Activity Report 2019

Project-Team ILDA

Interacting with Large Data

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)
# Table of contents

1. Team, Visitors, External Collaborators ........................................ 1
2. Overall Objectives ................................................................. 2
3. Research Program ................................................................. 2
   3.1. Introduction .............................................................. 2
   3.2. Semantics-driven Data Manipulation .................................. 3
   3.3. Generalized Multi-scale Navigation .................................. 4
   3.4. Novel Forms of Input for Groups and Individuals ................. 5
4. Application Domains .............................................................. 6
   4.1. Mission-critical systems ............................................... 6
   4.2. Exploratory analysis of scientific data ............................. 6
5. Highlights of the Year ............................................................ 7
6. New Software and Platforms .................................................. 7
   6.1. Smarties ................................................................. 7
   6.2. ZVTM ................................................................. 8
   6.3. MapMosaic ........................................................... 8
   6.4. Baia ................................................................. 8
   6.5. LODAtlas ............................................................. 9
   6.6. TouchTokens .......................................................... 10
   6.7. Platforms .............................................................. 10
7. New Results ........................................................................... 12
   7.1. Digital Ink and Data Manipulation ................................... 12
   7.2. Novel Forms of Input in Immersive Environments ............... 13
   7.3. Multivariate Network Visualization .................................. 13
   7.4. Visualization in Specific Application Areas ....................... 15
8. Bilateral Contracts and Grants with Industry ................................ 16
9. Partnerships and Cooperations ................................................ 16
   9.1. National Initiatives ...................................................... 16
   9.2. European Initiatives .................................................... 17
   9.3. International Initiatives ................................................ 17
      9.3.1. Inria International Labs ......................................... 17
      9.3.2. Inria International Partners .................................... 17
10. Dissemination ....................................................................... 18
    10.1. Promoting Scientific Activities ..................................... 18
       10.1.1. Scientific Events: Organisation ................................ 18
          10.1.1.1. General Chair, Scientific Chair ......................... 18
          10.1.1.2. Member of the Organizing Committees ................ 18
       10.1.2. Scientific Events: Selection .................................. 18
          10.1.2.1. Chair of Conference Program Committees ............ 18
          10.1.2.2. Member of the Conference Program Committees .... 18
          10.1.2.3. Reviewer .................................................... 18
       10.1.3. Journal ........................................................... 18
          10.1.3.1. Member of the Editorial Boards ......................... 18
          10.1.3.2. Reviewer - Reviewing Activities ....................... 19
       10.1.4. Invited Talks .................................................... 19
       10.1.5. Scientific Expertise ............................................ 19
       10.1.6. Research Administration ....................................... 19
       10.1.7. Learned societies .............................................. 19
       10.1.8. Hiring committees ............................................. 19
    10.2. Teaching - Supervision - Juries .................................... 19
10.2.1. Teaching 19
10.2.2. Supervision 20
10.2.3. Juries 21
10.3. Popularization 21

11. Bibliography ................................................................. 21
Project-Team ILDA

Creation of the Team: 2015 January 01, updated into Project-Team: 2016 December 01

Keywords:

Computer Science and Digital Science:
A3.1.7. - Open data
A3.1.10. - Heterogeneous data
A3.2.4. - Semantic Web
A3.2.5. - Ontologies
A3.2.6. - Linked data
A5.1. - Human-Computer Interaction
A5.1.6. - Tangible interfaces
A5.2. - Data visualization
A5.5.4. - Animation
A5.6.1. - Virtual reality
A5.6.2. - Augmented reality

Other Research Topics and Application Domains:
B9.5.3. - Physics
B9.5.6. - Data science
B9.6.7. - Geography
B9.7.2. - Open data
B9.11. - Risk management

1. Team, Visitors, External Collaborators

Research Scientists
- Emmanuel Pietriga [Team leader, Inria, Senior Researcher, HDR]
- Caroline Appert [CNRS, Senior Researcher, HDR]
- Olivier Chapuis [CNRS, Researcher]

Faculty Member
- Anastasia Bezerianos [Univ Paris-Sud, Associate Professor]

Post-Doctoral Fellow
- Vanessa Peña-Araya [Inria, Post-Doctoral Fellow]

PhD Students
- Eugénie Brasier [Univ Paris-Sud, PhD Student]
- Emmanuel Courtoux [Univ Paris-Sud, PhD Student, from Oct 2019]
- Marie Destandau [Inria, PhD Student]
- Anna Gogolou [Inria, PhD Student, until Nov 2019]
- Raphaël James [Univ Paris-Sud, PhD Student]
- Hugo Romat [Tecknowmetrix, PhD Student, until Oct 2019, granted by CIFRE]
- Tong Xue [Inria, PhD Student]

Technical staff
- Léo Colombo [Inria, Engineer, until May 2019]
- Adhitya Kamakshidasan [Inria, Engineer, until Sep 2019]
2. Overall Objectives

2.1. Overall Objectives

In an increasing number of domains, computer users are faced with large datasets, that are often interlinked and organized according to elaborate structures thanks to new data models such as those that are arising with the development of, e.g., the Web of Data. Rather than seeing the inherent complexity of those data models as a hindrance, we aim at leveraging it to design new interactive systems that can better assist users in their data understanding and processing tasks.

These “Data-centric Interactive Systems” aim at providing users with the right information at the right time, presenting it in the most meaningful manner, and letting users efficiently manipulate, edit and share these data with others. This entails minimizing the effort required to retrieve and relate data from relevant sources; displaying data using visual presentation techniques that match the data’s characteristics and the users’ tasks; and providing users with means of interacting with the data that effectively support their train of thought.

Our approach is based on the idea of bringing the fields of Web data management [27] and Human-computer interaction [54], [79] closer together, based on the strong belief that they have the potential to cross-fertilize one another. User interface design is essential to the management and understanding of large, interlinked datasets. Interlinked datasets enriched with even a small amount of semantics have the potential to help create interfaces that let users analyze and manipulate data in a more efficient manner by providing them with, e.g., more relevant query results and giving them efficient means to navigate and relate those results. Our ultimate, long-term goal is to design interactive systems that make it as straightforward to manipulate large webs of data as spreadsheets do for tabular data.

