

Combining and Uniting Business Intelligence with Semantic Technologies

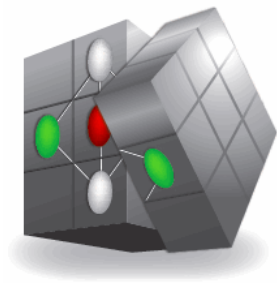
Acronym: CUBIST

Project No: 257403

Small or Medium-scale Focused Research Project

FP7-ICT-2009-5

Duration: 2010/10/01-2013/09/30



cubist

Your Business Intelligence

Generalized Requirements

Abstract: Based on the requirements from the use cases, which are provided in D7.1.1, D8.1.1 and D9.1.1, this report provides a use-case-independent, aggregated and generalized list of requirements for CUBIST.

Type	Report
Document ID:	CUBIST D1.1.2.
Workpackage:	WP1
Leading partner:	SAP
Author(s):	Frithjof Dau (SAP)
Dissemination level:	PU
Status:	Final Version
Date:	06 April 2011
Version:	1.0

Versioning and contribution history

Version	1. Description	Contributors
0.1	Draft with textual description of formal requirements	Frithjof Dau (SAP)
0.2	1) Corrections of textual description of reqs by HWU 2) introduction written 3) revised and extended FCA-analysis	Frithjof Dau (SAP)
0.3	1) slight changes in textual description of reqs 2) added formal, generalized requirements 3) added in appendix introduction into the reading of the FCA-diagrams	Frithjof Dau (SAP)
0.9	Reviews addressed: 1) Typos fixed 2) Changed nested diagrams in chapter 2 3) Extended introduction into reading diagrams in appendix 4) Added chapter with final remarks	Frithjof Dau (SAP)
1.0	Final minor corrections	Frithjof Dau (SAP)

Reviewers

Name	Affiliation
Kenneth McLeod	HWU
Simon Andrews	SHU

Table of contents

TABLE OF CONTENTS	3
1 INTRODUCTION	4
2 AN FCA-BASED EVALUATION OF THE FORMAL REQUIREMENTS	6
2.1 BY PRIORITY	6
2.2 BY PARTNER.....	6
2.3 BY REQUIREMENT TYPE.....	6
2.4 BY AFFECTED WORK PACKAGE	8
2.5 BY PARTNER AND AFFECTED WORK PACKAGE.....	9
2.6 BY PARTNER AND REQUIREMENT TYPE.....	9
2.7 REQUIREMENT TYPE AND WORK PACKAGE.....	10
3 TEXTUAL SUMMARY OF THE FORMAL REQUIREMENTS	12
3.1 REQUIREMENTS CONCERNING THE CUBIST ARCHITECTURE.....	12
3.2 STORAGE AND DATA REQUIREMENTS	12
3.3 ANALYTICAL FEATURES:.....	13
3.4 VISUAL DISPLAY OF DATA	14
3.5 USABILITY AND FRONTEND	15
4 LIST OF GENERALIZED REQUIREMENTS	16
5 CONCLUDING REMARKS	21
6 APPENDIX 1: INTRODUCTION INTO READING FCA DIAGRAMS	22
7 APPENDIX 2: COMPLETE LIST OF FORMAL REQUIREMENTS	26
7.1 PURPOSE OF THE PROJECT	27
7.2 MANDATED CONSTRAINTS	27
7.3 FUNCTIONAL REQUIREMENTS.....	28
7.4 DATA REQUIREMENTS	34
7.5 LOOK AND FEEL	35
7.6 USABILITY AND HUMANITY REQUIREMENT	36
7.7 PERFORMANCE REQUIREMENT	37
7.8 OPERATIONAL REQUIREMENT	37
7.9 MAINTAINABILITY AND SUPPORT REQUIREMENT.....	38
7.10 SECURITY REQUIREMENTS.....	39
7.11 LEGAL REQUIREMENTS.....	40

1 Introduction

The first phase of the CUBIST project is dedicated to its requirement analysis.

Different means for gathering and recording requirements are used in CUBIST. These means have been described in D1.1.1: “Directives for the Requirement Analysis”. Essentially, they are:

- 1) Personas
- 2) Utilization Scenarios
- 3) Formal Requirements
- 4) MockUps

Additionally Transactional Use Case Diagrams have been introduced to the use case partners as a further way of describing transactional-based business processes in their application scenarios.

For a detailed description of the means, we refer to D1.1.1. An overview of the means and their interdependencies is given in the diagram below.

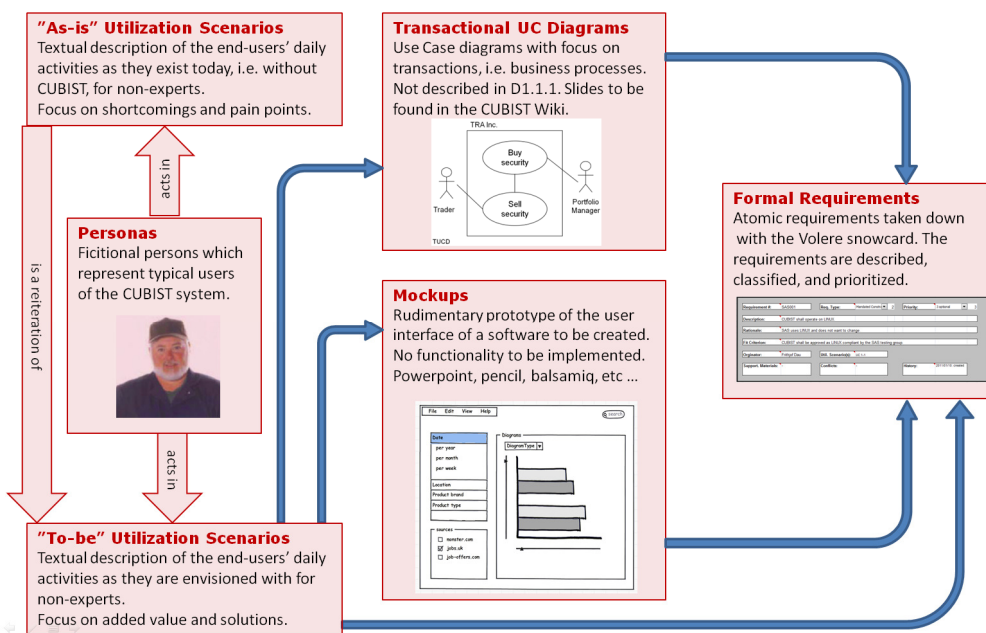


Fig 1: Overview over the requirement analysis means in CUBIST

Each use case partner has conducted a use-case-specific requirement analysis, in which utilization scenarios, personas and formal requirements are provided in use-case-dependent deliverables Dx.1.1 “Requirement Document“ (x=7,8,9). For the time being, the mock-ups are still to be created. Once finished, they will be provided in use-case-dependent deliverables Dx.1.2 “Initial Mockup” (x=7,8,9).

The three requirement documents Dx.1.1 are the basis for this deliverable, which analyzes the use-case-specific requirements and generalizes them to use-case-independent requirements.¹

Due to its importance, a paragraph from D1.1.1 shall be reiterated now:

Personas and utilization scenarios mainly serve to provide the CUBIST technological partners a better understanding of the use-cases. Strictly speaking, neither personas nor utilization scenarios are requirements; nonetheless, they serve as an input for the formal requirements.² In fact, in contrast to personas and utilization scenarios, formal requirements can be provided not only by use-case-partners, but by technological partners as well. In this document, we thus focus on the formal requirements.

In the last months, more than 100 formal requirements have been gathered from the use case partners and from SAP. The initial list of requirements has already been discussed in several telephone calls amongst the consortium, leading to some requirements being refined, re-prioritized, or dropped. Even though the initial list has already been refined, a paragraph from D1.1.1 shall be reiterated:

The requirements collected so far shall be understood to be **preliminary**. They are important for an initial, project-wide understanding of the aim of CUBIST, but they do not form a complete and final list. During the project, due to the research character of CUBIST and based on the agile method, requirements will be refined, added, or removed. It is planned that later in the project, requirements will be maintained in a backlog based on Bug Track, and regular calls in the “hot” development phase will be utilized for refining the requirements.

Comment [KM1]: Now mirrors style of quote above

In the following chapters, first a FCA-based analysis of the requirements is conducted. This analysis utilizes different attributes (like priority, requirement type or originating partner) of the requirements, but does not investigate the requirements contents. This is done in the subsequent chapter, where a textual summary of the requirements is given. Finally, a use-case independent list of requirements is provided.

¹ Similarly for the mock-ups, there will a deliverable “D1.1.3 Mockup” which generalizes the use-case-dependent mock-ups provided in Dx.1.2 to a use-case-independent mockup.

² A good example are the two (usability and humanity) formal requirements HWU069 “Visualisations should be detailed and convey large amounts of information” and HWU070 (“Visualisations should be simple”). These two requirements are at a first glance contradicting. Anyway, as it can be seen from the detailed requirements (see appendix), these two requirements serve two different Personas (namely biologists and computational scientist) with a different background in computer science and a different attitude towards interfaces and visualisations.

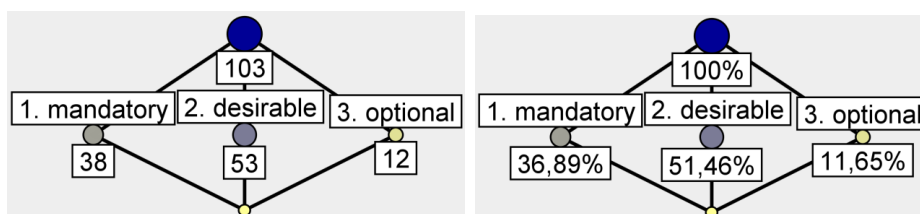
2 An FCA-based Evaluation of the Formal Requirements

In this chapter, an analysis of the formal requirements along some dimensions provided by the VOLERE snowcard, as altered for CUBIST, is conducted.

As CUBIST will employ FCA for conducting data analysis, it is self-suggesting to use FCA for evaluating the requirements in CUBIST (sloppily speaking, we are “eating our own dog food” in CUBIST). A very short introduction to reading the diagrams in this chapter is provided in the appendix.

