FP7 - Information and Communication Technologies

Grant Agreement no: 257184
Collaborative Project
Project start: 1 September 2010, Duration: 36 months

D4.3 - Collaboration and Decision Making
in Data-Intensive and Cognitively-Complex Settings:
Lessons Learnt from the Dicode Project

Due date of deliverable: 31 August 2013
Actual submission date: 30 August 2013
Responsible Partner: CTI
Contributing Partners: CTI, UOL

Nature: ❑ Report    □ Prototype    □ Demonstrator    □ Other

Dissemination Level:
❑ PU : Public
❑ PP : Restricted to other programme participants (including the Commission Services)
❑ RE : Restricted to a group specified by the consortium (including the Commission Services)
❑ CO : Confidential, only for members of the consortium (including the Commission Services)

Keyword List: Collaboration, decision making, sense making, Dicode Workbench, Dicode services, visualization, usability, software development, software integration, lessons learned.

The Dicode project (dicode-project.eu) is funded by the European Commission, Information Society and Media Directorate General, under the FP7 Cooperation programme (ICT/SO4.3: Intelligent Information Management).
The Dicode Consortium

Computer Technology Institute & Press “Diophantus” (CTI) (coordinator), Greece

University of Leeds (UOL), UK

Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (FHG), Germany

Universidad Politécnica de Madrid (UPM), Spain

Neofonie GmbH (NEO), Germany

Image Analysis Limited (IMA), UK

Biomedical Research Foundation, Academy of Athens (BRF), Greece

Publicis Frankfurt Zweigniederlassung der PWW GmbH (PUB), Germany
### Document history

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Status</th>
<th>Modifications made by</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01-08-2013</td>
<td>First draft</td>
<td>Spyros Christodoulou (CTI) Manolis Tzagarakis (CTI)</td>
</tr>
<tr>
<td>2</td>
<td>05-08-2013</td>
<td>Second draft</td>
<td>Nikos Karacapilidis (CTI)</td>
</tr>
<tr>
<td>3</td>
<td>12-08-2013</td>
<td>Third draft</td>
<td>Fan Yang-Turner (UOL) Lydia Lau (UOL) Vania Dimitrova (UOL)</td>
</tr>
<tr>
<td>4</td>
<td>20-08-2013</td>
<td>Fourth draft (sent for internal review)</td>
<td>Guillermo de la Calle (UPM) Natalja Friesen (FHG)</td>
</tr>
<tr>
<td>5</td>
<td>23-08-2013</td>
<td>Reviewer’s comments incorporated (version sent to the Dicode SC)</td>
<td>Nikos Karacapilidis (CTI)</td>
</tr>
<tr>
<td>6</td>
<td>30-08-2013</td>
<td>Final version (approved by SC, sent to Project Officer)</td>
<td>Nikos Karacapilidis (CTI)</td>
</tr>
</tbody>
</table>

### Deliverable manager
- Nikos Karacapilidis, CTI

### List of Contributors
- Nikos Karacapilidis, CTI
- Manolis Tzagarakis, CTI
- Spyros Christodoulou, CTI
- Fan Yang-Turner, UOL
- Lydia Lau, UOL
- Vania Dimitrova, UOL

### List of Evaluators
- Guillermo de la Calle, UPM
- Natalja Friesen, FHG

### Summary

This deliverable reports on practical lessons learned during the development of innovative collaboration and decision making support services in the context of the Dicode project. These lessons concern: (i) the methodology followed and process carried out for the development of the abovementioned Dicode services, (ii) the facilitation and enhancement of collaboration and decision making in data intensive and/or cognitively complex settings, and (iii) related technological and integration issues. Detailed evaluation reports, interviews and discussions within the development teams, as well as analysis of the use of the developed services by end-users through the associated log files, provided valuable feedback for the formulation and compilation of these lessons. By sharing insights gained in the context of the Dicode project, this deliverable aims to help people engaged in developing similar services.
## Table of Contents

1 Introduction ............................................................................................................................................. 5

2 Lessons learned ......................................................................................................................................... 5
   2.1 Software development methodology and process ................................................................. 5
   2.2 Facilitation and enhancement of collaboration and decision making ............................. 8
   2.3 Technical and Integration issues ................................................................................................. 12

