Visual Composition of Data Sources by End-Users

Carmelo Ardito, Giuseppe Desolda, Rosa Lanzilotti, M. Francesca Costabile
Dipartimento di Informatica, Università di Bari
Via Orabona 4, Bari
{name.surname}@uniba.it

Maristella Matera, Matteo Picozzi
DEIB, Politecnico di Milano
Piazza Leonardo da Vinci, 32
Milano, Italy
{surname}@elet.polimi.it

ABSTRACT
There is a huge and ever increasing amount of data sources available on the Web, which provide content through programmatic interfaces. Unfortunately, such data sources are accessible only through programming and therefore it is difficult for non-technical users to take advantage of such enormous data assets. The need therefore arises for paradigms to let laypeople, i.e., users without expertise in programming, explore and compose data sources. This paper discusses mechanisms for data source exploration and integration, which emerged from a study where laypeople were involved in discussions to gather their requirements about accessing and composing services. The paper also describes the prototypes that we defined to respond to the requirements highlighted by end users.

Categories and Subject Descriptors
D.2.2 Design Tools and Techniques: User Interfaces. D.2.10 Design: Methodologies. H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Design, Human Factors.

Keywords
Exploratory User Interfaces, Composition Paradigms, End-User Development, Service Composition.

1. INTRODUCTION AND MOTIVATION
The problem of facilitating the access to Web services and APIs through visual user interfaces has been attracting the attention of several researchers in the last years. There is indeed an ever increasing number of data sources that provide content in different formats through programmatic interfaces, while it is still difficult for laypeople, i.e., users without expertise in programming, to access and exploit the available content. This is true even if some data formats, for example, Linked data, are specifically meant to support data exploration and navigation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
AVT '14, May 27 - 29 2014, Como, Italy
Copyright 2014 ACM 978-1-4503-2775-6/14/05 …$15.00.
http://dx.doi.org/10.1145/2598153.2598201

Recent research projects have been dealing with the problem of easing the creation of effective presentations on top of Web services and APIs (e.g., [6]). They focused on the notion of Web Service Graphical User Interfaces (WSGUIs) [9], i.e., on a set of mechanisms to enrich the Web service specifications with annotations that could make the definition of visual interfaces easy. The idea was to automatically generate the presentation layer, starting from enriched service descriptors. Unfortunately, the proposed solutions were only able to generate dynamically dialogs for input and output of structured parameters; this compromised their adoption, especially considering the emerging trend of providing rich user interfaces going beyond the provision of forms and tables to query a service and access the query results respectively.

In the same period, the necessity arose to mash-up data services and Web APIs, by integrating their data sets by means of exploratory interfaces [3] and by synchronizing User Interface (UI) components, i.e., widgets already provided with their UI, integrating them at the presentation layer [5]. Several platforms were thus proposed both in academia and in industry. These approaches seemed to provide multiple advantages, and especially the possibility for laypeople to compose their own applications. However, even these approaches and the corresponding platforms did not succeed, mainly due to the complexity of the proposed composition paradigms, based on concepts closer to service protocols (such as data and control flow) than to the users’ mental model [7, 8, 10].

Our research focuses on the definition of composition platforms that, by exploiting end-user development approaches, enable even non technical users to take advantage of the huge quantity of data offered by data sources available online. This paper illustrates results we achieved by a study conducted with end users; they were involved in discussions to gather requirements about how to take advantage of online data sources, without the need of writing code. The motivation for this study comes out especially from some findings of our previous research, which was related to the definition of a meta-design framework to let end users add content items, extracted from different data sources, into personal information spaces [2]. Some past user studies helped us verify the usefulness for end users of the content made available by online data sources and the validity of our approach. The studies also let us identify possible improvements and extensions. In our first prototypes, end users were enabled to select content items from components that technology experts had previously packaged for them. We realized that end users, even when they do not have technical competences, might profit from packaging such components by themselves. In other words, we identified the need for a further
level of flexibility to be offered to end users to transform raw data extracted from remote data sources into components that are really meaningful to them. Of course, this further flexibility requires an adequate composition paradigm, which we decided to identify by investigating the users' mental model on how to explore and integrate data extracted from multiple sources. Section 2 reports on a study performed to collect end-user requirements. Section 3 then discusses the resulting prototype enabling people to explore and compose data sources. Section 4 concludes the paper.

2. USERS’ ACCESS TO DATA SOURCES: REQUIREMENTS

Very few studies are reported in the literature on understanding end-user requirements for accessing and integrating different services. The most significant study is reported in [10]: three separate focus groups were organized with the purpose of comparing three design alternatives of service-based systems enabling end-user composition. In our study we wanted to approach the problem differently, as we purposely did not ask end users to assess pre-defined composition paradigms for not introducing biases. Rather, we tried to elicit through discussion their mental model and attitude on accessing and composing services. Through interviews and a focus group the participants therefore discussed on how to compose data and services. The identified requirements were then used in three design workshops to model a composition process and to create some preliminary prototypes. In the sequel we will illustrate the study, outline the results and shortly illustrate the current version of our prototypes.