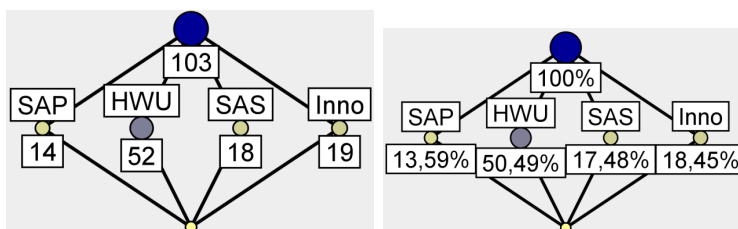
2.1 By Priority

In the first diagrams, the total list of all 103 requirements is clustered according to their priorities.



2.2 By Partner

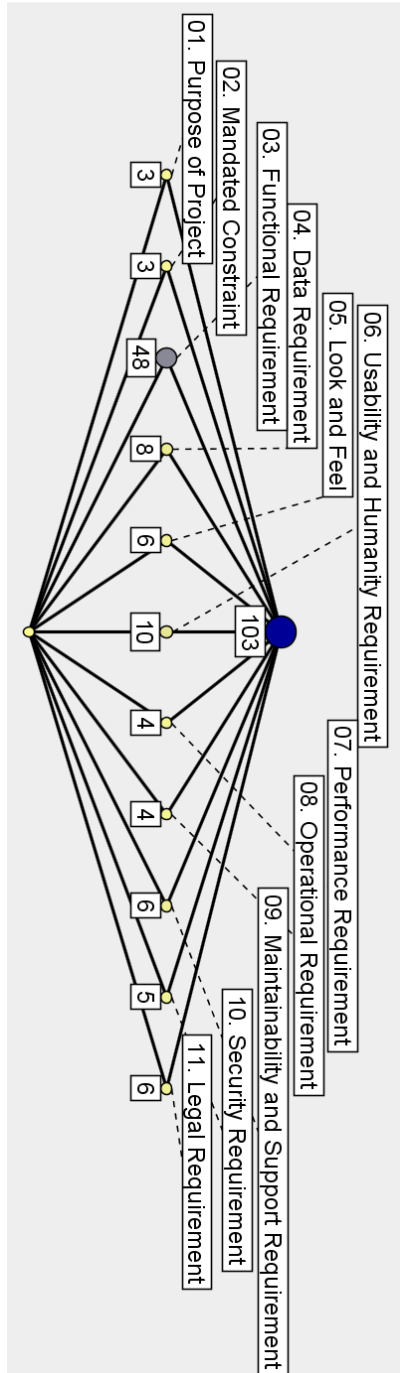
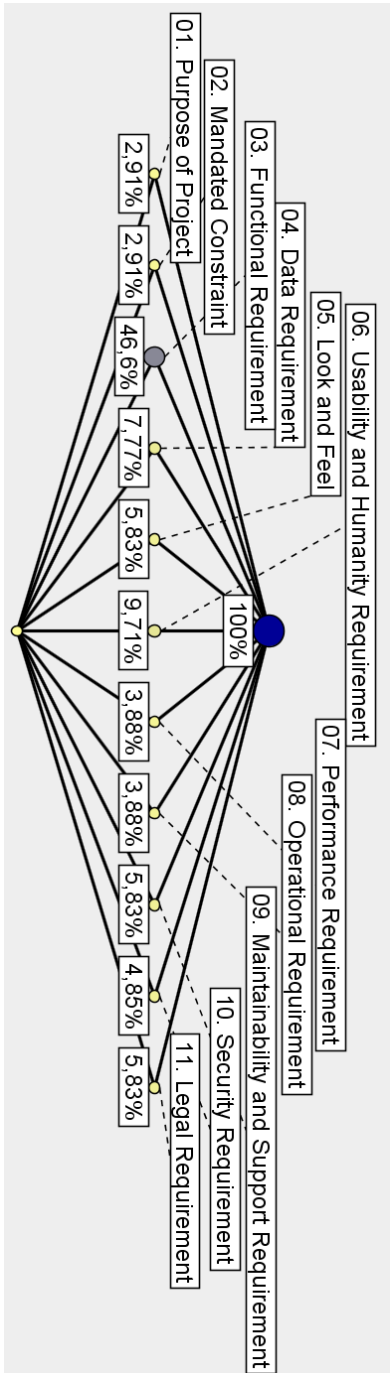
In the next diagrams, the requirements are clustered according to the CUBIST partners raising them. Note that HWU contributed more than half of the requirements.



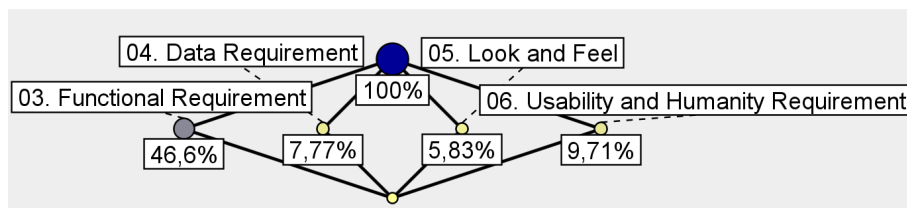
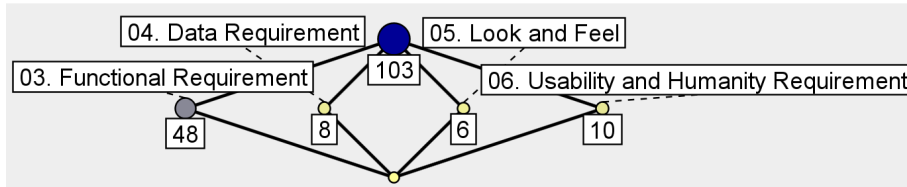
2.3 By Requirement Type

The next diagrams provide an overview over the requirements according to the requirement type. There are eleven different types, leading to a quite large diagram. It can be easily seen that functional requirements form the most prominent type, as they make nearly half of all requirements. They are followed by usability and humanity requirements, which are likely to be that prominent due the fact that CUBIST focuses on **user-friendly** BI and visual analytics. Though not that prominent (in terms of amount), it can be argued due to the focal points of

CUBIST that both data and look and feel requirements are important as well. For this reason, two more diagram focus on the four most important requirement types.

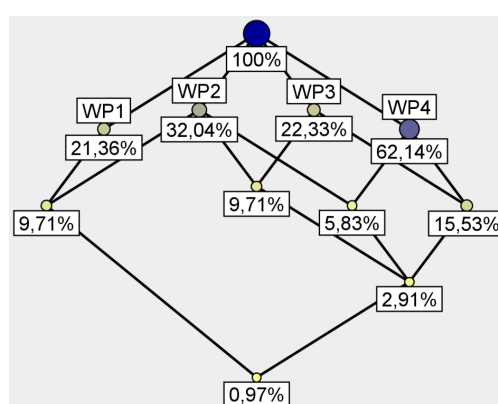
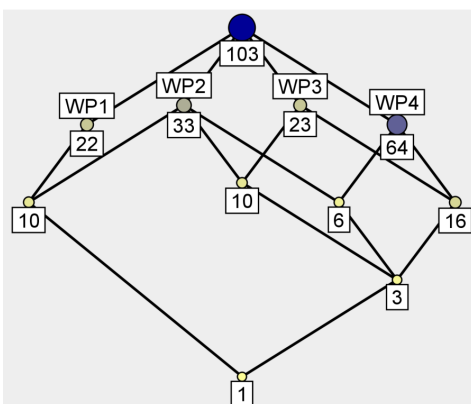


Due to the large number of requirement types, these diagrams are slightly too large. Next, two diagrams with a restricted set of the four most important types are provided.



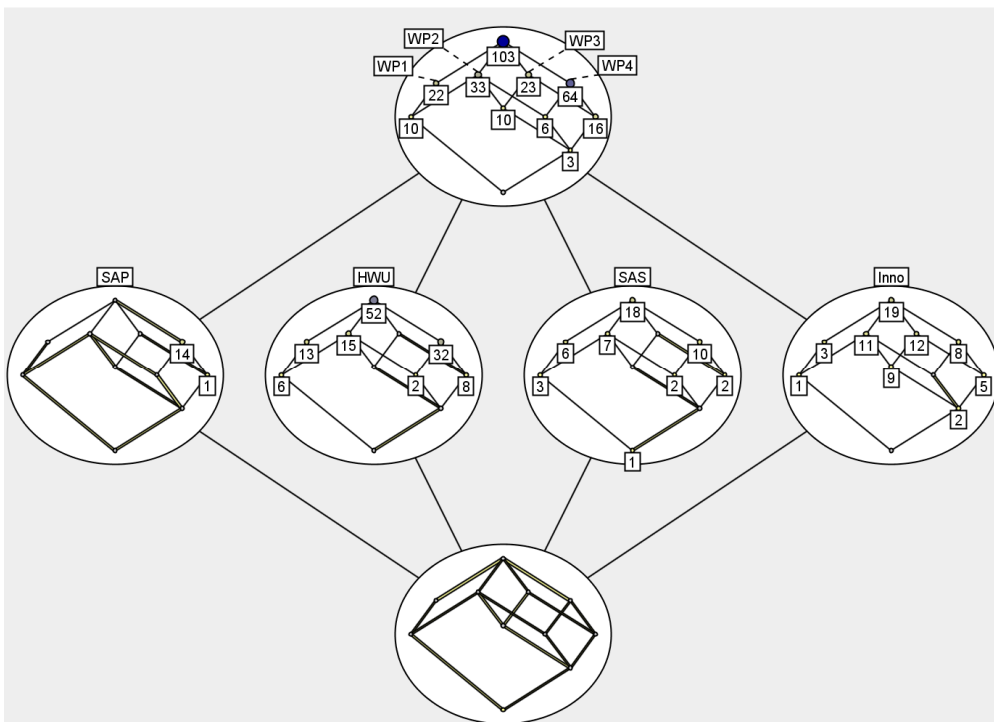
2.4 By Affected Work Package

For each technical work package, all requirements have been classified according to the WP they significantly affect. Multiple assignments are permitted. The next diagrams provide an overview of these assignments. These diagrams are the first diagrams revealing structural insights into the requirements. For example, besides one requirement which affects all WPs (this is the requirement SAS013: “CUBIST users shall be able to have complete control over the data managed by CUBIST software and the CUBIST software as such”), requirements concerning the overall architecture of CUBIST (WP1: “Technological Architecture”) might affect WP2 (“Semantic ETL and Data Integration”), but not WP3 (“Semantic Data Warehouse”) and WP4 (“Analyzing and Visualizing Data”). On the other hand, for WP2-WP4, for each combination of these three WPs (like “WP2 and WP3 and not WP4” or “not WP2 neither WP4 but WP3”) some requirements affect exactly this combination. Note moreover that more than half of the requirements concern WP4 (and possibly other WPs).