3 Conclusions .............................................................................................................................................. 14

4 References .................................................................................................................................................. 15

Appendix: Features of CommBAT ............................................................................................................ 16
1 Introduction

The overall aim of Dicode’s WP4 was the design and development of innovative collaboration and decision making support services which would enable the capitalization of the outcomes of WP3 (focusing on voluminous and complex data mining services), by exploiting the reasoning abilities of humans. In particular, these services would enable the synchronous and asynchronous collaboration of stakeholders through adaptive workspaces, and serve alternative data visualization schemas. Dicode’s WP4 services aimed to facilitate (both individual and group) sense- and decision-making by supporting stakeholders in locating, retrieving and meaningfully interact with relevant information.

Dicode’s collaboration and decision making support services were developed to address the requirements and needs of the project’s use cases (see deliverables D4.1.1, D4.2.1, D4.1.2, D4.2.2, D4.1.3 and D4.2.3). To do so, they focused on challenging issues that include the data intensive and cognitively complex characteristics of the use cases, as well as the need for integration with services developed in other work packages in order to further support end users.

In the next section, we report on a series of practical lessons learned while developing and deploying these services. Our aim is to share insights and thus assist developers of such services in similar settings. Some of the lessons reported are related to the development of CommBAT, a standalone Windows application (‘Community Behaviour Analytics Tool’), which supports users to use the log data of Dicode Collaborative Workspace in order to explore and understand collaborative sensemaking behaviour. Since CommBAT was under continuous improvement during the last few months of the project (a first description was given in deliverable D4.2.3), its main features are summarized in the Appendix of this deliverable.

2 Lessons learned

The lessons reported in this deliverable are classified according to three perspectives: (i) the process and methodologies followed for developing collaboration and decision making support services, (ii) the facilitation and enhancement of collaboration and decision making in data intensive and/or cognitively complex settings, and (iii) the technologies used to implement and integrate the developed services. These lessons can be considered as ‘best practices’ for people involved in building similar applications. We note that some of the lessons reported below are not exclusively related to the Dicode’s collaboration and decision making support services; they may also be valid for other categories of Dicode services (e.g. the Dicode’s data mining services – see also deliverable D3.3).

2.1 Software development methodology and process

In this section, we present the outcome of our experiences with regard to the software development methodology and processes. By the term ‘software development methodology and processes’ we refer to the structuring, planning, and controlling the process of software development in the context of the Dicode project.
Lesson 1: Agile development methodologies are well suited for teams developing independently software that needs to be integrated into a single product.

At the beginning of the Dicode project, it was decided to adopt agile development methodologies for managing the Dicode services development (see deliverable D5.1.1). More precisely, the project proposed the use of the Scrum software development framework\(^1\) that emphasizes cyclical development with rather short feedback loops called ‘sprints’. During a sprint, a potentially shippable product increment is created. In general, the duration of such sprints is decided by the development team.

The development team of Dicode’s collaboration and decision making support services fully embraced the Scrum framework. The duration of each sprint was decided to be about two weeks. During each sprint, specific functionalities of the abovementioned services were implemented and immediately tested in order to assess their ability to be integrated. Such organization of the development process proved to be important in the context of the project’s integration tasks (collaboration and decision support services were designed to be integrated into the Dicode workbench). The adopted agile methodology was properly supporting the project’s integration efforts. Such short sprints were crucial in identifying weaknesses and shortcomings and helped immensely in adjusting and configuring the developed services in the context of the adopted integration architecture. They also helped developers to respond quickly in identified bugs and fine tune the integrated services according to the needs of the project’s use cases.

Lesson 2: Frequent meetings with all technical partners and detailed meeting minutes provide the way to address complex integration issues.

During the Dicode project, all technical partners agreed to conduct regular meetings to discuss and decide on the design and implementation of the foreseen services. In the course of the Dicode project, eight technical committee meetings were conducted and detailed minutes were kept (minutes were immediately uploaded to the project’s wiki\(^2\)). Meetings were held using videoconferencing tools, where technical partners were discussing their current state of the work, the important design decisions they took since the last meeting, as well as solutions to the project’s integration issues.

