SENSE4US

First Functional Requirements Description

Project acronym:	**SENSE4US**
Project full title:	**Data Insights for Policy Makers and Citizens**
Grant agreement no.:	**611242**
Responsible:	**Steve Taylor**
Contributors:	**Juri Papay, Timo Wandhöfer, Beccy Allen, Rasa Uzdavinyte, Nasir Naveed, Aron Larsson, Miriam Fernandez**
Document Reference:	**D3.1**
Dissemination Level:	**PU**
Version:	**FINAL**
Date:	**7/10/14**

Project co-funded by the European Union under the Seventh Framework Programme
© Copyright 2014 University of Southampton IT Innovation Centre and other partners of the Sense4Us project

History

Version	*Date*	*Modification reason*	*Modified by*
0.1	22/05/2014	Initial Draft	Juri Papay
1.0	7/10/2014	Draft for Review	Steve Taylor
1.0	10/10/2014	Review	Timo Wandhöfer[1]
1.1	17/10/2014	Submitted Version	Steve Taylor

[1] Timo Wandhöfer is also listed as a contributor to this document. This is not a conflict of interest, as Timo is an end user partner, and as such his contribution to the deliverable was to verify the priority of the functional requirements (as did all other end user partners) and contribute to the survey of allocations, (as did all other partners). The review of this deliverable regards the whole of the functional specification process and the results generated by the process, and it was decided that an end user partner should be the chief reviewer, to evaluate the validity of the overall process and results.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

Table of contents

History **2**
Table of contents **3**
List of tables **5**
List of abbreviations **6**
Executive summary **7**
1 Introduction **8**
1.1 Functional Specification Process 8
1.2 Deliverable Structure 9
2 Determination of Functional Requirements from User Requirements **10**
2.1 EU requirements 10
2.1.1 High priority 10
2.1.2 Medium priority 13
2.1.3 Low priority 14
2.2 UK Requirements 15
2.2.1 High priority 15
2.2.2 Medium priority 18
2.2.3 Low priority 19
2.3 DE Requirements 20
2.3.1 High Priority 20
2.3.2 Medium Priority 22
2.3.3 Low Priority 22
2.4 Verification of Functional Requirements with End User Partners 23
3 Prioritisation of Functional Requirements **24**
4 Allocation of Functional Requirements to Sense4Us Research and Development Work **29**
5 Determine If and How Functional Requirements Can be Addressed **37**
6 Collate Functional Specification **38**
6.1 Policy Modelling 39
6.2 Presentation / Visualisation of Results 41
6.3 Search 46
6.4 Topic Analysis 51
6.5 User Environment / Interface 52
6.6 Social Media Analysis 56

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

7 Conclusions 59

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

List of tables

Table 1: Prioritised Functional Requirements *28*
Table 2: Allocation of Functional Requirements to Components *36*
Table 3: Policy Modelling Functional Specification *40*
Table 4: Functional Specification for Presentation / Visualisation of Results *45*
Table 5: Functional Specification for Search *50*
Table 6: Functional Specification for Topic Analysis *51*
Table 7: Functional Specification for User Environment / Interface *55*
Table 8: Functional Specification for Social Media Analysis *58*

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

List of abbreviations

Abbreviation	*Description*
PM	Project Month – PM1 (October 2013) is the first month of the project, etc. PM12 is Sept 2014.
WP	Workpackage
DoW	Description of Work
PR	Press Release
Research Partners	The partners responsible for the research components of the project: U Koblenz (WP4), Open University (WP5), U Stockholm (WP6)
End User Partners	The partners responsible for engagement with the external policy makers who are our target users. These are: Gov2u, Hansard Society, GESIS.
Factor	We primarily use a definition of "factor" to relate to the policy modelling task in WP6. Here, a factor is an element that has some causal relationship to the policy being modelled. For example, if we are modelling unemployment, then some example factors could be the unemployment level itself (the output), the geographical area (input) and the number of businesses in that area (input.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

Executive summary

This deliverable describes the process and results of work to find the initial functional requirements of the Sense4Us toolkit. Specifically, it describes the steps taken to transform the user requirements from D2.1 into the functional specification that is the overall goal of this deliverable.

The primary objective of the transformation process is that it should result in a clear and logical path from the end user requirements to a functional specification. This path is important to ensure that the functional specification actually answers the end user requirements.