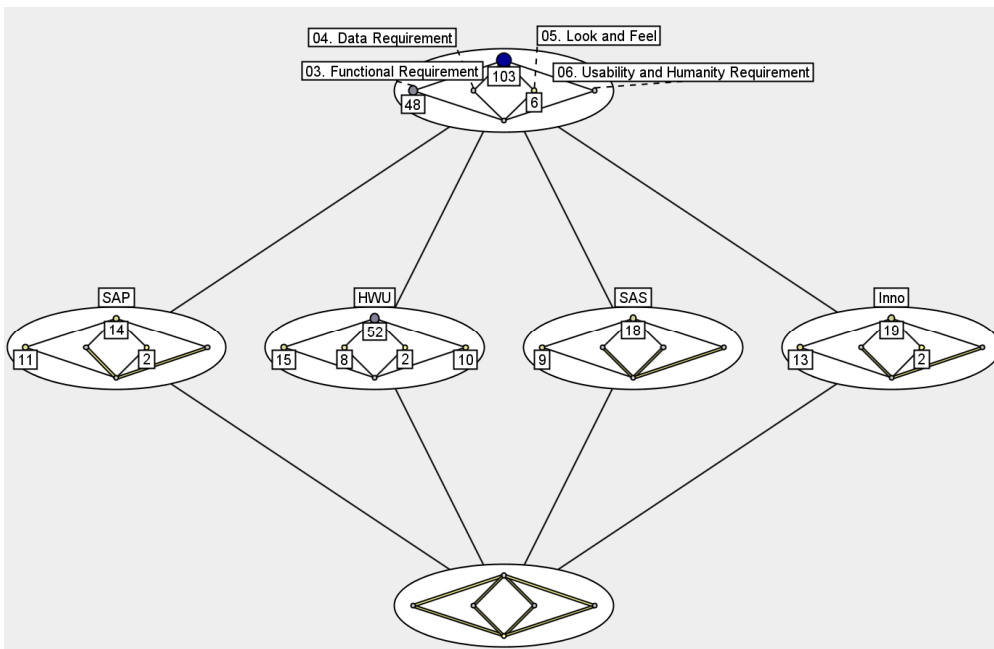
2.5 By Partner and Affected Work Package

Now we start to combine some of the diagrams. The next diagram combines as attributes the partners who raised requirements with the affected WPs. Note that SAP, being the only technological partner who contributed formal requirements, focused on requirements for WP4 (and one requirement concerning both WP3 and WP4), whereas the use case partners have requirements distributed over all work packages.



2.6 By Partner and Requirement Type

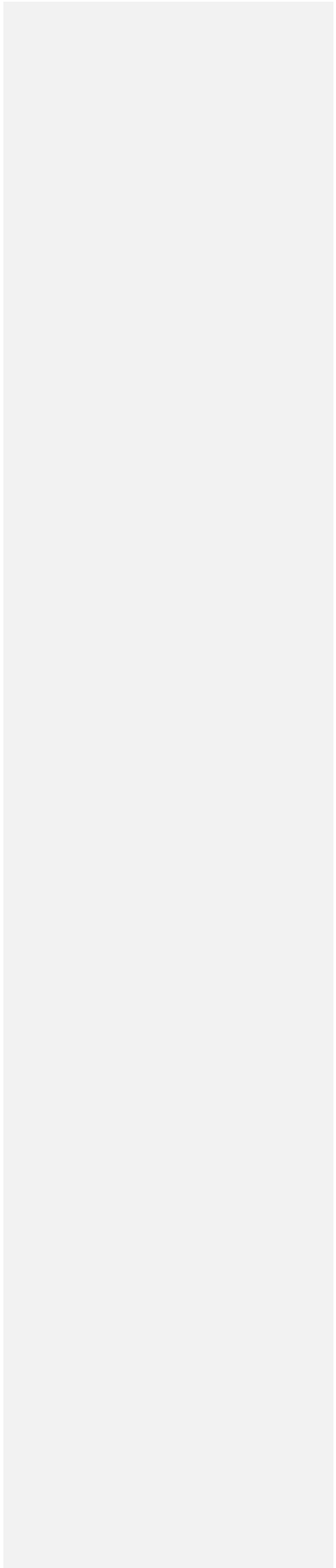
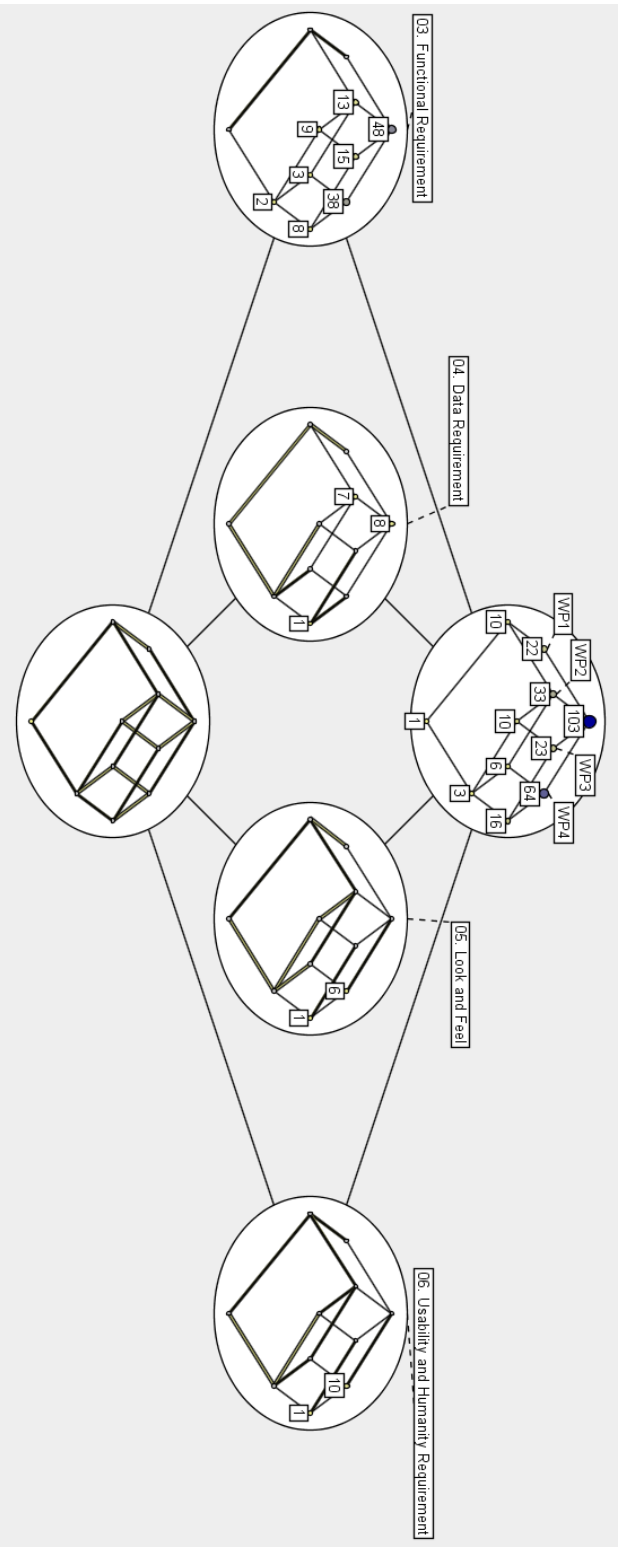
The next diagram combines as attributes the partners who raised requirements with the requirement types. Note that only HWU spans over all of the four most important requirement types, whereas other partners focus (amongst the four most important requirement types) on functional requirements and provide few look and feel requirements.



2.7 Requirement Type and Work Package

The next diagram combines the work packages with the requirement types. Note that:

- 1) There is no functional requirement which affects WP1 (“Technological Architecture”).
- 2) Unsurprisingly, nearly all data requirements affect WP2 (“Semantic ETL and Data Integration”). The one requirement which differs is HWU043 “Any inconsistencies / errors identified should be reported to EMAGE curators”, which is a requirement for the data, but affects WP3 and WP4.
- 3) Both look and feel requirements and usability and humanity requirements mainly affect the frontend. So it comes as no surprise that nearly all of these requirements affect WP4 exclusively, with only two of them affecting the underlying WP3 as well.



3 Textual Summary of the Formal Requirements

This chapter provides a more high level view of the requirements in a textual format.

3.1 Requirements Concerning the CUBIST Architecture

There are few requirements affecting the overall architecture for CUBIST.

- **Web-based:** For HWU and Innovantage, it is important that CUBIST is not restricted to Windows systems and does not require any installation, instead CUBIST - including a maintenance tool - must run in a web browser (HWU010, HWU011, HWU090, HWU091, HWU102, INN012, INN013).
- **Installation-based:** For SAS, it is fine when CUBIST has to be installed locally or on a server (SAS012).
- **Open source:** For HWU, developing CUBIST as open source is desirable (HWU121).

In these initial requirements, two use case partners (HWU, Innovantage) request a web-based architecture, whereas one use case partner (SAS) prefers a locally or server based installed software. This issue has been discussed in telephone conferences conducted during the requirement analysis phases, and SAS have agreed to a web-based implementation of CUBIST. The requirement SAS012 is left in the list, as its priority is not mandatory, but desirable. Moreover, it has been agreed that CUBIST should run in the most important browsers, but supporting *all* browsers is outside the realm of CUBIST (see INN013).

The open source requirement from HWU has been discussed in the CUBIST consortium as well. It is unlikely that the complete CUBIST solution will be developed as open source, but some parts of it may be open source. The requirement HWU121 is left in the list, as its priority is only desirable.

3.2 Storage and Data Requirements

This section deals with requirements concerning the data sources to be federated and the persistence layer of CUBIST.