3. Research Program

3.1. Introduction

Our ability to acquire or generate, store, process, interlink and query data has increased spectacularly over the last few years. The corresponding advances are commonly grouped under the umbrella of so called Big Data. Even if the latter has become a buzzword, these advances are real, and they are having a profound impact in domains as varied as scientific research, commerce, social media, industrial processes or e-government. Yet, looking ahead, emerging technologies related to what we now call the Web of Data (a.k.a the Semantic Web) have the potential to create an even larger revolution in data-driven activities, by making information accessible to machines as semistructured data [26] that eventually becomes actionable knowledge. Indeed, novel Web data models considerably ease the interlinking of semi-structured data originating from multiple independent sources. They make it possible to associate machine-processable semantics with the data. This in turn means that heterogeneous systems can exchange data, infer new data using reasoning engines, and that software agents can cross data sources, resolving ambiguities and conflicts between them [77]. Datasets are becoming very rich and very large. They are gradually being made even larger and more heterogeneous, but also much more useful, by interlinking them, as exemplified by the Linked Data initiative [49].
These advances raise research questions and technological challenges that span numerous fields of computer science research: databases, communication networks, security and trust, data mining, as well as human-computer interaction. Our research is based on the conviction that interactive systems play a central role in many data-driven activity domains. Indeed, no matter how elaborate the data acquisition, processing and storage pipelines are, data eventually get processed or consumed one way or another by users. The latter are faced with large, increasingly interlinked heterogeneous datasets (see, e.g., Figure 1) that are organized according to complex structures, resulting in overwhelming amounts of both raw data and structured information. Users thus require effective tools to make sense of their data and manipulate them.

Figure 1. Linking Open Data cloud diagram from 2007 to 2017 – http://lod-cloud.net

We approach this problem from the perspective of the Human-Computer Interaction (HCI) field of research, whose goal is to study how humans interact with computers and inspire novel hardware and software designs aimed at optimizing properties such as efficiency, ease of use and learnability, in single-user or cooperative work contexts. More formally, HCI is about designing systems that lower the barrier between users’ cognitive model of what they want to accomplish, and computers’ understanding of this model. HCI is about the design, implementation and evaluation of computing systems that humans interact with [54], [79]. It is a highly multidisciplinary field, with experts from computer science, cognitive psychology, design, engineering, ethnography, human factors and sociology.

In this broad context, ILDA aims at designing interactive systems that display [35], [61], [87] the data and let users interact with them, aiming to help users better navigate and comprehend large webs of data represented visually [7], as well as relate and manipulate them.

Our research agenda consists of the three complementary axes detailed in the following subsections. Designing systems that consider interaction in close conjunction with data semantics is pivotal to all three axes. Those semantics will help drive navigation in, and manipulation of, the data, so as to optimize the communication bandwidth between users and data.

3.2. Semantics-driven Data Manipulation

Participants: Emmanuel Pietriga, Caroline Appert, Anastasia Bezerianos, Marie Destandau, Hugo Romat, Tong Xue, Léo Colombaro.

The Web of Data has been maturing for the last fifteen years and is starting to gain adoption across numerous application domains (Figure 1). Now that most foundational building blocks are in place, from knowledge representation, inference mechanisms and query languages [50], all the way up to the expression of data presentation knowledge [70] and to mechanisms like look-up services [86] or spreading activation [43], we need to pay significant attention to how human beings are going to interact with this new Web, if it is to “reach its full potential” [44].
Most efforts in terms of user interface design and development for the Web of data have essentially focused on tools for software developers or subject-matter experts who create ontologies and populate them [56], [41]. Tools more oriented towards end-users are starting to appear [32], [34], [51], [52], [55], [64], including the so-called linked data browsers [49]. However, those browsers are in most cases based on quite conventional point-and-click hypertext interfaces that present data to users in a very page-centric, web-of-documents manner that is ill-suited to navigating in, and manipulating, webs of data.

To be successful, interaction paradigms that let users navigate and manipulate data on the Web have to be tailored to the radically different way of browsing information enabled by it, where users directly interact with the data rather than with monolithic documents. The general research question addressed in this part of our research program is how to design novel interaction techniques that help users manipulate their data more efficiently. By data manipulation, we mean all low-level tasks related to manually creating new content, modifying and cleaning existing content, merging data from different sources, establishing connections between datasets, categorizing data, and eventually sharing the end results with other users; tasks that are currently considered quite tedious because of the sheer complexity of the concepts, data models and syntax, and the interplay between all of them.

Our approach is based on the conviction that there is a strong potential for cross-fertilization, as mentioned earlier: on the one hand, user interface design is essential to the management and understanding of webs of data; on the other hand, interlinked datasets enriched with even a small amount of semantics can help create more powerful user interfaces, that provide users with the right information at the right time.

We envision systems that focus on the data themselves, exploiting the underlying semantics and structure in the background rather than exposing them – which is what current user interfaces for the Web of Data often do. We envision interactive systems in which the semantics and structure are not exposed directly to users, but serve as input to the system to generate interactive representations that convey information relevant to the task at hand and best afford the possible manipulation actions.

Relevant publications by team members this year: [22], [15], [17] and major ones in recent years: [7].

3.3. Generalized Multi-scale Navigation

Participants: Caroline Appert, Anastasia Bezerianos, Olivier Chapuis, Emmanuel Pietriga, Vanessa Peña-Araya, Marie Destandau, Anna Gogolou, Hugo Romat, Dylan Lebout.

The foundational question addressed here is what to display when, where and how, so as to provide effective support to users in their data understanding and manipulation tasks. ILDA targets contexts in which workers have to interact with complementary views on the same data, or with views on different-but-related datasets, possibly at different levels of abstraction. Being able to combine or switch between representations of the data at different levels of detail and merge data from multiple sources in a single representation is central to many scenarios. This is especially true in both of the application domains we consider: mission-critical systems (e.g., natural disaster crisis management) and the exploratory analysis of scientific data (e.g., correlate theories and heterogeneous observational data for an analysis of a given celestial body in Astrophysics).