During the early stages of the functional requirements analysis process, a process was decided upon to transform the requirements in D2.1 into a functional specification. The process has the following major steps:

1. Determine functional requirements from analysis of the user requirements in D2.1.
2. Prioritise the functional requirements, so the critical and most important ones are identified and addressed first.
3. Determine which of the project's components should address each requirement.
4. Determine if and how the allocated components should address the functional requirements.
5. Synthesise the results of the previous step into a requirements list and functional specification for each component.

The major result of this deliverable is a set of tables that describe the functional requirements and the functional specification of components (in Section 6). This set of tables will be used as input to the architecture task, which will result in the deliverable D3.2, due PM18. Because they contain specifications for each of the components in the research and integration work in the project, the tables will also be used as a benchmark by which to evaluate the research and integrating partners' work – we can use the tables as a checklist of functionality that the different components need, and evaluate the components as to whether they provide that functionality.

Another exploitable result of the work described in this deliverable is also the design and testing of the functional specification process itself. It is believed that the process can be useful and widely applicable to other projects.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

1 Introduction

This deliverable describes the process and results of work to find the initial functional requirements of the Sense4Us toolkit. Specifically, it describes the steps taken to transform the user requirements from D2.1 into the functional specification that is the overall goal of this deliverable.

The primary objective of the transformation process is that it should result in a clear and logical path from the end user requirements to a functional specification. This path is important to ensure that the functional specification actually answers the end user requirements. To achieve this clarity, it was decided that we should specify a formal process for the requirements analysis and the process by which to define the functional specification.

1.1 Functional Specification Process

During the early stages of functional requirements analysis, a process (named the "functional specification process") was determined to transform the requirements in D2.1 into a functional specification. The key to this process is functional requirements, which are derived from the requirements in D2.1. For reference, we define the two requirement types as follows:

- An end-user requirement is something needed by an end-user. This is often expressed as a gap, a need or a problem to be solved, and it may contain suggestions for functionality, but not necessarily. D2.1 has end user requirements.
- A functional requirement is what the Sense4Us toolkit needs to do to address end-user requirements. The functional specification is a collection of functionality directly aimed at addressing the functional requirements.

Given this definition, the functional requirements serve as a bridge between the end user requirements specified in D2.1 and the functional specification.

The functional specification process has the following major steps:

1. Determine functional requirements from analysis of the user requirements in D2.1.
2. Prioritise the functional requirements, so the critical and most important ones are identified and addressed first.
3. Determine which functional components[2] should address each requirement.
4. Determine if and how the allocated components should address the functional requirements.
5. Synthesise the results of the previous step into a requirements list and specifications for each component.

The result of this is the functional specification, the target outcome of this deliverable.

The above process was deliberately designed to be open and consultative:

- In steps 1 and 2 above, the end user partners were involved, firstly to check that the functional requirements determined from their end user requirements were sensible, and secondly to assess the prioritisation and identify any functional requirements that are critical. After step 2, we had a verified and prioritised list of functional requirements.

[2] The functional components referenced here broadly follow the workpackage structure of the project, e.g. Search (WP4), Social Media Analysis (WP5), Policy Modelling (WP6), Topic Analysis (WP4), Visualisation of Results (WP4, WP5, WP6, WP7) and User Interface (WP7). More details on this breakdown can be found in Section 4.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

- In step 3 (allocation of requirements to functional components) all partners were involved, to give a balanced opinion about which components should address the functional requirements.
- Step 4 (deciding on how the components addressed the functional requirements) was the responsibility of the partners who were developing the components.
- Finally, step 5 was the responsibility of IT Innovation, as project coordinator and integrating partner, to determine the overall functional specification given all these inputs.