Such meetings and minutes showed to be very important in properly developing and integrating the collaboration and decision support services. In particular, these frequent meetings gave the opportunity to resolve misconceptions as well as redesign the collaboration and decision support services in order to properly and meaningfully interoperate with the rest Dicode services. Furthermore, they provided a project-wide reference framework which helped the synchronization of the development of the Dicode services. Finally, the meeting minutes functioned as design documents that aided the contextualization of all implementation efforts.

Lesson 3: Scenarios play an important role to elucidate the requirements analysis and provide tailored support to address big data concerns.

The aim of the Dicode project was to support the needs of specific use cases, which were characterized by high volumes of complex data. The Dicode use cases spanned three

---

\(^1\) [http://www.scrumalliance.org/](http://www.scrumalliance.org/)

\(^2\) Available at [https://wiki.dicode-project.eu/display/DIC/Dicode+Technical+Committee](https://wiki.dicode-project.eu/display/DIC/Dicode+Technical+Committee)
different fields: biomedical research, medical decision making and online marketing. While all three use cases are characterized by high volumes of data, each of them had different needs with respect to collaboration and decision making. To further clarify their requirements early in the project, a set of scenarios were outlined. These scenarios illustrated typical situations in each of these application domains. They proved to be important in addressing the particular needs of each use case and in developing the appropriate services and configuration mechanisms required.

In addition, scenarios helped in establishing a common view and vocabulary between technical and non-technical project partners when discussing the foreseen services. They also aided in identifying both the parts of the collaboration and decision making support services that should be generic and the parts that should be configurable, thus providing the tailored support for each use case. Finally, they helped developers to properly conceive the different big data concerns in each use case. This approach led to the development of the appropriate mechanisms to cope with related issues.

**Lesson 4: A standardized way to discuss and document innovative services aids the establishment of a common vocabulary among developers and facilitates integration tasks.**

In the context of the Dicode project, services play an important role in properly supporting the different needs of the project’s use cases. During the initial project meetings, where technical and non-technical partners were discussing the design, role and use of the foreseen services, the term ‘service’ was used with many different meanings and in different contexts. This led to misconceptions about what the term refers to, even among the technical partners. In particular, the term ‘service’ was used in discussions to denote the final software as it would be perceived by end-users, as well as the technical software interfaces that would allow other third-party software to invoke its operations.

The technical partners early on in the project attempted to disambiguate the term ‘service’ in order to properly set the foundations for the design and implementation of the Dicode services. This led to a project-wide standardization of the meaning, description and documentation of the developed services. According to it, all Dicode services were documented in a consistent way (for instance, see deliverables D4.1.1, D4.2.1, D4.1.2, D4.2.2, D4.1.3 and D4.2.3). Such a consistent and coherent description of services also helped immensely the design of the technical integration strategy to be followed.

**Lesson 5: Multi-disciplinary collaboration is critical in the design of innovative tools**

To better understand how a group collaborates and makes decisions, we conducted user studies to understand activities that take place in Dicode’s Collaborative Workspace. Logs of activities offer an opportunity to track how users try to make sense, argue and engage others when they try to solve a problem. Our goal was to understand the sensemaking activities of users that lead to a decision and how a collaborative workspace evolves over time.

To reach our goal, we developed CommBAT to support analytics of the log data. CommBAT is an analytics tool developed to support collaborative sensemaking research. The development of such a tool requires expertise not only from data analysts, but also from experts in collaborative sensemaking, visualization and semantic technologies. Data analysts contribute to the design on how CommBAT should be used and what aspects of the data are important for the analysis. Researchers in collaborative sensemaking provide theoretical
support on the process of collaborative sensemaking. A sensemaking model is useful for the interpretation of the log data. Experts in visualization contribute to the presentation of the data and user interaction techniques. Semantic technologists help to build a rich semantic model to provide a foundation of data analytics. As a working prototype, CommBAT has made a progress towards understanding collaborative sensemaking behaviour with efforts of a team of multi-disciplinary expertise. The next step is to deepen the collaboration and work with wider communities to evaluate its features and improve the design.