A significant part of our research over the last ten years has focused on multi-scale interfaces. We designed and evaluated novel interaction techniques, but also worked actively on the development of open-source UI toolkits for multi-scale interfaces (http://zvtm.sf.net). These interfaces let users navigate large but relatively homogeneous datasets at different levels of detail, on both workstations [73], [29], [69], [68], [67], [30], [72], [28], [74] and wall-sized displays [63], [58], [71], [62], [31], [37], [36]. This part of the ILDA research program is about extending multi-scale navigation in two directions: 1. Enabling the representation of multiple, spatially-registered but widely varying, multi-scale data layers in Geographical Information Systems (GIS); 2. Generalizing the multi-scale navigation paradigm to interconnected, heterogeneous datasets as found on the Web of Data.
The first research problem has been mainly investigated in collaboration with IGN in the context of ANR project MapMuxing, which stands for multi-dimensional map multiplexing, from 2014 to early 2019. Project MapMuxing aimed at going beyond the traditional pan & zoom and overview+detail interface schemes, and at designing and evaluating novel cartographic visualizations that rely on high-quality generalization, i.e., the simplification of geographic data to make it legible at a given map scale [82], [83], and symbol specification. Beyond project MapMuxing, we are also investigating multi-scale multiplexing techniques for geo-localized data in the specific context of ultra-high-resolution wall-sized displays, where the combination of a very high pixel density and large physical surface enable us to explore designs that involve collaborative interaction and physical navigation in front of the workspace. This is work done in cooperation with team Massive Data at Inria Chile.

The second research problem is about the extension of multi-scale navigation to interconnected, heterogeneous datasets. Generalization has a rather straightforward definition in the specific domain of geographical information systems, where data items are geographical entities that naturally aggregate as scale increases. But it is unclear how generalization could work for representations of the more heterogeneous webs of data that we consider in the first axis of our research program. Those data form complex networks of resources with multiple and quite varied relationships between them, that cannot rely on a single, unified type of representation (a role played by maps in GIS applications).

Addressing the limits of current generalization processes is a longer-term, more exploratory endeavor. Here again, the machine-processable semantics and structure of the data give us an opportunity to rethink how users navigate interconnected heterogeneous datasets. Using these additional data, we investigate ways to generalize the multi-scale navigation paradigm to datasets whose layout and spatial relationships can be much richer and much more diverse than what can be encoded with static linear hierarchies as typically found today in interfaces for browsing maps or large imagery. Our goal is thus to design and develop highly dynamic and versatile multi-scale information spaces for heterogeneous data whose structure and semantics are not known in advance, but discovered incrementally.

Relevant publications by team members this year: [24], [20], [13], [14], [11], [19] and major ones in recent years: [10], [2].

3.4. Novel Forms of Input for Groups and Individuals

Participants: Caroline Appert, Anastasia Bezerianos, Olivier Chapuis, Emmanuel Pietriga, Eugénie Brasier, Emmanuel Courtoux, Raphaël James.

Analyzing and manipulating large datasets can involve multiple users working together in a coordinated manner in multi-display environments: workstations, handheld devices, wall-sized displays [31]. Those users work towards a common goal, navigating and manipulating data displayed on various hardware surfaces in a coordinated manner. Group awareness [48], [25] is central in these situations, as users, who may or may not be co-located in the same room, can have an optimal individual behavior only if they have a clear picture of what their collaborators have done and are currently doing in the global context. We work on the design and implementation of interactive systems that improve group awareness in co-located situations [57], making individual users able to figure out what other users are doing without breaking the flow of their own actions.

In addition, users need a rich interaction vocabulary to handle large, structured datasets in a flexible and powerful way, regardless of the context of work. Input devices such as mice and trackpads provide a limited number of input actions, thus requiring users to switch between modes to perform different types of data manipulation and navigation actions. The action semantics of these input devices are also often too much dependent on the display output. For instance, a mouse movement and click can only be interpreted according to the graphical controller (widget) above which it is moved. We focus on designing powerful input techniques based upon technologies such as tactile surfaces (supported by UI toolkits developed in-house), 3D motion tracking systems, or custom-built controllers [60] to complement (rather than replace) traditional input devices such as keyboards, that remain the best method so far for text entry, and indirect input devices such as mice or trackpads for pixel-precise pointing actions.
The input vocabularies we investigate enable users to navigate and manipulate large and structured datasets in environments that involve multiple users and displays that vary in their size, position and orientation [31], [45], each having their own characteristics and affordances: wall displays [63], [89], workstations, tabletops [66], [40], tablets [65], [84], smartphones [88], [38], [80], [81], and combinations thereof [39], [85], [62], [31]. We aim at designing rich interaction vocabularies that go far beyond what current touch interfaces offer, which rarely exceeds five gestures such as simple slides and pinches. Designing larger gesture vocabularies requires identifying discriminating dimensions (e.g., the presence or absence of anchor points and the distinction between internal and external frames of reference [65]) in order to structure a space of gestures that interface designers can use as a dictionary for choosing a coherent set of controls. These dimensions should be few and simple, so as to provide users with gestures that are easy to memorize and execute. Beyond gesture complexity, the scalability of vocabularies also depends on our ability to design robust gesture recognizers that will allow users to fluidly chain simple gestures that make it possible to interlace navigation and manipulation actions.

We also study how to further extend input vocabularies by combining touch [65], [88], [66] and mid-air gestures [63] with physical objects [53], [78], [60] and classical input devices such as keyboards to enable users to input commands to the system or to involve other users in their workflow (request for help, delegation, communication of personal findings, etc.) [33], [59]. Gestures and objects encode a lot of information in their shape, dynamics and direction, that can be directly interpreted in relation with the user, independently from the display output. Physical objects can also greatly improve coordination among actors for, e.g., handling priorities or assigning specific roles.

Relevant publications by team members this year: [9], [23], [16], [22] and major ones in recent years: [1], [10], [5], [3], [8].