- **Using existing data sources:** Of course, the respective data from the use case partners has to be used, e.g. data from SAS (SAS001), HWU sources (HWU041, HWU042), some of them optional (HWU046, HWU047).
Particularly for Innovantage, external, web-based, unstructured sources like LinkedIn, Bloomberg, Financial times, twitter, facebook etc are most important. Here, RSS feeds are to be utilized (INN010) or text mining and information extraction has to be used to extract core facts from different sources, including sentiments, and match information from different sources (INN004, INN006, INN007, INN008, INN009, INN001, INN011, INN021).
- **Persistence of textual annotations:** Found textual annotation must be persisted (HWU040) and propagated to frontend (HWU021).
- **Crediting information providers:** The information providers have to be credited (HWU120, SAS019), and sometimes backlinks to pages which provide more information of the retrieved data are needed (HWU033).

Comment [KM2]: Is this a new bullet point? Or should it be on the previous line, ie no carriage return
FD: No new bullet point, as this still refers to existing data sources.

- **Data usage and access:** It is important that the data usage is not restricted by CUBIST licensing agreements; instead, CUBIST users must have complete control over the data managed by CUBIST, and the data must not leave the enterprise environment (SAS013, SAS018, SAS015). Particularly, storing data in the cloud should be avoided (SAS016). Only system administrators and developers are allowed to change the data in the triple store (HWU110, INN014), and the original data sources are not to be changed by CUBIST (HWU111).
- **Extending existing data sources:** Extending existing datasources with public accessible SPARQL endpoints is a relevant side outcome of CUBIST for HWU (HWU122, HWU123).
- **Triple store for persisting:** As already stated in the DoW, the data is to be persisted in a triple store, which provides a SPARQL interface (SAS005, INN003).
- **Updating data:** The data is not to be persisted once, but has regularly to be updated (HWU044, SAS011), e.g. with pipelines (HWU100, HWU101).

3.3 Analytical Features:

Unsurprisingly, some of the formal requirements request that CUBIST provides analytical features (HWU001, SAS002, SAS006, INN001). In the following, a summary of the requested kinds of analytics is provided. Requirements dedicated more to the *visual* aspects of analytics are provided in a following section.

- **Detecting similarities and patterns:** Detecting similarities and patterns of entities (HWU027, HWU028, SAS003), and optionally computing implications and association rules (SAP012) shall be possible with CUBIST. By mutual agreement in CUBIST calls, this is to be understood to be done by core facilities of FCA (e.g. it is likely that CUBIST will not employ additional similarity measures for entities). Optionally, CUBIST shall facilitate the comparison of different results (HWU039).
- **Detecting inconsistencies/errors:** CUBIST shall be able to detect inconsistencies / errors in the data (HWU022, HWU043) and to suggest corrections. This is essentially done by a human using visual inspection of the visuals provided by CUBIST. Dealing with corrections is error-prone, thus recall the Data usage and access related requirements (see the previous section) which state that CUBIST does not change the underlying data set.
- **Spatial analytics:** CUBIST will have to deal with two different kinds of spatial annotations and queries, namely spatial annotations for mouse embryos in the HWU use case, and geo-spatial annotations for job offers in the Innovantage use case. Spatial annotations for mouse embryos must be persisted and analyzable (HWU035, HWU045, HWU046, HWU047). A spatial representation for job offers which is queryable via the radius is important for Innovantage (INN005).
- **Time-related analytics and detecting trends:** Another important dimension is time. Detecting trend information for different entities is important for HWU (HWU023, HWU024, HWU025). Though time-related requirements are only raised by HWU, SAS and Innovantage agreed in internal calls that time-related analytics is important for them as well.

Next, some analytics related requirements are provided.

- **Filtering results:** It is mandatory that CUBIST allows for filtering of results in order to restrict the analysis to a subset of all data (HWU031). Concerning the visual analytics, filtering is a a-priori reduction of the data: it is conducted before the data is visually displayed. This is different to the zooming facilities required in the visual display requirements.
- **Search functionalities:** Besides analytical features, some search functionalities are needed as well. The data in the triple store can be searched. Examples are searching by location (HWU035), co-expressions (HWU037), and genes (HWU034), and searching by job vacancies matching several criteria (INN018). A search interface in the start pane is desirable (INN017). It should be stressed that it was agreed in the consortium that only basic search functionalities will be developed in CUBIST; sophisticated search facilities lie outside the realm of CUBIST. Finally, in a concept lattice, a full text search in the lattice (SAP001) is needed as well.
- **Storage and loading:** It is desirable that CUBIST provides means to store queries and results (HWU030, SAP013, SAP014).

3.4 Visual display of data

A key aspect of CUBIST which is reflected by the formal requirements is CUBIST's visual representation of the data (SAS008, SAS004, INN002). In the following, requirements focusing on visual aspects of the analytics are provided.

- **Interactive and lively visual representations:** The visual representations must be interactive and lively (INN016, SAS009). For example, it must be possible select parts of the displayed lattice to be redisplayed, or to zoom into concept nodes (compare this to the filtering facilities in the last section) (SAP007, SAP008). Different visualizations can be chosen (HWU032) and can be combined and nested (SAP005). Particularly for Hasse-diagrams, CUBIST offers different layout algorithms (SAP011). As no layout algorithm is perfect, a manual correction of the layout must be possible (SAP009).
- **Amount of information to be displayed:** The interface and the visualizations should be simple (HWU068, HWU070), but on the other hand, sometimes comprehensive amounts of information have to be displayed to the user (HWU069). This is due to different technical backgrounds of different personas. Thus different levels of detail are needed.
- **Conceptual scales:** For a specific approach, namely conceptual scales, scales can be either predefined or created on the fly (SAP002, SAP003), and scales are suggested by CUBIST (SAP010).

3.5 Usability and Frontend

As argued in the CUBIST DoW, one of the greatest barriers for BI is the complexity of the user interfaces. So it comes as no surprise that some requirements focus on the simplicity and usability of the CUBIST frontend. This section deals with the usability of CUBIST.

- **Ease of use:** Particularly for non-technical users, CUBIST must be easy to use. A simple frontend and visualizations are needed (HWU068, HWU070). In the best case, CUBIST can be used after watching the screencast once (HWU064, HWU065).
- **Transparency:** CUBIST might reveal some surprising functionalities, e.g. with respect to detecting inconsistencies. For this reason, it is important that users always understand what CUBIST is actually doing, that is CUBIST must constantly provide feedback (HWU061, HWU062).
- **Language:** SAS wants no domain-specific terminology (SAS007³). Optionally, HWU wants to use biological metaphors (HWU060).
- **Frontend:** The frontend should be clean and uncluttered (HWU051), should use web metaphors, or optionally even look like existing webapps (HWU052, HWU066). The CUBIST frontend should be customizable (SAP006).
- **Documentation:** It is desirable that CUBIST is fully documented (SAS014, HWU103). However the consortium agreed that full documentation might take too many resources to be implement.
- **Performance:** The performance of CUBIST affects its usability. Four performance requirements provide concrete numbers for measuring the performance, e.g. a concept lattice for a context with 20 attributes is computed in less than 2 seconds, and most queries must have results displayed to user in less than 10 seconds (SAP004, HWU080, INN015, HWU081). Moreover, two requirements state that it is desirable that users are made aware of the progress for tasks (SAS010, HWU067). It should be noted that due to the research project nature of CUBIST, enterprise-level scalability is not a core requirement for CUBIST.

³ SAS007 mentions “the text strings deployed in the software”, so a possible solution is to maintain these strings in a separate file. must be maintainable

4 List of Generalized Requirements

Based on the textual summary of the requirements, this chapter provides a list of formal, generalized requirements for CUBIST.

section in textual description	Req. #:	Req. Type:	Priority:	Description:	Derived from:	WP1	WP2	WP3	WP4
01 web based	GEN001	02. Mandated Constraint	1. mandatory	CUBIST must be web based	HWU010, INN012	x	.	.	.
01 web based	GEN002	08. Operational Requirement	2. desirable	Works and looks similar with major browsers (with flash support, no other plugins required). No mobile browsers targeted	INN013, HWU011, HWU091	x	.	.	.
01 web based	GEN003	09. Maintainability and Support Requirement	2. desirable	Maintenance tool should be online	HWU102	x	.	.	.
01 web based	GEN004	08. Operational Requirement	1. mandatory	Must not rely on windows technology		x	.	.	.
02 installation based	GEN005	08. Operational Requirement	2. desirable	The installation procedure shall be scriptable and allow for automation	SAS012	x	.	.	.
03 open source	GEN006	11. Legal Requirement	2. desirable	CUBIST software should be open source to allow it to be used/extended after project ends	SAS012	x	.	.	.
04 using existing data sources	GEN007	04. Data Requirement	1. mandatory	Repository should include data sets from use case partners	HWU041, HWU042, HWU046, HWU047, SAS001	.	x	.	.
04 using existing data sources	GEN008	03. Functional Requirement	1. mandatory	Rule based extraction and text mining of facts from text documents	INN004, INN008, INN009	.	x	.	.
04 using existing data sources	GEN009	03. Functional Requirement	2. desirable	categorisation of text documents	INN006	.	x	x	.
04 using existing data sources	GEN010	03. Functional Requirement	2. desirable	Extraction of sentiment phrases of text documents	INN007	.	x	x	.
04 using existing data sources	GEN011	03. Functional Requirement	2. desirable	Ability to subscribe to RSS feeds	INN010	.	x	x	.
04 using existing data sources	GEN012	03. Functional Requirement	2. desirable	Learning from user feedback	INN011	.	x	x	.
04 using existing data sources	GEN013	03. Functional Requirement	2. desirable	matching entities extracted from texts with entities in the ontology	INN021	.	x	x	x
05 persistence of textual annotations	GEN014	04. Data Requirement	1. mandatory	CUBIST warehouse must include textual annotations (unrelated to FCA viz)	HWU021, HWU040	.	x	.	.