2.2 Facilitation and enhancement of collaboration and decision making

This subsection reports on experiences gained with respect to the facilitation and augmentation of collaboration and decision making in data-intensive and/or cognitively complex contexts.

Lesson 6: Innovative metaphors of collaboration, although useful, may confuse users and should be introduced in a way that is close to what users are expecting.

Whenever innovative metaphors to collaboration are provided to users, these must be carefully introduced, as there lurks the danger of tool rejection due to encountering of a new and unexplored ‘territory’. In general, when users get to use collaborative systems, they expect (based on their experience) to see functionalities offered in wikis, discussion forums and tagging systems, as these are the prevailing applications nowadays on the Web. Our experience showed that radical new ways to collaboration may initially cause confusion rather than excitement.

To address such concerns, the related Dicode services offered the ability to render the collaboration in a way that is familiar to users, by providing alternative collaboration and decision making ‘views’. Dicode even enabled a forum-like view of the collaboration (called time-order-view) that displayed the discussion in a way that is found in traditional Web-based discussion forums. This functionality proved to be very helpful as the spatial metaphor of collaboration workspaces (i.e. the “mind-map view”) was regarded as simply another way of viewing and conducting an ongoing collaboration.

Lesson 7: Alternative views of collaboration may significantly tame the complexity of data-intensive workspaces. In such environments, formality in managing collaboration should not be considered as a predefined and rigid property, but rather as an adaptable aspect that can be modified to meet the needs at hand.

Generally speaking, existing collaboration support tools provide only a fixed set of abstractions, with which participants may express their opinion and allow only one way of visualizing the associated discourse. More specifically, participants’ interaction is regulated by procedures that prescribe and - at the same time - constrain their work. This may refer to both the system-supported actions a user may perform (e.g. types of discourse or collaboration acts), and the system-supported types of collaboration objects (e.g. one has to strictly characterize a collaboration object as an idea or a position). In many cases, users have also to fine-tune, align, amend or even fully change their usual way of collaborating in order to be able to exploit the system’s features and functionalities. While such approaches to supporting collaboration are in general useful and used today in a wide range of situations, they are problematic when used in data intensive settings. Traditional collaboration support systems (such as Web-based forums) fail to cope with the great number of items that are
uploaded and discussed. One reason for the lack of the proper support is the fixed nature of
the available abstractions and static visualization options. There is much evidence that an
inflexible set of abstractions often resulted in failures (Scheuer et al., 2010).

To overcome such concerns in the Dicode project, the collaboration support services were
designed to enable multiple views of the same collaboration discourse. Each view introduces
a different set of abstractions and a unique way to visualize the collaboration. Dicode’s
collaboration support services provide the following meaningful views (for details, see
deliverables D4.1.1, D4.1.2 and D4.1.3):

- **Discussion-Forum view**: a collaboration space is displayed as a traditional web-based
  forum, where posts are displayed in ascending chronological order. Users are able to
  post new messages to the collaboration space, which will appear at the end of the list of
  messages. The aim of this view is to allow the collection and sharing of opinions
  without limiting the expressiveness of participants.

- **Mind-Map view**: a collaboration space is displayed as a ‘mind map’, where users can
  interact with the items uploaded so far. The map deploys a spatial metaphor
  permitting the easy movement and arrangement of items on the collaboration space.
  The aim of this view is to support ‘information triage’ (Marshall and Shipman, 1997),
  i.e. the process of sorting and organizing through numerous relevant materials and
  organizing them to meet the task at hand.

- **Neighbourhood view**: this view displays a specific item along with its neighbours. The
  “neighbourhood” of a specific collaboration item is defined as the set of items with
  which this item is directly connected via a relation in the “Mind-Map view”. The aim
  of this view is to allow users to focus on directly connected items and not be distracted
  by others.

- **Formal view**: this view enables the posting of predefined knowledge items, which
  adhere to a specific argumentation model. It invokes a set of dedicated scoring and
  reasoning mechanisms aiming to aid users conceive the outcome of a particular
  collaborative session and receive support towards reaching a decision.