4. Application Domains

4.1. Mission-critical systems

Mission-critical contexts of use include emergency response & management, and critical infrastructure operations, such as public transportation systems, communications and power distribution networks, or the operations of large scientific instruments such as particle accelerators and astronomical observatories. Central to these contexts of work is the notion of situation awareness [25], i.e., how workers perceive and understand elements of the environment with respect to time and space, such as maps and geolocated data feeds from the field, and how they form mental models that help them predict future states of those elements. One of the main challenges is how to best assist subject-matter experts in constructing correct mental models and making informed decisions, often under time pressure. This can be achieved by providing them with, or helping them efficiently identify and correlate, relevant and timely information extracted from large amounts of raw data, taking into account the often cooperative nature of their work and the need for task coordination. With this application area, our goal is to investigate novel ways of interacting with computing systems that improve collaborative data analysis capabilities and decision support assistance in a mission-critical, often time-constrained, work context.

Relevant publications by team members this year: [13], [19], [12].

4.2. Exploratory analysis of scientific data

Many scientific disciplines are increasingly data-driven, including astronomy, molecular biology, particle physics, or neuroanatomy. While making the right decision under time pressure is often less of critical issue when analyzing scientific data, at least not on the same temporal scale as truly time-critical systems, scientists are still faced with large-to-huge amounts of data. No matter their origin (experiments, remote observations, large-scale simulations), these data are difficult to understand and analyze in depth because of their sheer size and complexity. Challenges include how to help scientists freely-yet-efficiently explore their data, keep a trace of the multiple data processing paths they considered to verify their hypotheses and make it easy to
backtrack, and how to relate observations made on different parts of the data and insights gained at different moments during the exploration process. With this application area, our goal is to investigate how data-centric interactive systems can improve collaborative scientific data exploration, where users’ goals are more open-ended, and where roles, collaboration and coordination patterns [48] differ from those observed in mission-critical contexts of work.

Relevant publications by team members last year: [16], [24], [14], [18].

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Honorable mention for ActiveInk: (Th)Inking with Data at CHI 2019.

Best Papers Awards:


6. New Software and Platforms

6.1. Smarties

Functional Description: The Smarties system provides an easy way to add mobile interactive support to collaborative applications for wall displays.

It consists of (i) a mobile interface that runs on mobile devices for input, (ii) a communication protocol between the mobiles and the wall application, and (iii) libraries that implement the protocol and handle synchronization, locking and input conflicts. The library presents the input as an event loop with callback functions and handles all communication between mobiles and wall application. Developers can customize the mobile interface from the wall application side without modifying the mobile interface code.

On each mobile we find a set of cursor controllers associated with keyboards, widgets and clipboards. These controllers (pucks) can be shared by multiple collaborating users. They can control simple cursors on the wall application, or specific content (objects or groups of them). The developer can decide the types of widgets associated to pucks from the wall application side.

- Contact: Olivier Chapuis
- URL: http://smarties.lri.fr/
6.2. ZVTM

Zoomable Visual Transformation Machine

**KEYWORDS**: Big data - Visualization - Data visualization - Information visualization - Graph visualization

**FUNCTIONAL DESCRIPTION**: ZVTM is a toolkit enabling the implementation of multi-scale interfaces for interactively navigating in large datasets displayed as 2D graphics.

ZVTM is used for browsing large databases in multiple domains: geographical information systems, control rooms of complex facilities, astronomy, power distribution systems.

The toolkit also enables the development of applications running on ultra-high-resolution wall-sized displays.

- **Participants**: Caroline Appert, Olivier Chapuis and Emmanuel Pietriga
- **Contact**: Emmanuel Pietriga
- **Publications**: Rapid Development of User Interfaces on Cluster-Driven Wall Displays with jBricks - A Toolkit for Addressing HCI Issues in Visual Language Environments
- **URL**: http://zvtm.sf.net

6.3. MapMosaic

**KEYWORDS**: Geo-visualization - Data visualization

**SCIENTIFIC DESCRIPTION**: GIS software applications and other mapping tools enable users to correlate data from multiple layers and gain insight from the resulting visualizations. However, most of these applications only feature basic, monolithic layer compositing techniques. These techniques do not always support users effectively in their tasks, as we observed during interviews with GIS experts. MapMosaic is a novel approach based on dynamic visual compositing that enables users to interactively create and manipulate local composites of multiple vector and raster map layers, taking into account the semantics and attribute values of objects and fields in the compositing process. We evaluated MapMosaic’s interaction model against that of QGIS (a widely-used desktop GIS) and MAPublisher (a professional cartography tool) using the “Cognitive Dimensions” framework and through an analytical comparison, showing that MapMosaic’s model is more flexible and can support users more effectively in their tasks.

**FUNCTIONAL DESCRIPTION**: MapMosaic is a novel approach to combine geographical layers based on dynamic visual compositing that enables users to interactively create and manipulate local composites of multiple vector and raster map layers. It takes into account the semantics and attribute values of objects and fields in the compositing process. MapMosaic aims at better supporting GIS users in their tasks such as correlating data from multiple layers and gaining insight from the resulting visualizations.

**RELEASE FUNCTIONAL DESCRIPTION**: First public release.

- **Participants**: Maria Jesus Lobo, Caroline Appert and Emmanuel Pietriga
- **Contact**: Emmanuel Pietriga
- **Publications**: MapMosaic: Dynamic Layer Compositing for Interactive Geovisualization - An Evaluation of Interactive Map Comparison Techniques
- **URL**: http://ilda.saclay.inria.fr/mapmuxing/mapmosaic/index.html

6.4. Baia

**Before-and-after satellite image animation**

**KEYWORDS**: Geo-visualization - 2D animation
SCIENTIFIC DESCRIPTION: Before-and-after image pairs show how entities in a given region have evolved over a specific period of time. Satellite images are a major source of such data, that capture how natural phenomena or human activity impact a geographical area. These images are used both for data analysis and to illustrate the resulting findings to diverse audiences. The simple techniques used to display them, including juxtaposing, swapping and monolithic blending, often fail to convey the underlying phenomenon in a meaningful manner. Baia is a framework to create advanced animated transitions, called animation plans, between before-and-after images. Baia relies on a pixel-based transition model that gives authors much expressive power, while keeping animations for common types of changes easy to create thanks to predefined animation primitives.

FUNCTIONAL DESCRIPTION: Baia is a framework to create advanced animated transitions, called animation plans, between before-and-after satellite images.