section in textual description	Req. #:	Req. Type:	Priority:	Description:	Derived from:	WP1	WP2	WP3	WP4
06 crediting information providers	GEN015	03. Functional Requirement	2. desirable	Link back to original sources	HWU033	.	x	.	x
06 crediting information providers	GEN016	11. Legal Requirement	1. mandatory	Information provider must be credited (anywhere on the screen)	HWU120, SAS019	.	x	.	x
07 data usage and access	GEN017	09. Maintainability and Support Requirement	1. mandatory	CUBIST users shall be able to have complete control over the data managed by CUBIST software and the CUBIST software as such	SAS013	x	x	x	x
07 data usage and access	GEN018	10. Security Requirement	1. mandatory	Only system admin & CUBIST developers have write access to data	HWU110, INN014	x	x	.	.
07 data usage and access	GEN019	10. Security Requirement	1. mandatory	CUBIST should not attempt to change original data source	HWU111	x	x	.	.
07 data usage and access	GEN020	10. Security Requirement	1. mandatory	CUBIST software shall allow the user to host all the data locally (not locally on the PC; but in the environment)	SAS015	x	x	.	.
07 data usage and access	GEN021	10. Security Requirement	1. mandatory	Public cloud services shall not be used to store and process space control centre data unless complete confidentiality of such data is assured.	SAS016	x	.	.	.
07 data usage and access	GEN022	11. Legal Requirement	1. mandatory	The licensing agreement for CUBIST software shall not put any restriction on data use.	SAS018	x	x	.	.
08 extending existing data sources	GEN023	11. Legal Requirement	2. desirable	Some data sources provide an open sparql endpoint	HWU122, HWU123	x	x	.	.
09 triple store for persisting	GEN024	03. Functional Requirement	1. mandatory	CUBIST software shall store existing data in an RDF store and provide an extended SPARQL query interface to the data.	INN003, SAS005	.	x	x	.
10 updating data	GEN025	03. Functional Requirement	2. desirable	Warehouse should be updated inline with regular data updates	HWU044, SAS011	.	x	.	.
10 updating data	GEN026	09. Maintainability and Support Requirement	1. mandatory	Pipeline are implemented which automatically import new (only) and amended data	HWU100, HWU101	x	x	.	.

section in textual description	Req. #:	Req. Type:	Priority:	Description:	Derived from:	WP1	WP2	WP3	WP4
11 analytical features	GEN027	01. Purpose of Project	1. mandatory	CUBIST provides analytical, FCA-based features on top of existing data	HWU001, INN001, SAS002, SAS006	.	.	.	x
12 detecting similarites and patterns	GEN028	03. Functional Requirement	2. desirable	UBIST allows to identify patterns in entities	HWU027, HWU028, SAS003	.	.	.	x
12 detecting similarites and patterns	GEN029	03. Functional Requirement	3. optional	CUBIST allows to compare multiple experiments	HWU039	.	.	.	x
12 detecting similarites and patterns	GEN030	03. Functional Requirement	3. optional	For a given context, CUBIST can compute implications and association rules	SAP012	.	.	.	x
13 detecting invonsistencies/errors	GEN031	03. Functional Requirement	2. desirable	CUBIST allows to locate errors/inconsistencies in underlying data	HWU022	.	.	.	x
13 detecting invonsistencies/errors	GEN032	04. Data Requirement	3. optional	Any inconsistencies/errors identified should be reported to specific persons	HWU043	.	.	x	x
14 spatial analytics	GEN033	04. Data Requirement	2. desirable	Repository should include spatial annotations	HWU045, HWU046, HWU047	.	x	.	.
14 spatial analytics	GEN034	03. Functional Requirement	2. desirable	CUBIST provides specific means (like radius search) for exploring spatial annotations	INN005, HWU035	.	x	x	x
15 time-related analytics and detecting trends	GEN035	03. Functional Requirement	2. desirable	CUBIST allows to do time-related analytics and to detect trends	HWU023, HWU024, HWU025	.	.	x	x
16 filtering results	GEN036	03. Functional Requirement	1. mandatory	CUBIST allows to filter entities (before lattice is computed)	HWU031	.	.	x	x
17 search functionalities	GEN037	03. Functional Requirement	2. desirable	CUBIST provides simple search functionalities for searching the data in the triple store	INN017	.	.	.	x
17 search functionalities	GEN038	03. Functional Requirement	3. optional	CUBIST provides extended search functionalities for searching the data in the triple store	INN018, HWU034, HWU035, HWU037	.	.	x	x
17 search functionalities	GEN039	03. Functional Requirement	2. desirable	Full text search on displayed concept lattices	SAP001	.	.	.	x
18 storage and loading	GEN040	03. Functional Requirement	2. desirable	Queries & results may be saved & reloaded later	HWU030	.	.	.	x
18 storage and loading	GEN041	03. Functional Requirement	2. desirable	visualizations can be stored and reloaded	SAP013	.	.	.	x

section in textual description	Req. #:	Req. Type:	Priority:	Description:	Derived from:	WP1	WP2	WP3	WP4
18 storage and loading	GEN042	03. Functional Requirement	2. desirable	visualizations can be exported	SAP014	.	.	.	x
19 visual display of data	GEN043	03. Functional Requirement	1. mandatory	CUBIST software shall produce visual, FCA-based representations of data	INN002, SAS004, SAS008	.	.	.	x
20 interactive and lively visual representations	GEN044	03. Functional Requirement	2. desirable	Flexible, lively and interactive UI, particularly presentation of results	HWU032, INN016, SAS009	.	.	.	x
20 interactive and lively visual representations	GEN045	03. Functional Requirement	1. mandatory	different visualizations can be combined and "nested"	SAP005	.	.	.	x
20 interactive and lively visual representations	GEN046	03. Functional Requirement	1. mandatory	it is possible to zoom into a concept node or to select a part of the overall lattice to be displayed	SAP007, SAP008	.	.	.	x
20 interactive and lively visual representations	GEN047	05. Look and Feel	1. mandatory	it is possible to manually move nodes in the display of the concept lattice	SAP009	.	.	.	x
20 interactive and lively visual representations	GEN048	03. Functional Requirement	2. desirable	CUBIST offers different layout algorithms for the layout of a Hasse diagram	SAP011	.	.	.	x
21 amount of information to be displayed	GEN049	06. Usability and Humanity Requirement	2. desirable	visualisations should be simple and display only small amounts of data (for unskilled persons)	HWU068, HWU070	.	.	.	x
21 amount of information to be displayed	GEN050	06. Usability and Humanity Requirement	2. desirable	Visualisations should be detailed and convey large amounts of information (for power users)		.	.	.	x
22 conceptual scales	GEN051	03. Functional Requirement	2. desirable	For a given attribute, CUBIST can generate a conceptual scale on the fly	SAP002	.	.	.	x
22 conceptual scales	GEN052	03. Functional Requirement	2. desirable	For some attributes, conceptual scales can be predefined	SAP003	.	.	.	x
22 conceptual scales	GEN053	03. Functional Requirement	3. optional	CUBIST suggests scales and visualizations for given attributes	SAP010	.	.	.	x
23 ease of use	GEN054	06. Usability and Humanity Requirement	2. desirable	watching the screencast one time is sufficient for using CUBIST	HWU064, HWU065	.	.	.	x
23 ease of use	GEN055	06. Usability and Humanity Requirement	2. desirable	Interface and visualisations should be simple (for unskilled persons)	HWU068, HWU070	.	.	.	x
24 transparency	GEN056	06. Usability and Humanity Requirement	2. desirable	Provide constant feedback to users	HWU061	.	.	.	x

section in textual description	Req. #:	Req. Type:	Priority:	Description:	Derived from:	WP1	WP2	WP3	WP4
24 transparency	GEN057	06. Usability and Humanity Requirement	1. mandatory	Flag data when unsure of auto correction	HWU062	.	.	x	x
25 terminology	GEN058	06. Usability and Humanity Requirement	3. optional	Use use-case-specific metaphors when designing interface	HWU060	.	.	.	x
25 terminology	GEN059	03. Functional Requirement	1. mandatory	CUBIST software shall try to avoid (where possible) terminology of the underlying technology	SAS007	.	.	.	x
26 frontend	GEN060	05. Look and Feel	2. desirable	Interface should be clean and uncluttered	HWU051	.	.	.	x
26 frontend	GEN061	05. Look and Feel	2. desirable	The interface looks familiar by using standard web metaphors or resembling known web apps	HWU052, HWU066	.	.	.	x
26 frontend	GEN062	06. Usability and Humanity Requirement	2. desirable	the frontend can be customized	SAP006	.	.	.	x
27 documentation	GEN063	09. Maintainability and Support Requirement	2. desirable	A full manual for CUBIST is provided	HWU103, SAS014	x	.	.	.
28 performance	GEN064	06. Usability and Humanity Requirement	2. desirable	Make users aware of the progress for tasks that take > 1sec.	HWU067, SAS010	.	.	.	x
28 performance	GEN065	07. Performance Requirement	2. desirable	Most queries must have results displayed to user in less than 10 seconds	HWU080, INN015	.	.	x	x
28 performance	GEN066	07. Performance Requirement	2. desirable	a concept lattice for a context with 20 attributes is computed in less than 2 seconds	SAP004	.	.	x	x
28 performance	GEN067	07. Performance Requirement	3. optional	When results cannot be displayed in less than 50 seconds, should provide option to email results to user	HWU081	.	.	x	x

5 Concluding Remarks

It is the nature of a requirement analysis that many of the requirements are very implementation dependant, for this reason it has to be stressed that the focus of CUBIST is still on research topics. This is obvious both due the projects nature, as well as due to the three key objectives of CUBIST (taken from the DoW) which are research based, namely

- CUBIST will provide comprehensive methods to bring unstructured data into analytics.
- ST will enrich BI by enabling the discovery of new implicit information through logical reasoning.
- CUBIST will provide ways to visualize and explore hitherto undiscovered BI using FCA.

Contrasting the research focus of cubist, many of the requirements are too implementation dependant and are outside these objectives, i.e. they are more related to project process than research outcomes.

The set of requirements can be perceived to be over-ambitious for a STREP. Indeed, taking the limited resources in CUBIST into account, it is difficult to see how the requirements can be addressed in the limited time and person months allowed. To give an example, to create a new visualisation tool for FCA from scratch, that even partially satisfies the requirements is unlikely with the resources available. Existing tools such as Concept Explorer and ToscanaJ took several years to develop and are not capable of doing everything what is requested in the requirements of CUBIST.