Every collaboration workspace in Dicode can be operated and visualized in any of the above
views. Users may also switch from one view into another, in order to visualize differently the
discourse and use a more appropriate set of functionalities to manage the collaboration
items. By doing so, they can also easily spot issues that need attention. Such a flexible way of
visualizing and working with the discourse proved to be important in data intensive
settings, as the environment could adapt to the increasing number of available resources. It
was admitted that alternative views may also reveal previously unseen and potentially
valuable insights. Furthermore, these views are based on a unifying conceptual framework
which permits an ‘incremental formalization’ of discourses (Shipman and McCall, 1994). Incremental formalization of collaboration proved to be a successful approach in the context of Dicode. In our approach, formality and the level of knowledge structuring was not
considered as a predefined and rigid property, but rather as an adaptable aspect that can be
modified to meet the needs of the tasks at hand. By the term formality, we refer to the rules
enforced by the system, with which all user actions must comply. Allowing formality to vary
within the collaboration space, a stepwise and controlled evolution from a mere collection of
individual ideas, arguments, annotations and resources to the production of highly
contextualized and interrelated knowledge artifacts and actual decisions, can be effortlessly
achieved.
Lesson 8: Collaboration and decision making services should not be regarded as ‘application islands’. Seamless interoperability is a crucial factor for their adoption and success.

From the users’ initial needs and ongoing feedback, openness and seamless interoperability appeared to be a primary need and constitute one of the biggest challenges of today’s social software applications (Buytaert, 2008). Users want to gain “real ownership” over the information that they have provided and/or belongs to them (e.g. their profile information, projects, and friends). They want to be able to easily import/export data from one environment to another. They want to be able to synchronize information across different tools and visualize it in different ways via different applications. Thus, a seamless integration of distributed tools and services is instrumental when developing innovative collaboration solutions. As a plethora of resources are already available on the Web, collaboration services must explicitly address issues regarding the integration of these resources into their environments. Otherwise, the danger of becoming isolated may surface and ultimately lead to their rejection.

In this direction, the related Dicode services not only facilitate the synchronous and asynchronous collaboration of stakeholders through adaptive workspaces, but they can also efficiently handle the representation and visualization of the outcomes of the data mining services (through alternative and dedicated data visualization schemas) and enable the orchestration of a series of actions for the appropriate handling of data. In addition, they provide an interactive mechanism for indexing and searching of standard documents.

Lesson 9: Effective collaboration and decision making requires appropriate mechanisms tailored to the needs of each use case.

Collaboration and decision making services in Dicode, apart from providing a number of generic mechanisms and functionalities, have been carefully tailored to meet the needs of each one of the three project’s use cases. Concerning the collaboration support services, each stakeholder can express his concerns and thoughts through collaboration objects of diverse object types (each collaboration object, apart from its content, carries a specific semantic reflected by its object type). In addition to the predefined (common for all three use cases) collaboration objects types, stakeholders may use a number of different sets of object types dedicated to each specific use case.

Regarding the decision making services, the selection of the implemented decision making algorithms was based on a questionnaire filled in by senior decision makers acting in diverse data-intensive settings (from the three use cases). According to the results of this questionnaire, decision making algorithms are highly related to the specific problem under consideration. Depending on the specific problem, decision makers require support from algorithms that: (i) allow compensation among the attributes/criteria used for the evaluation of the alternatives (i.e. a good performance of an alternative concerning one attribute can compensate for a bad performance concerning another attribute), (ii) allow two or more alternatives to be incomparable, and (iii) do not allow compensation among criteria. Three Multi-Criteria Decision Making algorithms, fulfilling the aforementioned prerequisites, were implemented in the context of Dicode: the Weighted Sum Model (WSM), the Analytical Hierarchy Processing (AHP) and the Lexicographic Decision Making rule (LDM).
Lesson 10: Analysis of data-intensive collaboration requires innovative and efficient tools.