Before-and-after image pairs show how entities in a given region have evolved over a specific period of time. Satellite images are a major source of such data, that capture how natural phenomena or human activity impact a geographical area. These images are used both for data analysis and to illustrate the resulting findings to diverse audiences. The simple techniques used to display them, including juxtaposing, swapping and monolithic blending, often fail to convey the underlying phenomenon in a meaningful manner. Baia relies on a pixel-based transition model that gives authors much expressive power. The animation editor enables authors to easily represent common types of changes thanks to predefined animation primitives and to sequence different changes across time.

RELEASE FUNCTIONAL DESCRIPTION: First public release
- Participants: Maria Jesus Lobo, Caroline Appert and Emmanuel Pietriga
- Contact: Emmanuel Pietriga
- Publication: Animation Plans for Before-and-After Satellite Images
- URL: http://ilda.saclay.inria.fr/mapmuxing/baia/index.html

6.5. LODAtlas

KEYWORDS: LOD - Linked open data - Semantic Web

SCIENTIFIC DESCRIPTION: The Web of Data is growing fast, as exemplified by the evolution of the Linked Open Data (LOD) cloud over the last ten years. One of the consequences of this growth is that it is becoming increasingly difficult for application developers and end-users to find the datasets that would be relevant to them. Semantic Web search engines, open data catalogs, datasets and frameworks such as LODStats and LODLaundromat, are all useful but only give partial, even if complementary, views on what datasets are available on the Web. LODAtlas is a portal that enables users to find datasets of interest. Users can make different types of queries about both the datasets’ metadata and contents, aggregated from multiple sources. They can then quickly evaluate the matching datasets’ relevance, thanks to LODAtlas’ summary visualizations of their general metadata, connections and contents.

FUNCTIONAL DESCRIPTION: The Web of Data is growing fast, as exemplified by the evolution of the Linked Open Data (LOD) cloud over the last ten years. One of the consequences of this growth is that it is becoming increasingly difficult for application developers and end-users to find the datasets that would be relevant to them. Semantic Web search engines, open data catalogs, datasets and frameworks such as LODStats and LOD Laundromat, are all useful but only give partial, even if complementary, views on what datasets are available on the Web. LODAtlas is a portal that enables users to find datasets of interest. Users can make different types of queries about both the datasets’ metadata and contents, aggregated from multiple sources. They can then quickly evaluate the matching datasets’ relevance, thanks to LODAtlas’ summary visualizations of their general metadata, connections and contents.
- Participants: Caroline Appert, Marie Destandau, Ioana Manolescu, François Goasdoué, Sejla Cebric, Hande Ozaygen and Emmanuel Pietriga
- Contact: Emmanuel Pietriga
6.6. TouchTokens

**KEYWORDS**: Tangible interface - HCI

**SCIENTIFIC DESCRIPTION**: TouchTokens make it possible to easily build interfaces that combine tangible and gestural input using passive tokens and a regular multi-touch surface. The tokens constrain users’ grasp, and thus, the relative spatial configuration of fingers on the surface, theoretically making it possible to design algorithms that can recognize the resulting touch patterns. See associated scientific articles below.

**FUNCTIONAL DESCRIPTION**: TouchTokens allow interface designers to build low-cost tangible interfaces. The technique consists in recognizing multi-touch patterns that are associated with specific passive tokens. Those physical tokens can be made out of any material to get tracked on any touch-sensitive surface. Implementations of the recognizer (in both TUIO and Android) and vector descriptions of the tokens ready for 3D-printing or laser-cutting are available

- **Participants**: Caroline Appert, Rafael Morales Gonzalez, Emmanuel Pietriga and Gilles Bailly
- **Contact**: Caroline Appert
- **Publications**: TouchTokens: Guiding Touch Patterns with Passive Tokens - Passive yet Expressive TouchTokens - Custom-made Tangible Interfaces with TouchTokens
- **URL**: https://www.lri.fr/~appert/touchtokens/

6.7. Platforms

6.7.1. Platform: WILDER

Ultra-high-resolution wall-sized displays [31] feature a very high pixel density over a large physical surface. Such platforms have properties that make them well-suited to the visualization of very large datasets. They can represent the data with a high level of detail while at the same time retaining context: users can transition from an overview of the data to a detailed view simply by physically moving in front of the wall display. Wall displays also offer good support for collaborative work, enabling multiple users to simultaneously visualize and interact with the displayed data. To make them interactive, wall-sized displays are increasingly coupled with input devices such as touch frames, motion-tracking systems and wireless multitouch devices, in order to enable multi-device and multi-user interaction with the displayed data. Application areas for such visualization platforms range from the monitoring of complex infrastructures and crisis management situations to tools for the exploratory visualization of scientific data.

WILDER is the latest ultra-high-resolution wall-sized display set up at Inria Saclay, and is one of the nodes of the Digiscope EquipEx. We use this platform for multiple projects, both fundamental HCI research, and research and development activities for specific application areas such as geographical informations systems and astronomy (Figure 2).
Figure 2. Example application area for ultra-high-resolution wall-sized displays: geographical information systems, and astronomical data analysis.
7. New Results

7.1. Digital Ink and Data Manipulation

We investigated how pen and touch could be best combined to facilitate the digital annotation of documents. When editing or reviewing a document, people directly overlay ink marks on content. For instance, they underline words, or circle elements in a figure. These overlay marks often accompany in-context annotations in the form of handwritten footnotes and marginalia. People tend to put annotations close to the content that elicited them, but have to compose with the often-limited whitespace. Based on these observations, we explored a design space – which we call SpaceInk (UIST 2019 [9]) – of pen+touch techniques that make room for in-context annotations by dynamically reflowing documents. We identified representative techniques in this design space, spanning both new ones and existing ones, as illustrated in Figure 3. We evaluated them in a user study. The results of this study then informed the design of a prototype system which lets users concentrate on capturing fleeting thoughts, streamlining the overall annotation process by enabling the fluid inverleaving of space-making gestures with freeform ink.