Moreover, not all requirements are directly related to the project objectives. Some of the requirements have to do with future development and possible commercial implementations (like HWU103 “System should provide full manual describing maintenance options” which “allows management of resource by EMAGE staff”, or HWU121 “CUBIST software should be open source to allow it to be used/extended by EMAGE after project ends”). However, these requirements are still important, as they target the future exploitation of CUBIST.

Finally, some requirements are neither related to project objectives, nor to possible commercial implementations, but they target an “ideal” CUBIST system which cannot be achieved in this research project. Examples are requirements requesting sophisticated search functionalities. Though a combination in one system of BI-capabilities and search capabilities is certainly sensitive, it is unrealistic to expect that such search functionalities can be implemented with the limited resources available in CUBIST. Another example is INN011 “Learning from user feedback for skills, qualification, job description, company information etc. within job description”, which is outside the realm of CUBIST.

Having said that, it should be clear that to some extent, the requirements describe a vision of an ideal, non-prototypical CUBIST system that may be developed in the future. Thus they should be understood as *guidelines*, but not strict obligations, which have to be followed.

6 Appendix 1: Introduction into Reading FCA Diagrams

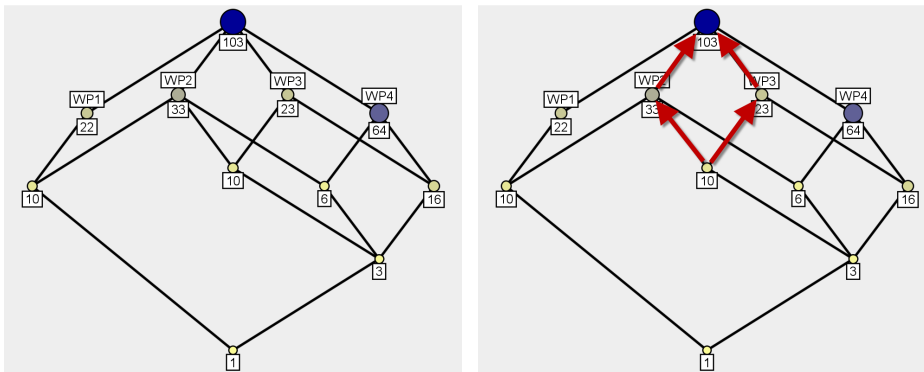
In this chapter, a short introduction into reading the diagrams is given.

In the diagrams, the requirements are clustered with respect to some attributes of the requirements, for example by their priorities or by the partner who raised the requirement.

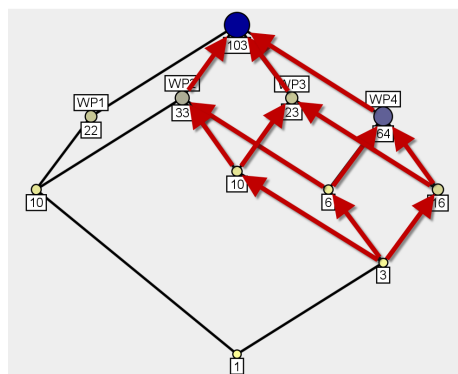
We start with the diagram where the requirements are clustered with respect to the affected work packages. Recall that all requirements have been classified according to which WP they significantly affect. Below the corresponding diagram is provided. The numbers below the nodes indicate the number of requirements; the texts above the nodes are the values of the attributes, i.e. the affected work packages. We have a closer look at the node labeled “10” (in the middle of the diagram). This node stands for ten requirements. The affected work packages are precisely those WPs in the diagram, which can be reached by starting from the node and then looking upwards. In the right diagram, we see that three nodes can be reached this way, namely:

- 1) the node labeled WP2
- 2) the node labeled WP3
- 3) the top node

Thus the 10 requirements affect WP2 and WP3, but neither WP1 nor WP4.

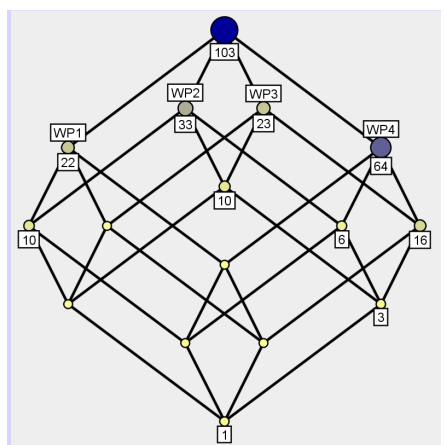


If we now look at the node labeled “3” (bottom right of diagram), we see that all three requirements are assigned to WP2, WP3 and WP4, but not to WP1.



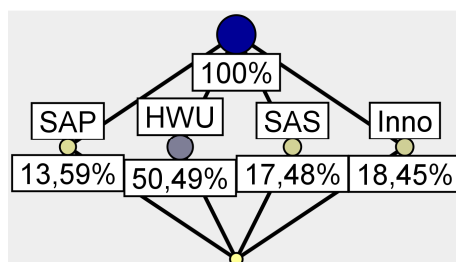
Similarly, finding out for a given work package the affected requirements, the diagram is read downwardly. For example, for WP1, 22 requirements exclusively affect WP1, ten further requirements additionally affect WP1, and one more requirement affects besides WP1 all other WPs (i.e. WP2, WP3 and WP4) as well.

Note that the diagram as such reveals some interesting insights into dependencies between the different work packages. We see that there is no node below WP1 and WP2 only, which means that there is no requirement that affects precisely those two WPs. In fact, it is possible to show instead a diagram with **all** possible combinations of WPs, thus having nodes with zero requirements assigned to it. This diagram is given below. In contrast to the diagram given before, it is now harder to realize the dependencies between different WPs.

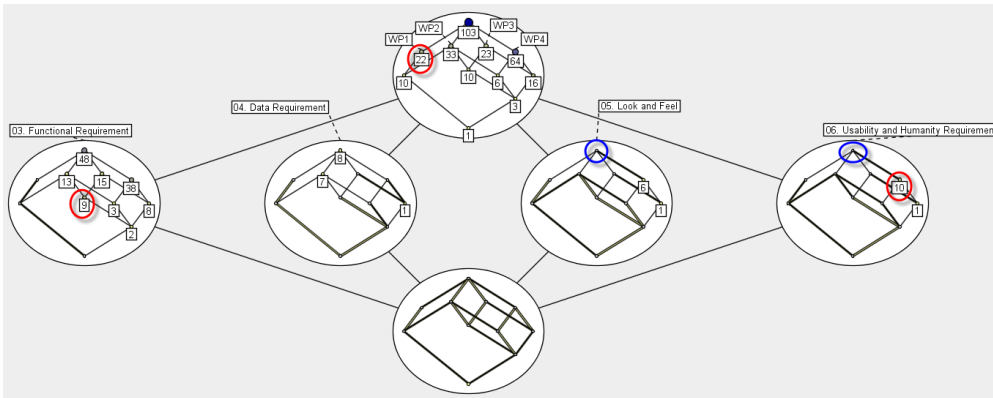


Comment [KM3]: I don't think this diagram (& associated text) contributes to the tutorial; I would cut it.

Besides giving absolute numbers, it is possible to indicate the percentage of requirements as well. For example, in the diagram below, 13.59% of all requirements are raised by SAP.



It is even possible to combine the diagrams for two attributes. The next diagram combines the affected work packages with the most important requirement types.



The “outer” diagram refers to the requirement types; the inner diagrams refer to the affected WPs. We have a look at three red encircled nodes to exemplify the reading.

The red encircled node labeled “9” in the diagram below counts the number of requirements

- 1) which affect WP2 and WP3 according to the inner diagrams and
- 2) which are functional requirements according to the outer diagram.

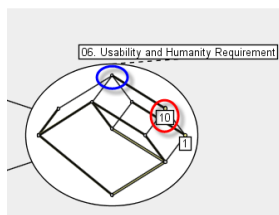
The red encircled node labeled “10” in the diagram below counts the number of requirements

- 1) which exclusively affect WP4 according to the inner diagrams and
- 2) which are usability and humanity requirements according to the outer diagram.

The red encircled node labeled “22” in the diagram below counts the number of requirements

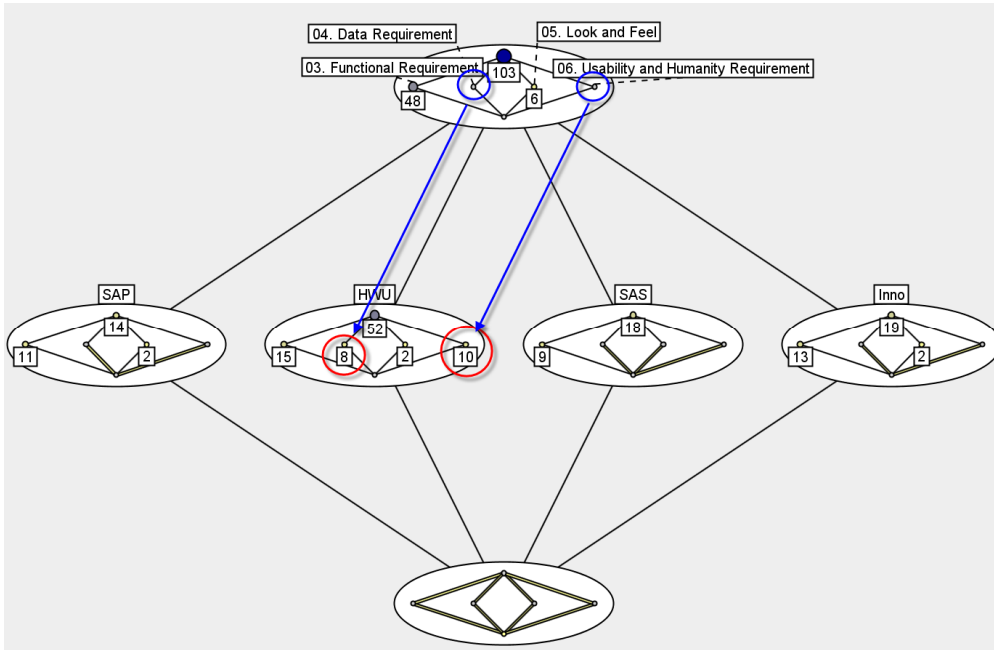
- 1) which affect WP3 and WP4 according to the inner diagrams and
- 2) do not belong to any of the requirement types 03 “functional requirement”, 04 “data requirement”, 05 “look and feel” and 06 “usability and humanity requirement” according to the outer diagram.