Data analytics can be ad-hoc and complex. Different analysts can take different approaches for their tasks: statistics, visualization or intuitive observation (Fisher et al., 2012). For example, for Dicode workspace log data, we conducted analysis using three different tools: Microsoft Excel for cleaning and filtering the data; Mathematica³ for statistical analysis; Tableau⁴ for data visualization. A typical process of this analysis could be (illustrated below):

- Read a Dicode log file (text format) into Excel to conduct basic cleaning and analysis (grouping, sorting or filtering etc.);
- Conduct further analysis in Mathematica, using the output for Tableau or Excel, and
- Read data from Mathematica or Excel into Tableau for visualization.

This process involves a lot of repetitive activities, such as copying and pasting, saving results in different file formats etc. Sometimes, these steps have to be repeated over and over again for different factors considered. To avoid multiple tools and repetitive activities, there is a need for a more efficient tool to support analysts and facilitate the process. To meet this need, we took an innovative approach, developing CommBAT, a tool that enables users to have the features of statistics, visualization and interactivity at one interface. As presented in the Appendix of this document, with its features, CommBAT can provide a more efficient way for users to conduct analysis and gain insight of the data. However, it needs to be tested with more researchers and analysts to evaluate its benefit and examine how to bring best practice to the users.

Lesson 11: Data analytics is an iterative exploratory task which requires multi-perspective view support.

Data analytics is not a ‘one-off’ but an exploratory activity. The ability to easily extract meaning from complex datasets has become something of a Holy Grail in the tech industry (Giurata, 2012). This is especially true if we intend to gain deep insight about user or group behaviour because it involves lots of different factors. We can look at individual user aspect about who is doing what and in what way; we can look at group aspect on how they interact with each other; we can look at the activity aspect about what activities are involved; we can look at the objects aspect about which object has gained most of the attention. A way to develop insight is to interact with these aspects and look at these factors in different perspectives.

CommBAT has enabled users to interact with all these factors and provide different views of the result. However, there is a lot future work to do to improve the tool, such as free combination of different factors and multiple selections of factor items.

³ http://www.wolfram.com/mathematica/
⁴ http://www.tableausoftware.com/
Lesson 12: The need of rich semantics model to support design and analysis of collaborative workspaces

One of our important research questions was to understand how a collaborative workspace evolves, for example: how were ‘ideas’ created? Were they all created at the beginning of the argumentation or at the end of argumentation? To answer such questions, we took a first step to utilize the semantic types to present knowledge objects in ‘Object Type view’. Object Type view tried to use the semantic features of Object to visualize the collaborative workspace. However, the current semantic types are quite simple and a rich semantic model is needed to support the design of the tool on how to semantically interpret the data.

2.3 Technical and Integration issues

In this subsection, we refer to technologies used to implement and integrate the developed services. In particular, we focus on our experiences concerning the integration of specific services for collaboration and decision making support, the technology that was used for this integration, as well as the integration of open source frameworks.

Lesson 13: Integrating data mining into collaboration support services makes the collaboration discourse more understandable and greatly facilitates collective sense and decision making in data-intensive environments.

Integration issues were a central part of the Dicode project. Discussions among project partners were focused on questions such as which integrations are useful in the context of the Dicode use cases, as well as how to technically achieve it. By elaborating the project’s use cases and analysing their particular needs with the use case partners, a decision was made to tightly integrate data mining within the collaboration and decision making support services. This was motivated by the observation that in many use cases, such as in the biomedical and marketing domain, data mining and collaboration and decision support services were consistently used in a very specific pattern: in these use cases, stakeholders used collaboration and decision making support services to plan the execution of data mining services as well as to comment on their outcomes. The aim of such integration was to make the implicit relationship of these services explicit, thus facilitating the associated workflow.

This integration allowed data mining algorithms to be first class abstractions in the context of collaboration workspaces – by introducing a new semantic type called “service”- and be part of the discourse elements. This means that data mining algorithms can be used in collaborative discourses as any other discourse element (e.g. ‘notes’, ‘comments’ and ‘ideas’). Furthermore, data mining algorithms available within a collaboration workspace can be easily executed (even by stakeholders who are not data scientists or analysts) and, after their completion, the results can be automatically uploaded into the collaboration workspace making them available for interpretation and further contemplation. In such a way, stakeholders are able to ask questions of the data based on their own expertise and easily find patterns, spot inconsistencies, or even get answers to questions they have not yet thought to ask.