Together with colleagues from the EPIC team at Microsoft Research (see Section 9.3.2.1), we also investigated the potential of digital inking for exploring heterogeneous datasets and trying to make sense of them. During sensemaking, people annotate insights: underlining sentences in a document or circling regions on a map. They jot down their hypotheses: drawing correlation lines on scatterplots or creating personal legends to track patterns. Based on these observations, we designed ActiveInk (CHI 2019 [22]), a system enabling people to seamlessly transition between exploring data and externalizing their thoughts using pen and touch as input channels. ActiveInk, illustrated in Figure 4, enables the natural use of pen for active reading behaviors, while supporting analytic actions by activating any of these ink strokes. Through a qualitative study with eight
participants, we observed active reading behaviors during data exploration and design principles to support sensemaking.

7.2. Novel Forms of Input in Immersive Environments

ILDA researchers have started to investigate input techniques for the specific context of immersive environments, based on, e.g., virtual or augmented reality. These hardware devices enable displaying large amounts of data in space to better support data analysis, and there is a growing group of research focusing on Immersive Analytics in the HCI community. The question of input for data manipulation in these environments is crucial, but challenging because users must be able to activate various commands or adjust various values while remaining free to move. In this context, using the whole body as an input device offers several advantages: 1) the body provides physical support as an interactive surface, which improves accuracy and makes it less tiring to interact; 2) using the body does not impair mobility and avoids handling devices; 3) proprioception makes it possible to interact eyes-free, including when choosing values in a range; 4) by leveraging spatial memory, the body helps memorizing commands, thus interacting in expert mode (i.e., perform quick actions without visual feedback). ILDA team members participated to a position paper on this topic, which analyzes various ways of interacting with the body, discussing their pros and cons as well as associated challenges for immersive analytics [23].

7.3. Multivariate Network Visualization

Edges in networks often represent transfer relationships between vertices. When visualizing such networks as node-link diagrams, animated particles flowing along the links can effectively convey this notion of transfer. Variables that govern the motion of particles, their speed in particular, may be used to visually represent edge data attributes. Few guidelines exist to inform the design of these particle-based network visualizations, however. Following up on our initial investigation of motion as an encoding channel for edge attributes in

---

Figure 4. ActiveInk, a collaboration with Microsoft Research, affords smooth transition between using a digital pen for high-precision selections of heterogeneous data coming from multiple sources, and for externalizing thinking via notes and annotations. Ink strokes are leveraged to perform operations on underlying data.
Figure 5. Experimental task used in the study about the influence of color and size of moving particles on their perceived speed in node-link diagrams. Participants had to adjust the speed of particles on a specific edge so that it would match that of particles on another edge, using a slider.
multivariate network visualization [75], we investigated the influence of color and size of moving particles on their perceived speed in node-link diagrams (INTERACT 2019 [20]). Empirical studies so far had only looked at the different motion variables in isolation, independently from other visual variables controlling the appearance of particles, such as their color or size. We ran a study of the influence of several visual variables on users’ perception of the speed of particles. Part of the experimental setup is illustrated in Figure 5. Our results show that particles’ luminance, chromaticity and width do not interfere with their perceived speed. But variations in their length make it more difficult for users to compare the relative speed of particles across edges.

Beyond questions of perception of information in multivariate network visualizations, we also investigated the problem of creating environments for the design of multivariate network visualizations, with a focus on expressive design. Expressive design environments enable visualization designers not only to specify chart types and visual mappings, but also to customize individual graphical marks, as they would in a vector graphics drawing tool. Prior work had mainly investigated how to support the expressive design of a wide range of charts generated from tabular data: bar charts, scatterplots, maps, etc. But multivariate network data structures raise specific challenges and opportunities in terms of visual design and interactive authoring. Together with company TKM (see Section 8.1), we developed an expressive design environment for node-link diagrams generated from multivariate networks called Graphies (TVCG 2019, [15]), illustrated in Figure 6. We followed a user-centered design approach, involving expert analysts from TKM, and validated the approach through a study in which participants successfully reproduced several expressive designs, and created their own designs as well.

7.4. Visualization in Specific Application Areas

Finally, we worked in collaboration with other researchers on projects aimed at investigating how visualization can support experts in different application areas.
In the area of geovisualization, we performed a comparison of visualization techniques to help analysts identify correlation between variables over space and time (TVCG/InfoVis 2019, [6]). Observing the relationship between two or more variables over space and time is essential in many application domains. For instance, looking, for different countries, at the evolution of both the life expectancy at birth and the fertility rate will give an overview of their demographics. The choice of visual representation for such multivariate data is key to enabling analysts to extract patterns and trends. We conducted a study comparing three techniques that are representative of different strategies to visualize geo-temporal multivariate data. Participants performed a series of tasks that required them to identify if two variables were correlated over time and if there was a pattern in their evolution. Our results showed that a visualization’s effectiveness depends strongly on the task to be carried out. Based on this study’s findings, we derived a set of design guidelines about geo-temporal visualization techniques for communicating correlation.

Together with researchers from INRA, we performed an exploratory study about the visual exploration of model simulations for a range of experts (CHI 2019, [16]). Experts in different domains rely increasingly on simulation models of complex processes to reach insights, make decisions, and plan future projects. These models are often used to study possible trade-offs, as experts try to optimize multiple conflicting objectives in a single investigation. Understanding all the model intricacies, however, is challenging for a single domain expert. This project introduced a simple approach to support multiple experts when exploring complex model results, working concurrently on a shared visualization surface. The results of an observational study focusing on the link between expertise and insight generation during the analysis process, revealed the different exploration strategies and multi-storyline approaches that domain experts adopt during trade-off analysis. This eventually led to recommendations for collaborative model exploration systems.

We collaborated with researchers in databases from Université Paris Descartes on progressive similarity search on time-series data (BigVis 2019 workshop, [24]). Time-series data are increasing at a dramatic rate, yet their analysis remains highly relevant in a wide range of human activities. Due to their volume, existing systems dealing with time-series data cannot guarantee interactive response times, even for fundamental tasks such as similarity search. This paper presented our vision to develop analytic approaches that support exploration and decision making by providing progressive results, before the final and exact ones have been computed. Findings from our experiment indicated that there is a gap between the time the most similar answer is found and the time when the search algorithm terminates, resulting in inflated waiting times without any improvement. These findings led to preliminary ideas about computing probabilistic estimates of the final results that could help users decide when to stop the search process.