1.2 Deliverable Structure

This deliverable is broadly structured following the steps described above. Firstly in Section 2 there is an analysis of end user requirements from D2.1 to determine the functional requirements from them. Next, Section 3 describes how the functional requirements are prioritised and Section 4 describes how the functional requirements were allocated to functional components. In Section 5, the project's component developers describe the ways their component can address these requirements. The major output of deliverable is in Section 6, where the answers to how to address requirements are collated by component and are presented in tables, thus forming a specification of functionality for each component. This serves as input to the Sense4us architecture that is the responsibility of the next deliverable, D3.2, due PM18, and the work presented in D3.2 will develop and enhance the work presented here. The deliverable concludes with some remarks about how the functional specifications will be used, and some comments in general about the efficacy of the process.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

2 Determination of Functional Requirements from User Requirements

The first task is to determine the functional requirements of the first Sense4Us toolkit by analysis of the end-user requirements described in D2.1. Clearly, the functional requirements derive from interpretation of the user requirements – the functional requirements are the result of analysis of what functionality is required to answer a gap or need requested by the end users. Therefore, this step must plainly describe the relationship between the end-user requirements from D2.1 and the functional requirements derived from them, so it is obvious how the functionality proposed by the toolkit addresses the end-user requirements.

The functional requirements derived from the end user requirements are described below. In many cases, we have seen the same functional requirement mentioned in more than one end user requirement. It is also the case that an end user requirement often contains more than one functional requirement. Given this, the functional requirements are separate from the user requirements because there is no one-to-one mapping between them, but the origin of the functional requirements are preserved so as to be clear where the functional requirements came from.

This section comprises an annotated version of the end user requirements from D2.1. The text from D2.1 is *italicised*, and comments are made on each end user requirement (in non-italics). Where functional requirements are identified they are highlighted in **bold and given an identity number (e.g. F-1)**, so as to easily identify each functional requirement in subsequent analysis etc.

2.1 EU requirements

2.1.1 High priority

Requirement EU-1: A top requirement, mentioned by all of the interviewed policy makers and clearly evident from the survey results is a wide scale search*: a search tool should bring up all the online information across:*

- *different **data formats** and **data sources**: the online sources normally used in the policy drafting at the EU level (internal and external sources of information: legal documents, proposals from the European Commission, OECD, Committee of Regions, e-committee/working documents and EC and EP impact assessments, academic papers, newspapers, studies from governments, studies)*

- The functional requirement here is clear: to be able to **search different data formats and interfaces to multiple data sources (F-1)**.
- It is unlikely that the project will discover all data sources needed by policy makers during its development phases so a tacit related functional requirement is to be able to **add new data sources (F-2)** that can be searched by the toolkit while the toolkit is in operation, to accommodate new data sources as they appear.

- ***key stakeholders**, especially permanent ones, that are key interest groups in the usual political cycle. The search tool could scan and gather information from the websites of these lobby organisations as well as the Transparency Register in order to help inform the EU policy making process in real time;*

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

- The functional requirement here is similar to the previous requirement – to **enable the user to add specific data sources and websites as search targets (F-3)** – i.e. their favourite lobby websites.

- *relevant press, social media and other forums for public opinion.*

- The functional requirements here are obvious:
  - **enable traditional online media (e.g. news sites) to be searched (F-4)**; and
  - **enable opinions from social media and other forums to be gathered (F-5)**.

***Requirement EU-2: The majority of policy makers preferred a cross disciplinary search**: policy makers would like to search data across policy areas/topics, related to an issue, for example, the Single Market and SME's and Environment, so that searching multiple databases or other sources across a range of disciplines would be unnecessary and in that way save time.*

This requirement is about searching across disciplines and databases, for information related to an issue. There is also an implied requirement for searching multiple data sources automatically with one search request from the user. Given this, the function required from this is:

- **enable searching multiple data sources with the same query and mouse click from the user (F-6)**; and
- **enable finding information for subjects related to those in the policy (F-7)** as well as those directly involved with the policy.

***Requirement EU-3: It is clear from the survey and interview findings that visualization, filtering and structure of the search results is of high importance**, bearing in mind that the key element when choosing a data set to trust or to follow is often determined by the authorship of the data. For example weight can be assigned to data sources corresponding to whether it is officially published or how well the author/organization is known and supported publicly. Important factors for the trust decision about a data source include:*

- *relevant actors/organisations (with contact details);*
- *relevant legal acts and related documents;*
- *relevant articles and studies;*
- *relevant social media discussions.*

This requirement discusses filtration of data sources, and the justification of this is given in the discussion about how the policy maker can decide whether they trust a data source or not. Given this discussion, it is clear that providing evidence so the policy maker can make a decision about whether they trust a data source or not is important. There are therefore two functional requirements from this:

- **enabling the user to filter search results - include or exclude results according to a wide range of criteria (F-8)**; and
- to provide evidence to enable the user to make a judgement about whether they trust a data source - **present to the user all available information about the provenance of some data that has been found by the tool (F-9).**

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

There is a subsidiary functional requirement based on the combination of the two functional requirements above: to **enable the user to restrict search results to those that they trust (F-10)**.