It should be noted that the re-used “inner” diagram about the affected work packages has been created on the basis of all requirements, whatever type they are. We have a closer look at the diagram in the node for usability and humanity requirements, as given below.



As this diagram refers only to a the small subset of usability and humanity requirements (out of all requirements), some of it nodes are information-wise not distinguishable anymore. Such nodes are connected with thick lines. Such nodes should be understood to be equally labeled. For example, the blue encircled node does contain exactly the same information as the red encircled node, that is it refers to ten (usability and humanity) requirements which affect WP4. Similarly the other blue encircled node in the overall diagram refers to six look and feel requirements which affect WP4.

This kind of reading might even span over different outer nodes. To see this, we look at the diagram of Section 2.6, which is reiterated below.



Consider the two blue encircled nodes for data requirements and usability & humanity requirements. They are not labeled, as these nodes do not accumulate these types of requirements from more than one partner. Instead,

- data requirements have been raised by HWU only
- usability and humanity requirements have been raised by HWU only

For this reason, the blue encircled nodes refer to the same number of requirements as the corresponding red encircled nodes, thus they are not labeled.

7 Appendix 2: Complete List of Formal Requirements

On the next pages, the complete list of all use-case-dependent requirements, classified according to the eleven requirement types, is provided.

7.1 Purpose of the Project

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU001	1 mand.	Provide analytical features on top of existing data	Lots of data, yet no analytical features	CUBIST provides a series of analytical features for EMAGE data							x
SAS	Alexander Mikhailian	SAS001	1 mand.	CUBIST software shall be used with space control centre data as delivered by SAS	This is one of the objectives of the DoW.	Via testing of the use case prototypes v1 and v2.	SAS			x			
SAS	Alexander Mikhailian	SAS002	1 mand.	Provide analytical features on top of existing data	Lots of data, yet no analytical features		SAS						x

7.2 Mandated Constraints

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU010	1 mand.	Must be web based	Many biological users will not control their computer & thus cannot install software	CUBIST runs in a web browser				x			
HWU	Kenneth McLeod	HWU011	1 mand.	Must not require installation of plug-in, extension etc.	Many biological users will not control their computer & thus cannot install software	CUBIST runs perfectly in a newly installed browser				x			
Inno	Hazzaz Imtiaz	INN012	1 mand.	Must be web based	installing software in each users computer is not an option	Test with Innovantage sales team	INNO			x			

Partner	Originator :	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
SAP	Frithjof Dau	SAP007	1 mand.	it is possible to zoom into a concept node	a given concept node corresponds to a specific subset of the information space, and a user might want to explore only the given subset	CUBIST allows to zoom into a concept node and to conduct further analysis (again with visual analytics) for the objects and attributes of the given concept node							x
SAP	Frithjof Dau	SAP008	1 mand.	it is possible to select a part of the overall concept lattice to be displayed	the overall concept lattice might be too large to be displayed in total	CUBIST offers possibilities to zoom into parts of the concept lattice, e.g. by displaying for a given node only nodes in the neighbourhood, by using a fisheye view for segment of the concept lattice							x
SAS	Alexander Mikhailian	SAS003	1 mand.	CUBIST software shall use FCA for suggesting new patterns in data	This is one of the objectives of the DoW.	Via testing of the use case prototypes v1 and v2.	SAS						x
SAS	Alexander Mikhailian	SAS004	1 mand.	CUBIST software shall produce visual representation of data	This is one of the objectives of the DoW.	Via testing of the use case prototypes v1 and v2.	SAS						x
SAS	Alexander Mikhailian	SAS005	1 mand.	CUBIST software shall store existing data in an RDF store and provide an extended SPARQL query interface to the data.	This is one of the objectives of the DoW.	Via testing of the use case prototypes v1 and v2.	SAS				x		
SAS	Alexander Mikhailian	SAS007	1 mand.	CUBIST software shall try to avoid (where possible) terminology of the underlying technology	Users come from different backgrounds	Via the review of text strings deployed in software.	SAS						x
SAS	Alexander Mikhailian	SAS010	2 des.	Make users aware of the progress for tasks that take > 1sec.	Users must realise that system is working	By review of design	SAS					x	x
HWU	Kenneth McLeod	HWU022	2 des.	Locate errors/inconsistencies in underlying data	Biological data is naturally inconsistent & incomplete. Users find it hard to deal with these issues	Suggests inconsistency "solutions" to user							x
HWU	Kenneth McLeod	HWU023	2 des.	Display trend information for expression level	Level changes over time	Ask human expert to verify three examples						x	x

Partner	Originator :	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU024	2 des.	Display trend information for expression pattern	Pattern changes over time	Ask human expert to verify three examples						x	x
HWU	Kenneth McLeod	HWU025	2 des.	Display trend information for co-expression level	Co-expression changes over time	Ask human expert to verify three examples						x	x
HWU	Kenneth McLeod	HWU027	2 des.	Identify genes with similar expression patterns	Provide spatially orientated co-expression information	Ask human expert to verify three examples							x
HWU	Kenneth McLeod	HWU028	2 des.	Describe similarities of genes involved in same process	Indicates what other processes, functions etc the genes may be involved in	Ask human expert to verify three examples							x
HWU	Kenneth McLeod	HWU030	2 des.	Queries & results may be saved & reloaded later	Allows users to return to previous state	A test query can be exported, then later imported successfully							x
HWU	Kenneth McLeod	HWU031	1 mand.	Filter results (before lattice is computed)	Users may find it helpful to deal with a subset of the full result set	Result set can be reduced by setting a range of parameters						x	x
HWU	Kenneth McLeod	HWU032	2 des.	Flexible presentation of results	Different personas have different computing background & experience	User can switch to different visualisation of same information		See D7.1.1. section 4					x
HWU	Kenneth McLeod	HWU033	2 des.	Link back to main EMAGE page	If users wish to drill down, they should be referred to EMAGE web pages	CUBIST links out to EMAGE web pages		http://www.emouseatlas.org/emage/search/url.html			x		x
Inno	Hazzaz Imtiaz	INN006	2 des.	Multilevel job categorisation from text mining the job description, rule based or unsupervised and also supervised or learning from user feedbacks considering a feature set.	Current job categorisation is based on a set of phrases and not very granular	query relevant RDF graphs and test the usecase prototype	INNO				x	x	
Inno	Hazzaz Imtiaz	INN007	2 des.	Extraction of sentiment phrases within job description	Sentiment phrases often contain important facts about the job which need to be indexed and queried	query relevant RDF graphs and test the usecase prototype	INNO				x	x	

Partner	Originator :	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
SAP	Frithjof Dau	SAP002	2 des.	For a given attribute, CUBIST can generate a conceptual scale on the fly	it is not feasible that for all attributes, conceptual scales are predefined								x
SAP	Frithjof Dau	SAP003	2 des.	For some attributes, conceptual scales can be predefined	For some attributes, automatically generated scales are unlikely to be appropriate	CUBIST offers a set of predefined scales in the frontend to the user							x
SAP	Frithjof Dau	SAP011	2 des.	CUBIST offers different layout algorithms for the layout of a Hasse diagram	there exist different layout algorithms, and depending on the use case and user context, one or the other is more appropriate	CUBIST offers to choose from different layout algorithms (e.g. spring algorithm or a chain decomposition)							x
SAP	Frithjof Dau	SAP013	2 des.	visualizations can be stored and reloaded	For good visualizations, users might want to be able to store them in order to reuse them later	"load" and "save" functionality available for visualizations							x
SAP	Frithjof Dau	SAP014	2 des.	visualizations can be exported	For good visualizations, users might want to export them (e.g. to png or gif) on order to print or distribute them	"export" functionality available for visualizations							x
SAS	Alexander Mikhailian	SAS006	2 des.	CUBIST shall research analytical and BI solutions to space control centre document management issues	A number of potential improvement areas related to document management has been identified.		SAS						x
SAS	Alexander Mikhailian	SAS008	2 des.	Provide a FCA-based visual outlook of the data in the triple store	Space control centre operations run on tight schedule. A graphical UI for the otherwise numeric data will simplify cognitive operations		SAS						x
SAS	Alexander Mikhailian	SAS009	2 des.	Provide an interactive environment to users	Users need to navigate the data and to receive instant feedback from software.	By review of design	SAS						x
SAS	Alexander Mikhailian	SAS011	2 des.	It shall be possible to incrementally update the datastore with more data	This is essential for the telemetry use case where data comes in at a constant rate.	By review of design	SAS						x

Partner	Originator :	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU034	3 opt.	Expanded means to investigate genes	Currently users forced to query in specific ways only	Can search for gene, gene + location... with or without a time restriction							x
HWU	Kenneth McLeod	HWU035	3 opt.	Expanded means to investigate locations	Currently users forced to query in specific ways only	Can search for location, gene + location... with or without a time restriction							x
HWU	Kenneth McLeod	HWU037	3 opt.	Expanded means to investigate co-expressions	Currently users forced to query in specific ways only	Can search for co-expression alone, co-expression in set of particular locations, co-expression for particular genes, co-expression for particular genes in a set of locations, co-expression for particular GO terms... all with(out) a time restriction							x
HWU	Kenneth McLeod	HWU039	3 opt.	Compare multiple experiments	Useful to go back and look at a range of experiments; comparing and contrasting their contents	Can enter multiple IDs and get the same presentation as HWU038 for each one							x
SAP	Frithjof Dau	SAP010	3 opt.	CUBIST suggests scales and visualizations for given attributes	Inexperienced users might be overstrained with choosing appropriate scales or visualizations	CUBIST offers a "suggest visualization" button for a selected attribute							x
SAP	Frithjof Dau	SAP012	3 opt.	For a given context, CUBIST can compute implications and association rules	for some use case partners, association rules are an appropriate means to explore the data sets	CUBIST offers a button "rules" with sliders for the confidence and support for association rules, and by pressing the button, the according rules are computed and displayed							x

7.4 Data Requirements

WHERE IS BOTTOM OF THIS TABLE?