In the field of Education, we contributed to EduClust, an online visualization application for teaching clustering algorithms (EuroGraphics 2019, [18]). EduClust combines visualizations, interactions, and animations to facilitate the understanding and teaching of clustering steps, parameters, and procedures. Traditional classroom settings aim for cognitive processes like remembering and understanding. We designed EduClust for expanded educational objectives like applying and evaluating. The application can be used by both educators to prepare teaching material and examples, and by students to explore clustering differences and discover algorithmic subtleties.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry


9. Partnerships and Cooperations

9.1. National Initiatives
9.1.1. Inria Project Lab (IPL)

ILDA participates to Inria Project Lab iCODA: Data Journalism: knowledge-mediated Content and Data Interactive Analytics, that started in 2017. A key issue in data science is the design of algorithms that enable analysts to infer information and knowledge by exploring heterogeneous information sources, structured data, or unstructured content. With journalism data as a landmark use-case, iCODA aims to develop the scientific and technological foundation for collaborative, heterogeneous data analysis, guided by formalized, user-centric knowledge. The project relies on realistic scenarios in data-journalism to assess the contribution of the project to this area. iCODA is at the crossroads of several research areas (content analysis, data management, knowledge representation, visualization) and is part of a club of partners of the world of the press. Equipes-projets Inria : Graphik, Ilda, Linkmedia, Cedar. Press partners: Le Monde, OuestFrance, AFP. Participants: Anastasia Bezerianos (PI), Emmanuel Pietriga, Tong Xue, Vanessa Peña-Araya, Nicole Barbosa Sultanum.

9.2. European Initiatives

9.2.1. Collaborations with Major European Organizations

Deutsches Elektronen-Synchrotron (DESY): Scientific collaboration on the design and implementation of user interfaces for array operations monitoring and control for the Cherenkov Telescope Array (CTA) project, to be built in the Canary Islands (Spain) and in the Atacama desert (Chile), 2 years, contract started May 2018.

9.3. International Initiatives

9.3.1. Inria International Labs

Inria Chile. From 2012 to 2015, Emmanuel Pietriga was the scientific leader of the Massive Data team at Inria Chile, working on projects in collaboration with the ALMA radio-telescope and the Millenium Institute of Astrophysics. He is now scientific advisor to Inria Chile’s visualization projects, and is actively involved in the collaboration between Inria Chile and the LSST on the design and development of user interfaces for operations monitoring and control.

9.3.2. Inria International Partners

Association of Universities for Research in Astronomy (AURA): contract, jointly with Inria Chile, on the design and implementation of user interfaces for telescope operations monitoring and control for the Large Synoptic Survey Telescope (LSST) project, under construction in the Atacama desert (Chile), started 2017. Participants: Emmanuel Pietriga (ILDA), Sebastian Fehlandt (Inria Chile), José Galaz (Inria Chile), Sebastian Pereira (Inria Chile), Grazia Prato (Inria Chile).

9.3.2.1. Informal International Partners

We have had multiple collaboration projects with Microsoft Research in Redmond, USA. Hugo Romat visited the EPIC team for three months, and this collaboration led to the following publications at CHI 2018 [75], CHI 2019 [22] and UIST 2019 [9]. Anastasia Bezerianos also continues working with that team on topics related to smartwatch interaction and visualization that appeared in TVCG 2019 (InfoVis 2018) [11].

Our long-term collaboration with University of Konstanz, Germany continues. After publications at TVCG/InfoVis in 2014 and 2018 [46], [47], Anastasia Bezerianos has co-authored a paper at Eurographics 2019 with these colleagues [18].

Finally, our ongoing collaboration with Northwestern University, USA continues. Anastasia Bezerianos and past PhD student Evanthia Dimara (PhD defended in 2017) have worked on publications in TVCG 2019 [12] [42].
10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair


10.1.1.2. Member of the Organizing Committees

- IEEE VIS (IEEE Visualization Conference) 2019, 2020: Anastasia Bezerianos (Organization Committee - Workshops co-Chair)

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees


10.1.2.2. Member of the Conference Program Committees


- IEEE VIS 2019, Visualization Conference (InfoVis): Anastasia Bezerianos

- TheWebConf (WWW) 2020, 29th Web Conference, research track Web Mining and Content Analysis: Emmanuel Pietriga


- EICS 2019, 11th ACM Symposium on Engineering Interactive Computing Systems, Tech Notes, Olivier Chapuis

10.1.2.3. Reviewer


- ACM UIST 2019, Interface Software and Technologies Symposium: Olivier Chapuis, Emmanuel Pietriga, Hugo Romat

- ACM DIS 2019, Designing Interactive Systems: Caroline Appert

- ACM MobileHCI 2019, International Conference on Human-Computer Interaction with Mobile Devices and Services: Caroline Appert, Olivier Chapuis

- IEEE VIS 2019, Visualization Conference (InfoVis): Emmanuel Pietriga, Vanessa Peña-Araya (short papers)

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- ACM ToCHI, Transactions on Computer-Human Interaction: Caroline Appert (associate editor)
10.1.3.2. Reviewer - Reviewing Activities
- ACM ToCHI, Transactions on Computer-Human Interaction: Olivier Chapuis
- IEEE TVCG, Transactions on Visualization and Computer graphics: Anastasia Bezerianos

10.1.4. Invited Talks
- Caroline Appert: Lowering the cost of transitions between devices, Microsoft Research Faculty Summit, Seattle, USA, July 2019.
- Anastasia Bezerianos: Interactive visualizations and large displays, Indiana University, USA, December 2019.
- Olivier Chapuis: Multi-user interaction on wall displays, Public lectures on interactive design, Chalmers University, Göteborg, Sweden, November 2019.
- Emmanuel Pietriga: Créativité et visualisation de réseaux sous formes de diagrammes noeuds-liens, École d’été de cartographie et visualisation, Lyon, France, July 2019.