It is not clear what "structure" the data means, but organising the data, in particular **enabling the user to sort data by different criteria (F-11)** is anticipated to be a functional requirement.

Requirement EU-4: It was strongly underlined by the interviewed policy makers that policy modelling part of the tool needs to be trusted and therefore transparent. *The policy modelling facility has to present a clear methodology of how it came to its conclusions, including the possibility to use scientific arguments if they apply as well as publicly available sentiments. Where evidence is present (e.g. sources, authors, etc), it should be presented.*

The functionality required here is not explicit, but the requirement alludes to two major functional requirements: helping the user understand (and have faith in) the policy modelling aspect of the project, and to help them include all factors (the things they think will have a bearing on the simulation) they wish to include.

To help users understand the policy modelling, there should be **easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)**. This must be in laypersons language, and avoid any technical language or jargon.

To help users believe the results of policy modelling, two approaches are suggested:

- **Transparency in the creation and operation of the policy models (F-13).** The user should be given the opportunity to participate in the creation of policy models, e.g. deciding on key factors and their relationships if they so wish, as this gives them a significant element of control and understanding of the workings of the model. In addition, the policy models should not be considered as black boxes, where the user has no idea of the relationship between the models' factors. The models should be able to be "opened up" and explored by the user if they wish to, so that intermediate stages and their data values can be examined. This gives the user the information to make a judgement about whether they believe the output of the models.
- Data that is selected as a factor for the modelling should have as much provenance information as is available displayed with it, so the user can make a judgement about whether they wish to trust this data. Therefore, the previously identified functional requirement: "**present to the user all available information about the provenance of some data that has been found by the tool (F-9)**" is reinforced here.

Some specific data sources are mentioned in the requirement, so the associated functional requirements for these are as follows:

- **Enable the searching for scientific information relevant to the policy subjects (F-14).**
- **Enable the collection of sentiments / opinions about the policy subjects (F-15).**

It is impossible to imagine all the relevant factors that could be included, as this is up to the policy maker, but **the policy maker should be able to examine a policy model and to add new factors if necessary (F-16)**. In general, the tool should help the users include a wide variety of factors in modelling – as many as they wish to.

Some factors related to the policy may be unknown to the user, and providing a means to discover factors related to their policy issue should be of great benefit to them, especially if the user did not know about a particular factor. The functional requirement related to this is

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

therefore: **enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17).**

Requirement EU-5: User functionality in policy modelling is of high priority, because this requirement highly determines the success and believability of the tool*. Most of the policy makers indicated that the problems in predicting the outcomes of a certain political development lay in the complexity of the issue and the dynamic range of factors that might affect it. There are also a number of actors at any one given time that exercise influence on the policy. Most of the respondents indicated that it is difficult to ensure that a complete set of relevant variables is in the model, as policy in a generic sense is a social discipline and factors may change fast. However, the tool should include all factors wanted by the user and policy models should be based on all the variables the user considers relevant.*

This requirement backs up the previous requirement, EU-4, by adding more rationale to enable the user to add new factors as they wish. The major functional requirement therefore reinforces F-17 above: **enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17).**

2.1.2 Medium priority

Requirement EU-6: Summary of social media analysis*: Get a summary of structured data available on social media on specific themes and be able to identify the social group and its sentiment and opinions.*

The functionality required to address this user requirement is clear:

- **to be able to acquire information from social media on the topics and themes of the policy in question (F-18)**;
- **to enable opinions from social media and other forums to be gathered (F-5)** (already identified in EU-1); and
- **to discover any publicly available characteristic information about the social media participants who are expressing the opinions (F-19)**.

Requirement EU-7: Unlimited time origin for data: *Historical data to be used mainly for countries with stable political systems and history (as opposed to constantly changing situations/ crisis in some EU countries). Instability of current and up-to-date data/evidence/situation/context in EU Member States makes simulations difficult.*

Some data may be relevant to a particular time frame, e.g. census data – this is a snapshot of a population at the time of the census, and this requirement alludes to being able to search for historical data that is relevant at a specified point in time. Therefore the functional requirement is **to enable the user to specify different time ranges for when the data was (or is) relevant in a search (F-20)**.