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU040	1 mand.	CUBIST warehouse must include textual annotations (unrelated to FCA viz)	Core dataset for CUBIST use case	See description					x		
HWU	Kenneth McLeod	HWU041	1 mand.	Repository should include EMAP anatomy ontologies	Required for textual annotations - HWU040	See description					x		
HWU	Kenneth McLeod	HWU042	1 mand.	Repository should contain EMAGE controlled vocabulary	Required to understand data in EMAGE; e.g. distinction between moderate & present	See description					x		
HWU	Kenneth McLeod	HWU043	3 opt.	Any inconsistencies/errors identified should be reported to EMAGE curators	Curators are responsible for maintaining accuracy of resource	See description						x	x
HWU	Kenneth McLeod	HWU044	2 des.	Warehouse should be updated inline with regular EMAGE updates	EMAGE contents constantly changes; new versions are released quarterly						x		
HWU	Kenneth McLeod	HWU045	2 des.	Repository should include spatial annotations	Core dataset for EMAGE	See description					x		
HWU	Kenneth McLeod	HWU046	2 des.	Repository should include spatial representation of mouse	Core dataset for EMAGE	See description					x		
HWU	Kenneth McLeod	HWU047	3 opt.	Repository should include EMAGE's pre-computed spatial clusters	Required for HWU032	See description					x		

7.5 Look and Feel

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
SAP	Frithjof Dau	SAP009	1 mand.	it is possible to manually move nodes in the display of the concept lattice	the automatic layout of the concept lattice might be inappropriate or insufficient, and the user want to manually correct the layout	CUBIST allows nodes to be manually moved							x
HWU	Kenneth McLeod	HWU051	2 des.	Interface should be clean and uncluttered	Makes interface easier to use			HWU020	HWU052				x
Inno	Hazzaz Imtiaz	INN016	2 des.	Look and Feel of the UI in general needs to be lively, funky which includes colour of the panels and fonts, font needs to be appropriate	Should have a user friendly interface	Test with Innovantage sales team	INNO						x
Inno	Hazzaz Imtiaz	INN017	2 des.	Start page can have a basic search interface on left pane and current jobs on right pane, i.e. reed.co.uk.	Should have a user friendly interface	Test with Innovantage sales team	INNO					x	x
SAP	Frithjof Dau	SAP006	2 des.	the frontend can be customized	different users and use case partners have different priorities and preferences for the look & feel of the frontend	CUBIST offers a set of options for the display of concept lattices							x
HWU	Kenneth McLeod	HWU052	3 opt.	Looks like prominent web apps	Mimicking resource will make CUBIST seem friendlier & easier to use	See description			HWU051				x

7.6 Usability and Humanity Requirement

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU060	3 opt.	Use biological metaphors when designing interface	Biologists understand biological metaphors but not computing ones								x
HWU	Kenneth McLeod	HWU061	2 des.	Provide constant feedback to users	Biologists are often not computer experts & need lots of support. Feedback improves the transparency too.	Users are able to describe what CUBIST is doing whilst using the system		HWU020					x
HWU	Kenneth McLeod	HWU062	1 mand.	Flag data when unsure of auto correction (HWU022).	System needs to be transparent and provide constant feedback to users	Users can tell when CUBIST has changed the underlying context (e.g. with fault tolerance algorithms)		HWU020				x	x
HWU	Kenneth McLeod	HWU064	2 des.	Users will be able to use system once they have watched the screencast	Users are unwilling to read manual or work through a tutorial	Users can use CUBIST when their only introduction is the screencast							x
HWU	Kenneth McLeod	HWU065	2 des.	Users should only need to watch screencast once	Users are unwilling to use tools that seem hard to use	Users can use CUBIST 2 weeks later without any help or reminders							x
HWU	Kenneth McLeod	HWU066	2 des.	Use standard web metaphors	Users are familiar with web								x
HWU	Kenneth McLeod	HWU067	2 des.	Make users aware of progress	Users must realise that system is working			HWU020					x
HWU	Kenneth McLeod	HWU068	2 des.	Interface should be simple	Biologist are not comfortable with complex interfaces, and traditionally ignore them	Evaluation shows biologists describe the interface as "simple"			HWU052; HWU069				x
HWU	Kenneth McLeod	HWU069	2 des.	Visualisations should be detailed and convey large amounts of information (for personal computational scientist)	If they don't might as well scan web pages	Computational biologists indicate all "key" information is in visualisation			HWU068; HWU070				x
HWU	Kenneth McLeod	HWU070	2 des.	Visualisations should be simple (for persona biologist)	Biologists are not used to dealing with lattices and other forms of visualisation CUBIST is likely to employ	Biologists can interpret visualisations			HWU069				x

7.7 Performance Requirement

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
SAP	Frithjof Dau	SAP004	2 des.	a concept lattice for a context with 20 attributes is computed in less than 2 seconds	users expect near real-time display of concept lattices							x	x
HWU	Kenneth McLeod	HWU080	2 des.	80% of queries must have results displayed to user in less than 50 seconds	Web based interfaces need quick response times, or they appear to be broken	See description						x	x
Inno	Hazzaz Imtiaz	INN015	2 des.	Most queries must have results displayed to user in less than 10 seconds	Web based interfaces need quick response times, or they appear to be broken	Test with Innovantage sales team	INNO					x	x
HWU	Kenneth McLeod	HWU081	3 opt.	When results cannot be displayed in less than 50 seconds, should provide option to email results to user	Many biological tasks are computationally expensive. Standard procedure is to email user when finished	See description						x	x

7.8 Operational Requirement

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU090	1 mand.	Must not rely on windows technology	Significant number of linux & mac boxes in use	5 test queries can be run using linux & mac machines				x			
HWU	Kenneth McLeod	HWU091	2 des.	Works and looks similar with major browsers (with flash support). No mobile browsers targeted	Significant number of users use firefox, safari & IE browsers	CUBIST looks the same in all 3 browsers; works for 5 test queries.				x			
Inno	Hazzaz Imtiaz	INN013	2 des.	Works and looks similar with major browsers (with flash support). No mobile browsers targeted	users use firefox, safari, internet explorer, chrome	CUBIST looks the same in all 3 browsers; works for 5 test queries.	INNO			x			
SAS	Alexander Mikhailian	SAS012	2 des.	The installation procedure shall be scriptable and allow for automation	Easiness of system administration and operational support is of paramount importance	By testing	SAS						x

7.9 Maintainability and Support Requirement

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU100	1 mand.	Pipeline to important more data automatically	EMAGE data is continually expanding	CUBIST contains latest version of EMAGE data set				x	x		
HWU	Kenneth McLeod	HWU101	1 mand.	Pipeline to automatically import new data will only import new/amended data	Pulling entire dataset every time is too expensive	See description				x	x		
SAS	Alexander Mikhailian	SAS013	1 mand.	CUBIST users shall be able to have complete control over the data managed by CUBIST software and the CUBIST software as such	CUBIST users shall have complete control over the data managed by CUBIST software and the CUBIST software as such	By architecture review. See also the relevant Security requirement	SAS			x	x	x	x
SAS	Alexander Mikhailian	SAS014	2 des.	The system shall be documented.	Inputs for user training and subsequent reference documentation shall be available.		SAS			x			
HWU	Kenneth McLeod	HWU102	2 des.	Maintenance tool should be online	Allows management of EMAGE specific aspects by EMAGE people, even if run on server elsewhere	An EMAGE admin can control the system remotely				x			
HWU	Kenneth McLeod	HWU103	2 des.	System should provide full manual describing maintenance options	Allows management of resource by EMAGE staff	An EMAGE admin can change the system				x			

7.10 Security Requirements

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU110	1 mand.	Only system admin & CUBIST developers have write access to data	Data should not be changed by anyone other than trusted admin	See description				x	x		
HWU	Kenneth McLeod	HWU111	1 mand.	CUBIST should not attempt to change original data source	Data should not be changed by anyone other than trusted admin	CUBIST will only have read only access anyway				x	x		
Inno	Hazzaz Imtiaz	INN014	1 mand.	Only system admin & developers have write access to data	Data should not be changed by anyone other than trusted admin		INNO			x	x		
SAS	Alexander Mikhailian	SAS015	1 mand.	CUBIST software shall allow the user to host all the data locally (not locally on the PC; but in the environment)	Some data related to space control centre operations has restrictions with regards to the location of data. For instance, operational data is not allowed to leave the operations environment.	By architecture review. See also the relevant Maintainability and Support Requirement.				x	x		
SAS	Alexander Mikhailian	SAS016	1 mand.	Public cloud services shall not be used to store and process space control centre data unless complete confidentiality of such data is assured.	CUBIST users shall have complete control over the data managed by CUBIST software.	No public cloud service is used.	SAS			x			

7.11 Legal Requirements

Partner	Originator:	Req. #:	Priority:	Description:	Rationale:	Fit Criterion:	Use Case(s):	Support. Materials:	Conflicts:	WP1	WP2	WP3	WP4
HWU	Kenneth McLeod	HWU120	1 mand.	Information provider (ie researcher & journal) must be credited (anywhere on the screen)	Standard practise to acknowledge knowledge creator; also, some experimental results (images) are copyright	A user can determine who performed the experiment					x		x
SAS	Alexander Mikhailian	SAS018	1 mand.	The licensing agreement for CUBIST software shall not put any restriction on data use.	Space control centre data utilisation and distribution is already strictly controlled.	Absence of an data-related statements in the CUBIST software licensing agreement.	SAS			x	x		
SAS	Alexander Mikhailian	SAS019	1 mand.	The originator of the data has to be credited.	This means in practice that the data provided by B.USOC or ESA has to be credited.	By testing					x		x
HWU	Kenneth McLeod	HWU121	2 des.	CUBIST software should be open source to allow it to be used/extended by EMAGE after project ends	Technology in CUBIST may prove useful, and EMAGE wish to be able to use and extend it after project ends	All code and documentation is publically available, and licensed for public use				x			
HWU	Kenneth McLeod	HWU122	2 des.	Open access to EMAGE SPARQL (SPARQL endpoint at HWU for data)	EMAGE is a public resource funded by the UK tax payer	Any interested party can use the EMAGE SPARQL query mechanism				x	x		
HWU	Kenneth McLeod	HWU123	3 opt.	Open access to EMAGE data in CUBIST warehouse	Registration provides access to the warehouse for interested parties	Once registered, users have full read access to EMAGE data in warehouse				x	x		