By allowing the integration of service items into collaboration workspaces and their treatment as any other discourse element, the discourse contextualizes their use, execution and outcomes, thus greatly contributing to its understanding. In addition, as the specific settings under which the services have been executed are stored as part of their metadata
(service parameters), the collaboration workspace makes it easy to support the provenance of the results. Such contextualization and understanding were not possible in contemporary data-intensive collaborative environments, which separate the execution of data mining services from their use in the context of collaboration support systems. This ultimately hindered participants from fully comprehending the discourse and made it difficult to assess the origins of the associated resources.

**Lesson 14: Open Source visualization libraries are mature enough to support visualization in data intensive environments.**

As the amount of the digital information nowadays is rapidly increasing, one of the major challenges is to make effective use of this vast amount of information. Visual data analysis and information visualization, facilitated by interactive interfaces, enable the detection and validation of expected results, highlight unexpected discoveries in data, validate new theoretical models, provide comparison between models and datasets, enable qualitative and quantitative querying, facilitate decision making and, in general, enable effective data processing and management in data-intensive environments (Hansen et al., 2009).

Driven by: (i) the growing number of open source visualization libraries which have recently emerged (and become popular), (ii) the prerequisites of the Dicode project to embed open source libraries, and (iii) the suggestions of the Project Officer (as expressed in the project reviews), open source visualization libraries have been effectively integrated to provide part of the functionality of the collaboration and decision making support services. In particular, popular open source frameworks have been used to:

- visually outline information and depict the collaboration process in the form of a ‘mind-map’;
- implement user-friendly components to allow sharing of the collaboration in a number of social software applications;
- provide a graphical representation of the collaboration data and metadata through a number of plots and charts;
- implement a number of supplementary visualization features to enable provenance of information and reduce the data intensiveness in the collaboration environment.

The exploitation of these open source frameworks led us argue that open source visualization libraries are mature enough to support visualization in data intensive environments. Moreover, the required manpower to adapt the related libraries to the needs of the Dicode project proved to be pretty small when compared to the manpower required to implement all functionality from scratch, while the quality of the final implementation remained high.

**Lesson 15: REST-based services can support the tight and fine-tuned integration required in data intensive environments.**

From early on, the Dicode project investigated different technologies that can be exploited to implement the required integration of the foreseen services. In particular, we experimented with and compared existing integration architecture styles in the context of the project’s data intensive use cases (see deliverable D2.1). This review of existing technologies led the technical partners to adopt a REST-based approach to service development and integration.

For instance, the REST-based approach was used for the development of the Dicode workbench (for executing actions such as user logging in a collaboration workspace,
creating/updating a collaboration workspace, creating/updating a user’s profile, creating a collaboration object from a resource stored in the Dicode repository, etc.). The REST calls implemented for the specific integration proved to be simple and lightweight as they were based on normal HTTP requests. The developer effort required was the minimum possible as no extra toolkits or libraries were involved in the implementation, while the human readable output of the REST calls was a great advantage during the debugging process and the tailoring of the REST calls to meet the exact requirements of the consumer service. The only issue worth mentioning concerns the encryption standards and the algorithms that were used for the data encryption (as the developers of the services used different technologies to implement their services). A close coordination of the developers’ efforts was required in order to follow the best possible common standard.

3 Conclusions

Developing collaboration and decision making support services to be used in data-intensive and/or cognitively-complex environments is certainly a challenging task. During the design and development of such services in the Dicode project, challenges encountered concerned different areas, such as the development methodology, the facilitation and enhancement of collaboration, as well as the integration technologies. By addressing these challenges, valuable experiences were collected and lessons were learned. For each lesson reported in this deliverable, the context in which it arose was presented and its importance was explained.

