coimbra, Portugal *asa@dei.uc.pt*

Licinio Roque
Department of Informatics Engineering,
University of Coimbra, Coimbra, Portugal *lir@dei.uc.pt*

Paulo Rupino da Cunha
Department of Informatics Engineering,
University of Coimbra, Coimbra, Portugal *rupino@dei.uc.pt*

Abstract

Our design research goal is to improve the user experience and effectiveness of an integrated IT solution for supporting the creative and collaborative Data Science (DS) life-cycle process. The work is being done as a Design Science Research (DSR) project, in real-life context. Within a fast-pace development environment, with scarce access to end-users, we combined hands-on sessions and semi-structured user interviews into a fast-forward design insights technique ([aka *insightz*]) to capture: i) people interests and expectations about the tool (leading to design improvements) and ii) stakeholders' insights about the DS process itself (leading to process and business innovation). We propose these *insightz* workshops and the user research approach as a design technique to define and to communicate design principles and guidelines between different stakeholders, namely, UI/UX and engineering teams.

Keywords: Human-Computer Interaction, User Experience Design, Data Science, Design Science Research, Information Systems Design.

1. Introduction and Motivation

Companies increasingly need to deal with large scale data being created by information systems, hardly copping with the expanding volume, velocity and variety of it, albeit the speedy evolution of technologies for data storage, analysis and visualization [10, 17, 19].

Powerful computational capabilities make Machine Learning (ML) and Artificial Intelligence (AI) available and effectively viable to be used in a multitude of domain sectors. Data Science (DS) is a fruitful emerging interdisciplinary research area with growing technical, economic, social and ethical impact in multiple domain and use scenarios [6, 17, 28]. DS is becoming the new needed literacy for a broader range of professionals [8]. However, existing tools and technologies require a specific set of skills and competences, not always present in IT teams or widespread over other professional stakeholders and domain specialists [4, 8, 19]. The shortage of data scientists, "the sexiest job of XXI Century" [7] and "the hottest profession" [20] along with the increasing difficulties to deal with diverse (big) data quality urges for integrated IT solutions to support the full complex DS process.

In this overall context, we teamed up with a tech-based company and undertook a design research project to understand i) the nature of the full DS pipeline and ii) how to design a digital tool/service to support it; capturing, discovering and inventing user experiences.

This paper makes two contributions: i) an overview of a design research project in IS, with data scientists as end-users ii) a pragmatic user research technique, named *insightz*, to discover improvement opportunities for innovation design (of both product and process).

Next section outlines our design research approach, as a *research trough design* effort, highlighting the user research sessions flow and the iterative evaluation of the prototype(s) and process. In Section 3, we detail the protocol for the hands-on user sessions, used to test the prototype(s); to capture user needs and improvement opportunities and as *what-if* workshops to discover innovative ways to enhance product, process and interaction design. Section 4 discusses some research shortcomings and found constrains in managing the integration of the user research results into the prototype development and deployment, within an agile development pipeline. Last section addresses conclusions and future work.

2. Design Research Approach

Our partner is a tech-based company using AI and ML to fight fraud and manage risks, mainly for Banking, Insurance and e-Commerce Clients. The real-life problem was to improve the user experience and effectiveness of an in-house built pilot aimed to help company data scientists, software engineers and clients collaborate on DS activities. On this project, we followed a Design Science Research (DSR) approach to create new insights through the design of an innovative artefact [15]. We incorporate the guidelines and principles of DSR [11, 15, 26, 27], the construction of a digital prototype and the evaluation of its effectiveness for success criteria, as defined by company stakeholders and project goals.

Our design research approach aims to contribute using *research through design*, as proposed by Frayling [9], while emphasizing the "wicked problem" [22] nature of designing for DS, particularly, with a user-centric and a socio-technical perspective over the information systems design and development. Our research approach mainly included i) a literature review to situate our field work [2]; ii) a series of user research sessions, scripted and adapted to address our particular problem; iii) an iterative evaluation of artifacts designed and benchmarking criteria and iv) a communication strategy among teams and stakeholders.

In Table 1, we outline the research design approach followed, adapted from [21], positioning our research as "*Improvement*", according to the DSR Knowledge Contribution Framework [11]. We started by conducting a vast literature review focusing on how to design an interactive system in the areas of Information Systems (IS) and Human-Computer Interaction (HCI) [2]. We found potential in using hybrid design methods and shared research and practice perspectives, while addressing our specific project problem, to frame the integration of diverse subject matters and techniques, from both areas, into a novel *invention*.

In this effort of design research, we kept in mind the "four lenses", proposed by [29, 30], as criteria to a high quality contribution, retaining and trying to report enough detail about the design process in order to be somewhat reused in similar contexts. Moreover, the *relevance* of the design attempt comes from working on a real-world problem, with an underlying context of design and development in a fast-pace environment, with scarce access to end-users.