2.1 Participants and procedure

The interviews involved 9 people (4 male) aged 19-60 years old. The focus group involved 5 people (4 male) aged 22-25 years old. All participants had different backgrounds; they had low-medium knowledge of computer use, familiarity with Internet, but no programming experience; they were interested to collect contents retrieved from different sources in the Web (e.g., Wikipedia pages, YouTube videos, maps) in order to satisfy their work or daily life needs.

An HCI researcher conducted 5 single interviews, another conducted 4 single interviews and a third one moderated a focus group. Interviews and the focus group followed the same procedure. It was decided that no prototype should be shown, to avoid influencing the participants. Initially, the HCI researcher conducting the interview (the moderator in the focus group) gave a 10-minute presentation to explain what a web API is and the motivations for composing data coming from different sources, by also showing examples available on the Web. The researcher also briefly described our current platform. Then, the discussion started. Questions were formulated in two steps: a) illustrating, through a simple scenario that users could easily understand, how data from different sources are actually accessed using a web browser, and b) asking how the participants would like to perform such activities through a self-contained platform. Participants provided comments and, in the meanwhile, sketched interface elements, showing in practice how they would like to carry out each activity through an interactive system. The researchers took notes of users’ comments, opinions and discussions and collected all the sketches created in the study.

2.2 Results and discussion

The participants commented about platform benefits in their own daily life with a lot of enthusiasm. For example, a dance teacher said she frequently prepares choreographies and she happens to look for videos on YouTube and Vimeo. She clearly said that the platform would simplify the search, allowing her not to jump from a web page to another.

All participants said that they would like to use very simple and familiar interaction mechanisms, e.g., clicks of buttons and icons, menu item selection, drag-and-drop. Indeed, they specified these mechanisms in their sketches. They all indicated that the results coming from each source would be shown in a resource box in their own workspace.

On the basis of the participants’ answers and sketches, two proposals actually emerged for the union of data sets taken from different sources. 10 out of 14 participants said that they would like to perform the union through a button, labeled with “+” or “Add” and included in the title bar of the resource box to be extended. This button opens a popup window proposing a wizard procedure that guides the user through the steps for combining two sources. The 4 remaining participants proposed that the platform interface could provide a search box in the title bar of the source to be extended, in which the user types a keyword for the category of content the new source should provide, e.g., video, audio, text. A drop-down menu is visualized listing the name and, possibly, an icon of the retrieved sources, e.g., YouTube and others. The user drags the source of interest from the ones listed and drops it into the box of the source to be combined.

Three alternatives were identified from participants’ answers and sketches for the join of two or more data sources. The user started from a resource box showing data items coming from a specific source, e.g., Facebook. The first alternative, proposed by 6 participants, was to provide a button near each field of a data item, whose click activates another wizard procedure. 4 out of 14 participants proposed an alternative, in which the user positions the mouse cursor on the data item he/she wants to extend. The user clicks on the mouse right button to open a context menu showing a list of target sources from which he/she selects the source of interest. The third alternative was suggested by 4 participants: the platform shows a search box to search for the new source to be joined (a mechanism similar to the one used for the union operation). The user then drops one selected source in the field of a specific data item.

Finally, all participants expressed the desire to select the visualization of the returned items through a button opening a popup window where alternative visualizations are shown, so that the user can select the preferred one.

The study results showed that users like to be guided step by step in the composition of data sources. Participants indeed suggested a wizard procedure for performing both union and join of data. This is in line with the study reported in [10], where participants preferred the system-assisted composition approach over the other two, which were based on the definition of control flow and data flow respectively. The wizard, indeed, hides complex technical details and guides the users in selecting data sources and in combining or linking data items. On the other hand, mechanisms strictly based on parameter coupling, as it is for data and control flow approaches, are the “programmer’s way
of building software artifacts”, which does not necessarily match the end users’ mental models ([7, 11]).

At the end of user requirement gathering, a design team was set up. The team included the three HCI researchers involved in the data collection, an interface designer and a system developer. Based on the collected information, during a first design workshop the team identified the interaction process. In the two successive design workshops, prototypes of the enabling user interface were created and refined, also considering the results of both heuristic evaluations and tests with two end users. These prototypes are described in the next section.

3. EXPLORING AND INTEGRATING DATA SOURCES

Our current platform enables end users to create interactive workspaces, consisting of collections of services providing information in an integrated and synchronized fashion. Our approach requires a customization phase, performed only once, during which professional developers and domain experts collaborate to customize the general-purpose tools to the requirements of a specific domain [2]: 1) they register services relevant for the domain; 2) they define visual templates adequate to visualize domain specific information, which end users then exploit to organize the presentation of the desired content. Some studies have indeed shown that software environments for data source exploration and composition are more effective for non-technical people if customized to their specific domain [7, 11]. We now concentrate on describing how our current prototype, developed to satisfy the requirements emerged during the study reported in the previous section, supports end users in creating and extending integrated components to be included in their workspaces.