The overall aim of identifying and discussing these lessons is to share the gained insights with developers who are engaged in developing innovative Web-based collaboration and decision making support services, in order to better structure and plan their work, lessen their development time and deploy such services in a meaningful way in today’s data-intensive and cognitively-complex settings.
4 References


Appendix: Features of CommBAT

The Dicode Collaborative Workspace supports knowledge workers on collaboration, sensemaking and decision making in data-intensive tasks. Specifically, the collaborative workspace supports users in working with their argumentation via different views, e.g. ‘forum view’, ‘mind-map view’ and ‘formal view’, which permit an incremental formalization of collaboration. This collaborative approach to problem solving involves a process of brain storming, sensemaking and decision making. To understand this process, we are interested in:

- How does a group make sense a problem together in a collaborative workspace?
- What kind of sensemaking behaviours are involved in a collaborative argumentation space?
- Can we discover any patterns which might lead us to automatically facilitate the process from brain storming to decision making?

Driven by the above research questions, CommBAT is developed to support users to explore the log data of Dicode Collaborative Workspace. After a few iterations of development, CommBAT has following features:

1. It provides a quick overview of a Dicode Collaborative Workspace in terms of users, activities and knowledge objects in a visual form (Figure 1). CommBAT presents collaborative workspace factors (Users, Sensemaking, Activities, Object types, Days of Activities and Knowledge Objects) in different visual forms: Pie Chart, Column Chart, Bar Chart, Bubble Chart etc. to allow users to gain a quick overview about the statistics of those factors. For example, in the workspace shown in Figure 1 (ID: 27478), there were 3 types of sensemaking activities: contributing, reviewing and organizing. The Pie Chart of sensemaking activities shows that about 75% of activities were “organizing knowledge objects” (e.g. moving items in the ‘mind-map view’).

![Figure 1: CommBAT overview - one collaboration workspace.](image)

2. It allows users to explore different aspects of sensemaking behaviour in an interactive way (Figure 2). CommBAT presents the factors of Users, Sensemaking, Activities, Object types and Days of Activities in a sorted order and in two views: List view and
Chart view. List view allows users to filter a particular item and get the statistics for that item. For example, in Figure 2, at the top, the factor Object type “Ideas” is filtered; at the bottom, the Sensemaking activity “Reviewing” is filtered. Statistics results of all other factors are presented according to the filtered item instantly.

Figure 2: CommBAT – Comparing different factors of sensemaking (top: records related to Object Type: “Ideas”; bottom: records related to Sensemaking Activity: “Reviewing”).

3. It can help users to discover patterns by comparing the different views of Knowledge Object: Object Type view, User Activity view and Activity Timeline view (Figure 3). CommBAT presents the most important factor of the workspace, Knowledge Objects, in three different views. Object Type view presents the semantic type of object in different shapes and colours in an order of their creation. It shows how the knowledge items are created in terms of semantic types. User Activity view presents all objects with a statistics of user activities. It shows how many users have contributed to the activities of the objects. Activity Timeline view presents the sensemaking activities of objects over time. It can show the time span of an object with different sensemaking activities.

The different views of knowledge items can also help users to examine the details of a particular Knowledge Object through the three different Chart views (Figure 3). The
details of a particular item on these three views can be retrieved by clicking the item on the Chart. The detail view of an object shows three kinds of information: all the logs of that object, the relation object (lines) that linked to this object, and the related objects that are linked in (inbound) or linked out (outbound) objects shown in a tree view.

Figure 3: Object Detail (top: Object Type view with Object Logs; middle: User Activity view with Object Relations; bottom: Activity Timeline view with Object Logs; all three figures show the details of object: 27481).

4. It can help users to quickly compare the community behaviours among different workspaces via Multiple Document Interfaces (MDI). CommBAT supports MDI, which allows users open multiple log files of workspaces and compare them in one
interface. Figure 4 illustrates examples of two workspaces opened at the same time: the top part of the figure shows a comparison between a workspace of Use Case 1 and one of Use Case 2, while the bottom part a comparison between a workspace of Use Case 2 and another one of Use Case 3.

Figure 4: CommBAT: Comparison of two collaboration workspaces (top: Dicode Use Case 1 (BRF: 24292) & Dicode Use Case 2 (IMA: 27478); bottom: Dicode Use Case 1 (BRF: 27231) & Dicode Use Case 3 (PUB: 28232).