2.1. Data Science Lifecycle Pipeline

Our design work was conducted over the design and development of a DS prototype, addressing a 'end-to-end DS lifecycle process' (DSL), as presented in Figure 1.

The proposed DSL model is an iterative sequence of six stages, where the individual steps towards an end-to-end process are organized and visualized, so to communicate the DS process and workflow among company teams and clients. Some stages are more sequential (e.g. *2. Data Cleaning & Enrichment* is usually followed by *3. Modelling & Training*; *6. Feedback & Model Tuning* recurrently urges for changes in *3. Modelling* and further *4. Experimentation & Evaluation*, prior to a new *5. Deployment*). This DSL model supports the work of internal collaborators and transparently accommodates the migration of the outcome of the phases *3. Modelling & Training* and *4. Experimentation & Evaluation* into runtime (the production environment), represented by the phase *5. Deployment* and back again to the design phases of the process, directly addressing the core competitive advantage of our partner company.

From each of the phases, the professional can back and forward into another phase, iteratively. This means understanding the data which Clients can provide to the company;

cleaning that data and validating that it is correct, usually resorting to a mix of shell scripts and programming languages. Then, data scientists have to iteratively compute features, enrich them and tune several parameters in order to improve the performance (or quality) of the ML model. After trained and evaluated for best quality on the indicators accorded, the ML model is integrated into a Plan, along with business rules and a Workflow of services, to be deployed into the Client environment (typically, by engineers).

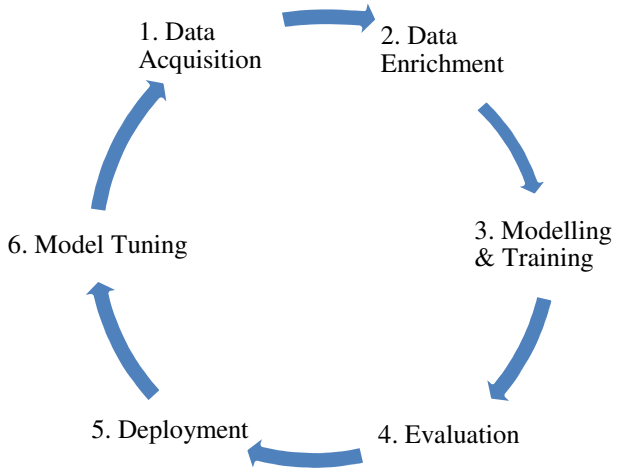


Figure 1. Data Science Lifecycle (DSL): a generic workflow of activities (as presented in [2]).

For each successful DS project, data scientists and engineers must work together with each other and with the Client's domain specialists. Each step of DSL has a corresponding digital support on the DS prototype(s) interface and has evolved from version to version, accommodating the feedback from the user research sessions. The corresponding design was iteratively evaluated in order to allow for job completion and the re-iteration of each phase. Each design aspect was considered from the particular user perspective and user profile, so that the overall collaborative and creative workflow would be quickly learned and used, by data scientists and engineers, be it experienced or newcomers to development.

The validated final prototype (V3) has evolved into a successful product. The model training, usually taking up two months to complete, is now done in a matter of weeks. The deployment phase, that could take up to one year to finish, is now made through a one-click system, batching the effort to put a new model in production, making this a truly competitive advantage, particularly in the finance sector, with millions of transactions to be scored in nanoseconds and prone to rapid change of fraud models.

2.2. End-User Sessions and Observation

While designing a DS service/tool for the DSL pipeline, we are aware of the diversity of personas and the need to understand the end-users, to get to know them, their roles, needs and expectations. What is learned from the users can be precious to understand the main needed improvements and new features, for strategic thinking and decision making. User priorities and how users think about the tool may differ from what we, researchers, business stakeholders and development teams think. Further, as [1], we too understand usability to be a combination of factors, "not a single, one-dimensional property of a product, system, or user interface". To plan and conduct a set of end-user sessions, with more than one interest in mind, we revisited HCI techniques and curricula in [16]. We, henceforth, evaluated the DS prototype(s) by conducting a series of user sessions, to demonstrate its effectiveness towards the main goals of the project and as opportunities of process innovation.

Several individual user sessions, were planned and conducted, in a jointly effort from our team and the company UI/UX design team. For selecting the participants, this team balanced the time constrains of collaborators (working for real Clients, using the DS prototype) and the

need to have a wide range of user profiles (e.g. experienced and novices) and a fit sample of company collaborators in age, genre and technical background diversity.

Table 1. DSR approach for defining an HCI project to design and improve a prototype for Data Science (adapted from Peffers [21]).