We suppose the end user is a male teenager, named Mario, keen on rock music. He is going to interact with a desktop design environment to create from scratch his workspace containing information about musical events (Figure 1). Mario has already included six services, each visualized in a resource box similar to the gadgets of the now dismissed iGoogle service: Youtube, Facebook contacts, SongKick, Flickr, Google Scholar and Pipl. Every service has been included in the workspace by means of the “add service” item listed in a menu displayed by selecting the gearwheel button in the upper right margin of the workspace. A wizard procedure has guided Mario: each service has been selected from a popup window, which shows the list of the registered services, classified by category (e.g., videos, photos, music, social); he has then chosen some of the data attributes returned by the service and has defined a visualization by mapping the chosen attributes onto the elements of a visual template. In case of YouTube, the visual template is a list of data items; for each item, a thumbnail and a title of the video are shown.

By typing a keyword in the search box shown in the title bar of each resource box, Mario can now retrieve contents from each service. In Figure 1, he has searched for videos about the Negramaro rock band on YouTube, the Facebook contacts of the user Mario Rodio, the musical events held in Liverpool from SongKick, the pictures of Times Square available in Flickr, papers about Mashup in Google Scholar and people named Mario in Pipl. He can also synchronize the content of the different resource boxes, according to an event-driven logic already provided by our platform [1].

If Mario is not satisfied by the results shown in a resource box, he can add, in the same box, data of another service. He can, for example, perform a Union between YouTube and Vimeo. In the title bar of the YouTube box, he clicks on the gearwheel button and selects the “Add results from a new source” command, to activate the wizard procedure. He selects Vimeo from a popup window showing sources; then, the Vimeo data attribute tree is visualized and Mario drags some Vimeo attributes into the YouTube box to define a mapping between the schemata of the two services. In the last step of Union, the wizard asks Mario to name the new service, e.g., “MyVideos”. As shown in Figure 2, the items in the MyVideos box come from both YouTube and Vimeo.

Figure 1. The workspace created by an end user.

Figure 2. An alternative visualization of some data.
Finally, Mario can change visualization. In Figure 2, Mario has selected the map visual template through the specific command available in the menu activated by the gearwheel button. The musical events, previously shown in a list that provides the details of each event, are now visualized as red pins on a map (top right in Figure 2). Each pin is placed at the location of the event, whose details can be visualized by clicking on the pin.

It is worth noticing that operations such as synchronization of components in the user workspace as well as union and join operations for data integration were implemented in a previous platform developed by some of the authors [4]. However, user tests showed that their actual use was clear for some users, but created difficulties to non-technical people, who are the main target of our work. Thus, we performed the studies illustrated in this paper to identify a more adequate interaction paradigm, also enhanced with assistance mechanisms for the selection of data sources.

The prototype presented here was also evaluated with end users to verify its usability. Five people were recruited: 3 had already participated in the requirement gathering and 2 were new ones. They were asked to create their workspace as shown in the examples above. During task execution, they commented that they were exploring the interface looking for widgets usually allowing content manipulations in well-known Web 2.0 applications, such as Google Drive. In particular, they appreciated finding commands in a menu activated by a gearwheel button, which was familiar to them. In general, they appreciated the interaction mechanisms provided by the platform. Only one participant who was involved in the requirement gathering showed some initial difficulties, because he tried to activate every function by the mouse right button. This is a behavior more typical for a person already acquainted with different computer applications rather than web applications, since for the latter the right button is unusual. However, after a more careful interface exploration, he realized how to work with the available mechanisms and was able to perform his tasks. Lack of space prevents us from giving more details, which could be however discussed at the conference, where a demo of the prototypes will be available.

4. CONCLUSION
The studies illustrated in this paper have been conducted to better understand end-user requirements. New prototypes have been developed on the basis of such requirements. We are going to evaluate them in a field study involving students and teachers of a high school in the Apulia region, in Southern Italy. In order to customize the platform to this new domain, we are collaborating with some teachers of the school to identify suitable services and visual templates. In the study, teachers and students will be asked to create their own workspaces, through which they will also share materials to better support their learning activities. The final aim of the studies performed in different domains is the identification of principles for the definition of sound paradigms for the exploration and composition by end users of online data sources.

5. ACKNOWLEDGEMENTS
This work is partially supported by the Italian Ministry of University and Research (MIUR) under the VINCENTE grant and by the Italian Ministry of Economic Development under the LOGIN grant. We thank students M. Rodio and R. Peskopia for contributing to system implementation. We are also grateful to all people involved in interviews and focus group performed in the reported studies.

6. REFERENCES