10.1.5. Scientific Expertise
- Evaluator for NSERC Discovery Grants (Canada): Anastasia Bezerianos, Olivier Chapuis

10.1.6. Research Administration
- Deputy Director of the Laboratoire de Recherche en Informatique (LRI): Olivier Chapuis
- President of Inria Saclay - Île de France’s Commission for Technological Development (CDT): Emmanuel Pietriga
- Deputy head of Pôle “Données, Connaissances, Apprentissage et Interaction” at École Doctorale Paris Saclay (ED STIC): Caroline Appert
- Member of Conseil de Département (LRI): Anastasia Bezerianos
- Member of Commissions Consultatives de Spécialités d’Université (CCSU) at Université Paris-Sud/Paris Saclay: Anastasia Bezerianos

10.1.7. Learned societies
- Association Francophone d’Interaction Homme-Machine (AFIHM), in charge of the relation with the SIF: Olivier Chapuis.
- SigCHI Paris Local Chapter, chair: Caroline Appert.

10.1.8. Hiring committees
- MCF Hiring committee, Univ. Paris Sud - LIMSI: Caroline Appert
- MCF Hiring committee, Univ. Technologique Troyes - Tech-CICO: Caroline Appert

10.2. Teaching - Supervision - Juries

10.2.1. Teaching
- Ingénieur: Emmanuel Pietriga, Data Visualization (INF552), 36h, 3A/M1, École Polytechnique.
- Ingénieur: Caroline Appert, Data Visualization (INF552), 18h, 3A/M1, École Polytechnique.
- Ingénieur: Vanessa Peña-Araya, Data Visualization (INF552), 16h, 3A/M1, École Polytechnique.
- Master: Emmanuel Pietriga, Data Visualization, 24h, M2 Informatique Décisionnelle, Univ. Paris-Dauphine.
Master: Anastasia Bezerianos, Career Seminar, 21h, M2 Interaction, Univ. Paris-Sud.
Master: Anastasia Bezerianos, Augmented Reality and Tangible Interaction, 12h, M2 Interaction, Univ. Paris-Sud.
Master: Anastasia Bezerianos, Interactive Information Visualization, 10.5h, M2 Interaction, Univ. Paris-Sud.
Master: Anastasia Bezerianos, HCI Project, 21h, M2 Interaction and HCID, Univ. Paris-Sud.
Master: Anastasia Bezerianos, Intro aux Systèmes Interactives, 15h, M1 Informatique, Univ. Paris-Sud.
Master: Eugénie Brasier, Programming of Interactive Systems, 10h30, M2 Interaction and HCID, Univ. Paris-Saclay.
Master: Eugénie Brasier, Career Seminar, 10h30, M2 Interaction and HCID, Univ. Paris-Saclay.
Master: Emmanuel Courtoux, Fundamentals of HCI, 21h, M2 Interaction and HCID, Univ. Paris-Saclay.
Master: Emmanuel Courtoux, Digital Fabrication, 21h, M2 Interaction and HCID, Univ. Paris-Saclay.
Master: Raphaël James, Programming of Interactive Systems, 12h, Univ. Paris-Saclay.
Master: Raphaël James, Fundamentals of HCI, 21h, Univ. Paris-Saclay.
Master: Tong Xue, Design of Interactive Systems, 21h, Univ. Paris-Saclay.
Licence: Eugénie Brasier, Algorithmique, 18h, PeiP2, Polytech.
License: Raphaël James, Algorithm/C/C++, 2h, Polytech.

10.2.2. Supervision

PhD in progress : Emmanuel Courtoux, Tangible Collaborative Interaction for Wall-sized Displays, since October 2019, Advisors: Olivier Chapuis, Caroline Appert
PhD in progress : Eugénie Brasier, Interaction techniques for remote manipulation in multi-display environments, since October 2018, Advisors: Caroline Appert
PhD in progress : Raphaël James, Environnements de réalité physique et augmentée utilisés dans l’analyse visuelle collaborative, since October 2018, Advisors: Anastasia Bezerianos, Olivier Chapuis, Tim Dwyer
PhD in progress : Tong Xue, Interactive Visualization for Data Journalism, since October 2018, Advisors: Anastasia Bezerianos, Emmanuel Pietriga
PhD in progress : Marie Destandau, Interactive Visual Exploration of Webs of Data, since October 2017, Advisors: Emmanuel Pietriga
PhD (defended) : Anna Gogolou, Iterative and expresssive querying for big data series, October 2016 - November 2019, Advisors: Anastasia Bezerianos, Themis Palpanas
PhD Internship: Nicole Barbosa Sultanum (from University of Toronto), Navigation and exploration of multivariate and dynamic graphs, 3 months, Advisor: Anastasia Bezerianos
Master 2 Internship: Marzieh Rafiei, Wall displays in Virtual Reality, 6 months, Advisors: Olivier Chapuis
Master 2 Internship: Emmanuel Courtoux, Tangible interaction for wall displays, 6 months, Advisors: Olivier Chapuis, Caroline Appert
10.2.3. Juries

PhD: Thibault Raffaillac, Université de Lille: Emmanuel Pietriga (examinateur)
PhD: Yujiro Okuya, Université Paris-Sud: Caroline Appert (president)
PhD: Gaëlle Richer, Université de Toulouse: Caroline Appert (reviewer)
PhD: Khanh Duy Le, Chalmers University at Göteborg: Olivier Chapuis (reviewer)

10.3. Popularization

10.3.1. Articles and contents

- 2019-06-12: Ciel & Espace - Sanctuary, le projet français qui vise à faire de la Lune un lieu de la mémoire humaine, [https://www.cieletespace.fr/actualites/sanctuary-le-projet-francais-qui-vise-a-faire-de-la-lune-un-lieu-de-la-memoire-humaine](https://www.cieletespace.fr/actualites/sanctuary-le-projet-francais-qui-vise-a-faire-de-la-lune-un-lieu-de-la-memoire-humaine)

11. Bibliography

Major publications by the team in recent years


Publications of the year

Articles in International Peer-Reviewed Journals


**International Conferences with Proceedings**


[20] **Best Paper**


[21] **Best Paper**


[22] **Best Paper**

Conferences without Proceedings


References in notes


[76] H. ROMAT. From data exploration to presentation : designing new systems and interaction techniques to enhance the sense-making process, Université Paris-Saclay, October 2019, https://tel.archives-ouvertes.fr/tel-02390645


Seoul, South Korea, ACM, April 2015, pp. 3255-3264 [DOI : 10.1145/2702123.2702129], https://hal.archives-ouvertes.fr/hal-01144312