Activity	Method	Outcome
Identify the Problem & Motivation	Team up with a global leading company, with a real-life problem; Systematic Literature Review (to map Data Science work and Design methods / development approaches to apply); Initial sessions with Development Team and Expert DS/Stakeholder.	Analysis of key problems related to DS in practice (in the context); Characterization of the context (complex problem; fast-pace changes; sparse access to end-users; agile development in place; researchers with no control over design decisions & development). → how to impact design improvement?
Define Objectives of Solution	Literature Review on relevant fields, with a twofold perspective over the IT-Artefact design: as a tool and as a service; Observe Data Scientists' work; Workshop with company stakeholders.	Definition of knowledge base (IS, HCI and UX and Service Design); Definition of requirements and software architecture. → how to capture and timely incorporate user insights into development?
Design & Development	Prototyping to test the applicability of insights and usability: 1. To digitally support the DS workflow (from model creation to model deployment, in real time production environments); 2. For DS experts and non-IT specialists (wide range of non-data professionals as targeted end users/consumers)	Initial prototype for Hands-on Sessions; Validated IT-artefact (V1) in initial controlled settings (for supervised ML); Initial set of the design insights and guidelines for the DS lifecycle / pipeline.
Demonstration	Demonstrate the usefulness of prototype for: 1. Supporting the full Data Science pipeline 2. Fraud fighting (domain dependent goal); 3. Usage in other domains. 4. Business/Domain Specialists (non-DS experts)	Validated IT-Artefact (V2) in identified real use case scenarios: fraud detection and general end-to-end Data Science (from data loading and cleaning, ML modelling and evaluation, to deployment into production)
Evaluation	Initial usability sessions organized as semi-structured user interviews, targeting mixed user profiles (e.g. junior and expert users; data scientists and delivery teams), over a hands-on scenario usage (controlled settings); Second Sessions (follow up) to assess the prototype in continuous usage, as semi-structured interviews; SUS questionnaires to assess attractiveness of the system	Consolidated use of the prototype by company collaborators: teams use the system on (at least) four real-life projects, using, testing and evolving the design, involving a total of ten company collaborators (six data scientists and four engineers) in twelve sessions; Field tested artefact, ready to systematic use (V3, to productizing)
Communication	Presentation of internal research Design guidelines and list of issues (JIRA) Use-cases shared among stakeholders Academic conference and journal contribution	Structured Deck of Design Insights for DS, for evaluation (in progress); Peer-reviewed publications Factsheets on key features

2.3. Iterative Evaluation and Prototype Improvement

The evaluation of the resulting artefact is central to a DSR project, to proof with rigor its relevance for practice. Albeit the current body of knowledge and the guidelines available, we struggled in practice to select and justify the DS system design and which evaluation strategies and techniques to use [15, 25, 27]. In our case, this was partially due to the real-life context, of a fast-pace growing company, implementing an agile development process and with scarce access to end-users (namely, client personnel and over busy collaborators). Within this context, we planned and defined a strategy to capture how expectations, appropriation and some emotional attractiveness evolved as our system had more field usage. Along the project, user sessions feedback was iteratively incorporated to improve the prototype, over two iterations (from V1 to V2 and to V3).

Some issues are still waiting for the opportunity to enter production, due to time and resources constrains. Overall, the project's "usability thread", with its flow of user insight sessions, found its place in the iterative environment, with frequent releases, following the agile methodologies (SCRUM) in use in the company.

2.4. Capturing User Feedback and Communication Strategy

After each user session (workshop), we and the company's UI/UX team would come together to map insights and compare notes, identifying recurrent themes, later transcribing and listing all the issues found to a Google Sheet, for data analysis and stakeholders sharing. Later, once selected and prioritized, a proper issue was created on the tool internally used to manage the software development (JIRA). A shared list of issues and of use-cases was maintained as a communication strategy among the several teams and relevant actors (e.g. research team, business stakeholders, front-end designers and programmers). Also, the initial user sessions were tracked by follow-up user sessions, planned to be conducted after three to four months afterwards. For these second (and third) sessions, the script was basically the same but a new version of the DS prototype was used and tested. Plus, the experience users had with the DS tool changed and improved, since they had to continually use the system after the previous session, in their daily job activities in the DS projects for real Clients.

This continuous flow of asking and consulting with the users allowed to capture how needs, pains and expectations also change over time. For instance, from session to session, users would report less on the lack of functionalities or issues not solved and more on the pain of repetitive tasks, the explosion of alternative models to evaluate and to compare and/or the time needed to accomplish a rather complex goal, not remembering that a few months earlier they would use the same amount of time to accomplish just one of the sub-tasks, in the overall complex workflow. By using the system and, periodically, called upon to test and think about the process and what they were doing, the activities were changing on the user mental model, and tasks were relegated to sub-tasks (or even, to automated – unconscious – tasks). Attention shifted from the technical operations with the tool to a higher level of collaborative and creative work and activities to be done, involving users from the different teams.

3. Hands-on Sessions: Towards User Insights

3.1. Project Settlement

First, we had initial meetings with the Development team leader, that explained the main goals and overall objectives, providing an historical perspective over the initial problem settings and how the company decided to pursue the design and development of an integrated tool for the DSL process. When our project started, an in-house built initial exercise already existed. It was designed and developed by a member of the programming team with a twofold profile: he was a software engineer that worked for some time as a data scientist for the company. In order to deal with his own 'pains and problems', an initiative was proposed and further nurtured by the company, evolving into a full fleshed research project, with results aiming the development of a new product/service. This important stakeholder was extremely committed to find and understand the user's 'needs and pains' and in position to map them

into the software development management system into the production pipeline. This is commonly the case in tech-based companies, were new products and services often develop from promptly building a solution for a quotidian handled problem.

The original DS pilot was designed as a web-based application to integrate the DSL activities. It was centered on the data scientist perspective of the work, with functionalities targeting essentially the phases *2. Data Cleaning & Enrichment* and *3. Data Modelling & Training*. One of the obvious issues was the lack of support for the Deployment teams, working in the Clients premises, a role unexperienced by the pilot designer/developer.

The team suspected that user experience could be improved and sought help to capture insights on the user real interests and expectations to innovate the process and product.

3.2. Establishing a User Session's Protocol

We started by discussing with the Development and UI/UX teams, establishing the purpose, goals and practical objectives for the flow of user sessions. Since the participants would be collaborators, currently working on real-life projects, there were time and resource constrains to be approved by Project Leaders. Due to constrains relating to users' convenience and respective Client projects, we searched for a lean process to conduct the DS prototype evaluation. The overall agreed purpose was to meet with each of the users for one hour and retain most of the information, knowledge and insights about the artefact and the DS process itself. A UI/UX person was designated to accompany our team (external) and attend each session, functioning as the privileged liaison and as the "*user experience champion*" inside the company, for the duration of the project.

The sessions were to be recorded (audio and screen video), for further content analysis, targeting additional design insights for DS and a way to operationalize the user design proposals into system requirements or research purposes. Plus, it was decided that the findings were to be reported to the development team and to the business stakeholders with the procedures and tools in use in the company.

We meet with the UI/UX responsible to plan the "*usability thread*", a new thread of tasks to be included in the agile methodology in use. Figure 2 lists the pragmatic script defined to be followed with each participant and in each session.

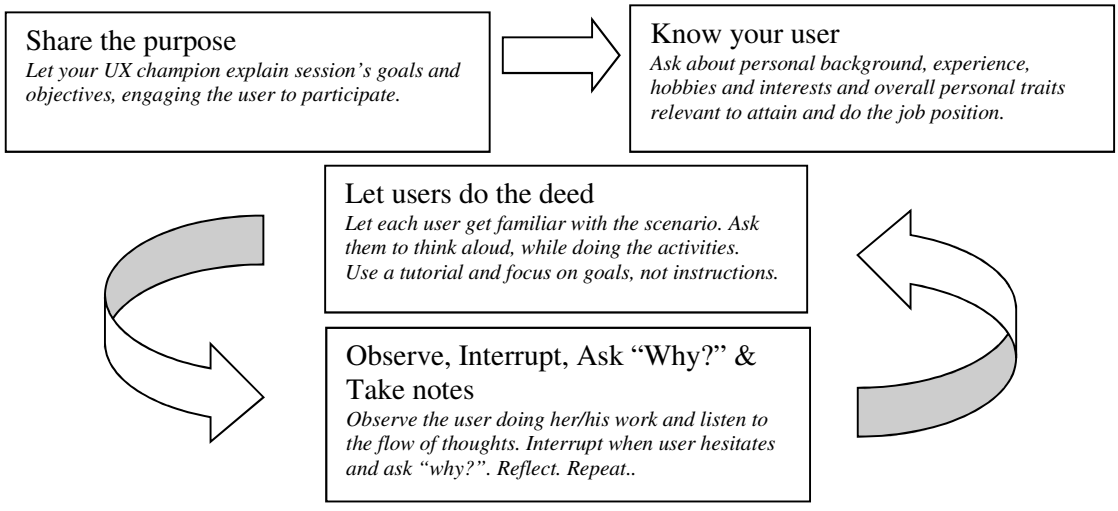


Figure 2. A generic protocol, scripting the hands-on sessions, named '*insightz* workshop'.

Each session was scripted as an experimentation session (over the use of a prototype) to function as a mix of lighter usability testing and user research session. With scarce access to end-users, we combined hands-on sessions and semi-structured user interviews into a fast-forward design technique ([aka *insightz*]) in order to capture: i) user interests and expectations about the tool (leading to design improvements) and ii) stakeholders' insights about the DS

process itself (leading to process and business innovation).

In the user's interviews, these main goals were kept in mind, while addressing three investigation areas:

Q1: Can the system be used for the intended job? The user understands concepts and how to use the system? Can user achieve the tasks? Is there a difference between experienced and new users?

The DS prototype(s) is a software artefact, used in a context to perform specific tasks. Users were asked to follow a step-by-step tutorial, with tasks related to the actual job activities, in a use-case controlled set. Due to NDA concerns, we provided a dataset of related data but proposing a non-familiar scenario. This allowed for a controlled user test-bed, challenging each user to stop and ponder at the 'new' problem and how to address it, while using the system. An effective future usage can only arise if the system fulfils the requirements from practice. Our prototype was specified with such requirements in mind and users were asked about needs, functionality and expectations for the DS system.

Q2: To what extent users make errors while using the software, how serious the errors are, and to what extent users recover from the errors? Does the user remember for new uses?

The sessions were also used to systematically test the technical functionality and the user interface controls, due to its relevance to the perceived usability of the DS system. Users were asked to follow a small tutorial for the DS process, using a controlled dataset, which allowed to compare performance results and elicit areas prone to errors. Counting the time for each task was not relevant, but we counted the number of errors and problems revealed by each user. If one problem or issue was "found" or reported by more than one user, its priority increased. That provided a pragmatic way to serialize the long list of founded issues (bugs, improvements and new features), select a smaller set for prioritization by stakeholders and eventually decided to enter production.

Q3: To what extent the user likes using the system? Does the user identify and/or report desirable changes? To the system? To the way of doing? What real-life use-case scenarios do users report?

The DS system has to support a complex process and we need to understand human and contextual factors involved in the collaborative work, such as learning and memory processes or information processing. The overall subjective satisfaction of users while using the prototype was also captured with post-session SUS questionnaires [3, 5, 23], that assess its overall attractiveness and user's willingness to continuing using the system. Users were also asked to answer an *AttrakDiff* questionnaire adapted from [12–14], in order to better know how users would rate the usability and the design(s) of the interactive prototype, capturing eventual changes over time, related with the increased usage on daily basis.

3.3. Experimentation Sessions: Hands-on Testing and Dialogue

We conducted a set of twelve hands-on user sessions and semi-structured interviews, involving ten end-users in total. In each session/workshop, as previously accorded, we count with the UI/UX champion, also responsible for the final UI mockups.

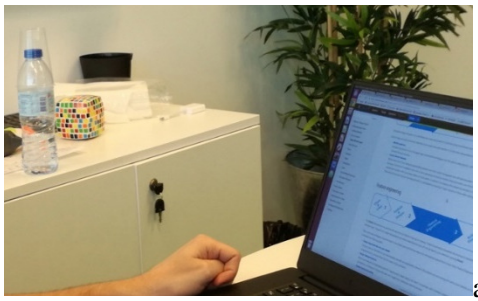
Participants were selected based on respective technical profile and project role; the type of Client project (e.g. e-commerce, merchant, banking). They were all company collaborators from the Research (business and product development), Data Science (data scientists) and Deployment (engineers) teams. Some have been working for the company for a few years (2-4 years) while for others, the user sessions functioned as on-boarding sessions, both to the company and to the tool. This was coordinated with the Humans Resource department. Only one user reported having previous experience with this kind of graphical tools, having use them for some of the DS activities (e.g. data analysis or model comparison), albeit no knowledge of eventual existing solutions to an end-to-end pipeline (actually, the end-to-end solution was company's core competitive advantage).

Each session started by having the UI/UX champion (an internal collaborator and participant's colleague) explaining the purpose of the session, emphasizing how the company is keen to ask the users and use their feedback to further improve the tool. This was used to engage the participant, create awareness of a transversal effort within the company towards a

user-centric perspective of IS design and development. This repeated process actually add to enhance user expectations, for best and worst, and established the UI/UX person as a "user champion" among stakeholders and software development team.

Each session carried on with us inquiring about the user background and learning more about user profile. We asked about specific attributes, such as age and formal education, but also about how they ended up doing their particular job (e.g. previous professional experience and extra education) and about hobbies and interests that they though contributed and about what they considered particular personality characteristics relevant to attain the job. This information was used to flesh and create several *Personas*, related to the different participants, with sufficient detail to the further creation of use-case scenarios by the development team. This was an effort to "learn from a sample of one or few" [18] since our participants are representative of future end-users.

Then we engaged on a loop of contextual usage of the prototype and a disruptive effort from the research team to interrupt the flow of thoughts with accomplishing the goal of each activity, "kindly pushing" user to explicitly articulate tactical knowledge, her/his own needs and interests or alternative ways of doing tasks and activities. Figure 3a) illustrates a particular aspect of the documentation and how it was accessed by the user, during one of the hands-on sessions. Figure 3b) illustrates the work screen, totally filled with two windows: on the right, the documentation about the particular module and functionality, with the tasks proposed (as a Tutorial), which steps the user has to follow and replicate on the left, by using the corresponding module interface to test.


a)

b)

Figure 3. Example of the experimentation technical environment of each session (January 2017).

Firstly, each session was observed as a user testing session (Q1 and Q2). The user would follow the steps in the tutorial, that mimic the actual steps to accomplish when doing the related job (either as data scientist or engineer). We handed a dataset purposefully different from those participants usually use, in order to provoke a discomfort. Each user would stop, ponder about the problem and tried to frame it in order to be able to solve it. This allowed us to question the mental models working and ask user to think-a-loud and share what (s)he is thinking, while trying to do or trying to accomplish (Q3). Sometimes, these two differ.

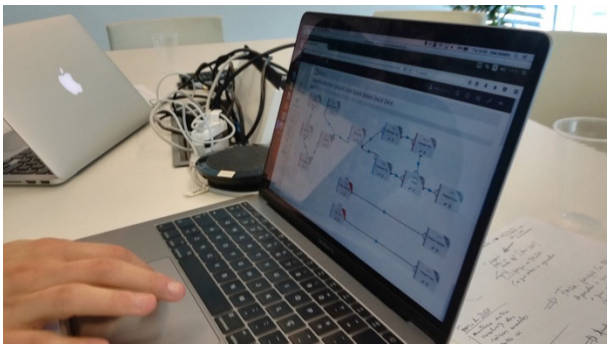


Figure 4. Testing the Drag & Drop interface (April, 2017).

We further asked what they could or wished to be doing in alternative (e.g. about the tool or about the flow of the DSL). Figure 4 shows one of this standard testing moment of the Drag & Drop mechanism, on a particular interface. Our user observation and the timely questions allowed to identify where visual aid and/or spatial help was needed, for improvement.

In Figure 5, we present two examples of interfaces that, following good practices of UI design, intend to provide feedback about the immediate results of the user's actions (e.g. signaling that '*There was an error. Showing details:*' and '*The fields will be removed. Save your changes?*', respectively), but resulting in screens of impossible legibility. In both examples, it is expected an informed decision by the user, but the interface is arduous and creates doubt instead of clarifying options for decision making, thus not really helping.

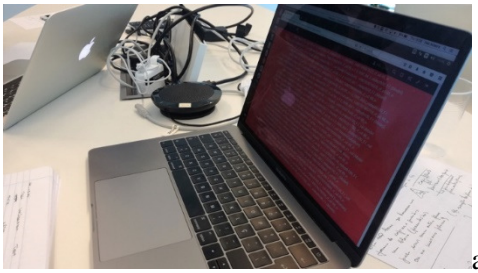
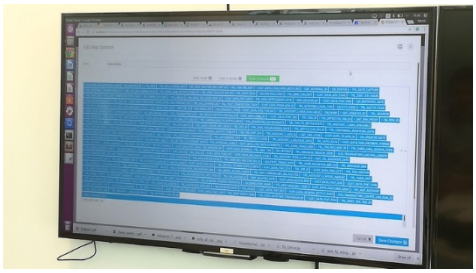

a)

b)

Figure 5. Feedback screens not very helpful and needing of usability enhancement (February 2017).

These are two examples of the kind of issues identified in the initial sessions, pointing specific and incremental improvements, that where addressed in the next versions, directly contributing to the next designed version of the prototype. Bugs identified were a strong priority, but these type of 'design bugs', albeit following design recommendation of providing contextual help were high prioritized for production.

Further considering the Q3, Figure 6 shows two exemplary cases where users resort and take advantage of auxiliary documents used to save relevant information and needed to complete the task. Directly observing these usage examples allowed to identify a problem that the DS system or the concrete module is not solving, creating a pain. Additionally, these moments allowed to identify a *potential* to i) complete the current interaction design (incremental design) or ii) to create and design a whole new feature, implicating changes in the artefact, but also, possibly, in the DSL process mapped (innovation design).

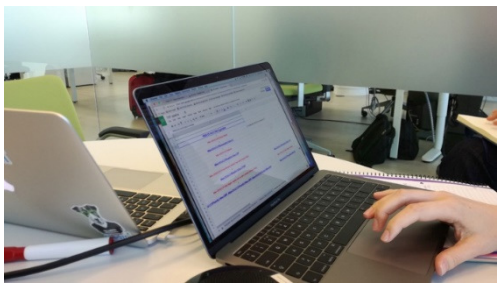
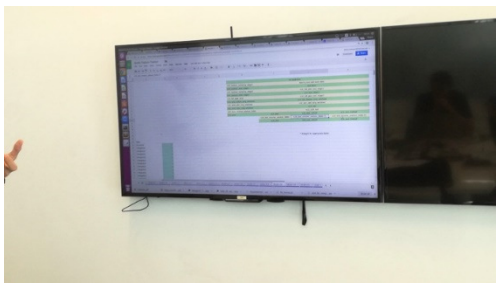



Figure 6. Users resort to auxiliary files, containing relevant information as memory aids (January 2017).

At these moments, sessions worked as "*what-if*" scenarios and were used to explore and discuss alternative solutions to accomplish the task, achieve the goal or improve the overall usability, satisfaction and/or memorability of the prototype. They would not relate directly with a non-supported functionality but more commonly illustrate a workaround or a way of doing the activity different from the implemented one. Alternative solutions to the specific problem or pain could then be discussed with the full engagement of the end-user, eventually translating into alternative UI mockups, for next prototype version.

Another important aspect of the project, closely related to the iterative approach of the agile software development methodologies, is the fact that we were able to also conduct *follow-up sessions*. Users involved in a first session (January 2017) were participating in second sessions, scheduled after passing some months (April and May 2017). In between sessions, they were using the DS prototype on respective Client project, sometimes reporting problems and alternative suggestions to the UI/UX person.

This continuous usage (afterwards the experimentation session), allowed a deeper appreciation of the issues found (e.g. how they impact in a daily basis) and of the effective improvement this kind of system introduces in the DSL daily work and activities. A long list of issues was mapped and partially are already solved, used to enhance the prototype over two iterations, within the duration of the project. Some of the issues, although, still waited for time and resources constrains to disappear and the opportunity to enter the production pipeline, albeit the identified need and its specific impact, in short and long term.

4. Shortcomings and Lessons Learned

We are aware of some of the limitations of our design research effort, particularly, being just a one-case scenario to test the fast-forward technique of getting user insights into development [18]. On one hand, the implementation of the DSR project had a positive impact on the overall system performance and actually improved usability and user experience of the DS prototype(s). The *insightz* workshop protocol proved to be practical within the fast-pace context of IS development and lean enough to be included in the sprint flow of work of the company. The issues mapped were pertinent and the effort to capture user observations and to reflect, not only about the tool, but also about the underlying process of data science activities, eventually added to the system improvement, from version to version. However, the drawback is the difficulty to track implications from the user sessions to the overall DSL and IT artefact improvement or to generalize its application on different contexts of IS design.

The company did establish a Product Team and a UX team (separating roles from front-end designers), fully aware of the advantages of a user-centric perspective and the use of socio-technical approaches to the design and development of a DS product. The *insightz* workshops remained a working thread in the development process, finding its way in an effective marriage of UX design and the agile software development sprints.

The project and, particularly, the flow of sessions added to the buy-in of the different teams. Front-end professionals and programmers were further aware of the effective tool usage and users' workarounds, when the UI or functionalities did not cope with the user interests, needs and/or expectations. Albeit, we found limitations when integrating the UX results into the development process. In each session, along with the team, we should have involved a person from the development team, since that would add to a strategy of effective communication of user insights into production.

We believe that researchers and practitioners can replicate the user session protocol in other contexts to generate rich data and information from hands-on sessions, guided by the three lenses of investigation presented. Yet, we need further work on how to support user observation (with IT tools, for instance), to capture the richness of user discussions upon the 'trial and error' effort and effectively translate these into insights and added systems requirements. Plus, there is a stronger need to address the collaboration between the development and deployment teams and the users, moving from the current user observation of "one user and one profile" at a time. The sessions remained with individual users and it was not possible to observe the interaction, supported by the prototype, between the different actors involved (and profiles: business analysts, engineers, data scientists, marketing and sales) in collaboration. Although we managed to do sessions were the participant would receive the result of another user participant (with different profile), the interweave of different user-profile experiences was difficult to capture.

The concept of designing interactive systems often emerges as an interdisciplinary approach, combining different methods and tools from various disciplines, as argued, for instance, by Service Design [24]. Mixing different tools, techniques and even design approaches is possible. Yet, we struggle to find guidance on which ones to use and when and

how to implement them, in fast-pace environments, where design happens while in use, in a flow of continuous use, evaluation and design improvement.

Some tools lack efficiency or amounts to too much work, falling behind the fast-pace of agile methodologies. IS and UI/UX teams need to have a strong understanding of the strengths and limitations of each tool and technique, including when and how to apply them during the development lifecycle. Other times, the issues and improvements found (even when agreed between users, stakeholders and decision makers), have to wait for an opportunity into the software production pipelines, due to resources constrains of various types. This can demotivate users, expecting that "issues are to be solved soon" or minimize the priority of user experience/usability threads in the development cycle. Dealing with and find solutions to handle these expectations was a great part of the design research project.

5. Conclusions

This paper presents a DSR approach to an ongoing information system development project, aiming to design and evolve an integrated IT system to support the complex and creative end-to-end DS pipeline. The project is being done in collaboration with an IT technology-based company rendering DS services, providing us with access to a real-life context for research.

Considering the resources and time constrains, we planned and conducted a set of user sessions, based on reflective experimentation, to capture user feedback upon iterative versions of an IT tool for DS. In this paper, we propose an *insightz* workshop, combining hands-on test and semi-structured interviews as a lean design technique to engage users in sharing insights, discuss alternative solutions to the design, reflecting on both tool and work practice.

During the DSR project, a prototype was designed, evaluated and refined over three versions, helping establishing a UX research and design process aligned with the company development process that follows agile methodologies. Further work will be to organize the design process followed and the lessons learned about mixing and scaling design methods across UX design teams, IT/IS development teams and business stakeholders. We plan additional exploratory sessions with stakeholders and decision makers to reflect on the solution-seeking process and the co-creation of design alternatives, working on the *extensibility* of these contributions.

Acknowledgements

The research work has been partially funded by the European Commission, under the Portugal 2020 structural fund (CENTRO2020), for the period of 2014-2020 (Ref. 2016/017728).

References

1. Abdelaziz, T.M., Maatukandfatmarajab, A., Maatuk, A.M., Rajab, F.: An Approach to Improvement the Usability in software Products. Int. J. Softw. Eng. Appl. 7 (2), (2016)
2. Almeida, A.S., Roque, L., Cunha, P.R. da: How to Design An Interactive System for Data Science: A Literature Review. In: Andersson, B., Johansson, B., Carlsson, S., Barry, C., Lang, M., Linger, H., and Schneider, C. (eds.) 27th International Conference on Information Systems Development. Lund University, Lund, Sweden (2018)
3. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: Adding an adjective rating scale. J. usability Stud. 4 (3), 114–123 (2009)
4. Baskarada, S., Koronios, A.: Unicorn data scientist: the rarest of breeds. Program-Electronic Libr. Inf. Syst. 51 (1), 65–74 (2017)
5. Brooke, J.: SUS: A Retrospective. J. Usability Stud. 8 (2), 29–40 (2005)
6. Chen, H., Chiang, R.H.L., Storey, V.C.: Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quartely. 36 (4), 1165–1188 (2012)
7. Davenport, T.H., Patil, D.J.: Data Scientist: The Sexiest Job of the 21st Century. Harv. Bus. Rev. 90 (10), 70+ (2012)
8. Dichev, C., Dicheva, D., Salem, W., Salem, W., Salem, W., Salem, W.: Towards Data Science Literacy. Procedia Comput. Sci. 108 (June), 2151–2160 (2017)

9. Frayling, C.: Research in Art and Design. In: Royal College of Art Research Papers. pp. 1–5. (1993)
10. Fu, X., Asorey, H.: Data-Driven Product Innovation. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 2311–2312. ACM, New York, USA (2015)
11. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Q. 37 (2), 337–355 (2013)
12. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff – Measure the attractiveness | UID, https://www.uid.com/en/publications/attrakdiff, Accessed: September 21, 2018, (2006)
13. Hassenzahl, M., Diefenbach, S., Göritz, A.: Needs, affect, and interactive products - Facets of user experience. Interact. Comput. 22 (5), 353–362 (2010)
14. Hassenzahl, M., Tractinsky, N.: User experience - A research agenda. Behav. Inf. Technol. 25 (2), 91–97 (2006)
15. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. 28 (1), 75–105 (2004)
16. Hewett, T.T., Baecker, R., Carey, T., Gasen, J., Perlman, G.: Curricula for Human-Computer Interaction ACM Special Interest Group on Computer-Human Interaction Curriculum Development Group.
17. Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P., Marrs, A.: Big data: The next frontier for innovation, competition, and productivity | McKinsey Company. (June), 1–22 (2011)
18. March, J.G., Sproull, L.S., Tamuz, M.: Learning from Samples of One or Fewer. Organ. Sci. 2 (1), 1–13 (1991)
19. McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D., Barton, D.: Big data: the management revolution. Harv. Bus. Rev. 90 (10), 61–67 (2012)
20. Mockus, A.: Operational data are missing, incorrect, and decontextualized. In: Perspectives on Data Science for Software Engineering. pp. 317–322. (2016)
21. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. J. Manag. Inf. Syst. 24 (3), 45–77 (2008)
22. Rittel, H.W.J., Webber, M.M.: Dilemmas in a General Theory of Planning. (1973)
23. Sauro, J.: Measuring Usability With The System Usability Scale (SUS), *Measuring Usability*, https://measuringu.com/sus/, Accessed: September 21, 2018, (2011)
24. Schneider, J., Stickdorn, M.: This is service design thinking. BIS Publishers, Amsterdam (2011)
25. Sonnenberg, C., Vom Brocke, J.: Evaluation Patterns for Design Science Research Artefacts. In: Helfert, M. and Donnellan, B. (eds.) Practical Aspects of Design Science. EDSS 2011. Communications in Computer and Information Science. pp. 71–83. Springer, Berlin, Heidelberg (2012)
26. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a Framework for Evaluation in Design Science Research. In: European Journal of Information Systems. pp. 423–438. Palgrave Macmillan UK (2012)
27. Venable, J.R.: Design science research post Hevner et al.: Criteria, standards, guidelines, and expectations. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 109–123. (2010)
28. Wamba, S.F., Akter, S., Edwards, A., Chopin, G., Gnanzou, D., Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., Gnanzou, D.: How `big data' can make big impact: Findings from a systematic review and a longitudinal case study. Int. J. Prod. Econ. 165 234–246 (2015)
29. Zimmerman, J., Forlizzi, J.: Research through design in HCI. In: Ways of Knowing in HCI. pp. 167–189. (2014)
30. Zimmerman, J., Forlizzi, J., Evenson, S.: Research through design as a method for interaction design research in HCI. In: Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07. pp. 493–502. ACM, New York, NY, USA (2007)