Requirement EU-8: "Search save" facility: *Policy makers indicated their wish to save any search made on broad or specific terms/topics and have a search history, so they could come back to it later.*

This requirement is about users being able to store and retrieve the history of their searches, data and analyses. It is not stated that segregation of users' data and histories is required (so

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

that data is protected and one user cannot access another's data), but it is reasonable to assume so, because users can then have their own working space and the data within it will be unavailable for anyone else to see or tamper with. Given all this, the functional requirement is to **provide user accounts, with storage and retrieval of data and workflows all within a user's account (F-21).**

Requirement EU-9: Cross-border EU Member State data search*: Be able to find policy related data that is relevant at the national level in EU Member States on legislative proposals initiated at EU level*

The functionality required here is to **locate information related to the topic in question, but restricted to specified countries (F-22)**.

Requirement EU-10: Include existing regular information sources, namely impact assessment studies*: Make Sense4us an advanced version of the impact assessment currently conducted by the European Commission at early stage of the policy-making process. It was even suggested that perhaps the Sense4us tool could become a common base for impact assessments, instead of having two separate ones in the EC and EP - if both assessment stages would be loaded and combined in the tool.*

This requirement is explicit about one function required, and this is providing the ability for the user to specify their favourite sources, so that they may be searched automatically in any search without having to specify them every time. Therefore functional requirement F-3, "**enable the user to add specific data sources and websites as search targets (F-3)**" is reinforced here.

The requirement also discusses impact assessment studies, and expresses the need for Sense4Us to **provide a facility to assess the impact of a policy decision (F-23)**. A related (optional) functional requirement concerns the statement about impact assessment being made common between the European Commission and the European Parliament. It is not clear what the specific function required arising from this is, but it is reasonable to assume as a starting point that the function required is **the ability to use data from both the European Commission and the European Parliament (F-24)**.

2.1.3 Low priority

Requirement EU-11: Search word flexibility*: The search engine should be able to use both broad and specific terms, not be limited linguistically when certain word searches provide no results, and the most commonly used key words in a certain policy field should be searchable. Without this functionality the tool will not provide added value to current popular search engines.*

Given the statements about broad and specific terms and the removal of linguistic limitations, together with the implied wish to avoid searches with no results, the functional requirement here is to **augment search terms the user has specified and perform additional and related searches to the ones the user is running (F-25)**, with the intention of providing additional searches related to the one the user is executing so as to provide more information around the subject in question.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

Requirement EU-12: Language options: *Multilingual search was deemed useful by interviewees, but in practice most data is still searched in English. This requirement is therefore optional. That said, support for multiple languages could be very useful when exploring insights drawn from social media and online forums.*

The functional requirement here is to be able to **support multiple languages, in the user interface labelling, searching and analyses (F-26).**

2.2 UK Requirements

2.2.1 High priority

Requirement UK-1: Aiding an assessment of the trustworthiness of sources *– understanding the trustworthiness of sources is mostly based on the researcher's judgement. A tool that could help this process by providing additional upfront information on key elements of these decisions, e.g. peer reviewed, most regularly cited or availability of primary datasets would be useful. Assessing the trustworthiness of sources was a problem common to most of the end users we spoke to and there is no tool currently in use that sufficiently addresses this problem. Additionally, encouraging users to engage with new data sources may be difficult without supporting information as to its veracity and value.*

From this requirement, it is clearly important that the user can make a decision about whether they trust a source data set or not, and they must be assisted by the tool. The tool cannot make recommendations or decisions about the reliability of a data set or source, but it can present relevant information about a data set to the user to help them make a judgement about the believability of a data set. It is expected that the information the user will want to see regarding believability will be related to a data set's provenance – its origins and processing, e.g. the creator of the data set, the methods used to gather the data, any processing, who did processing, the data set's publisher, or whether the data is accredited by any official organisations etc. Therefore, the functionality required has already been identified previously: to **present to the user all available information about the provenance of some data that has been found by the tool (F-9)**.

Requirement UK-2: Multiple filters *– multiple filters would be useful to quickly sift through different types of information and to allow different trust factors to be included in search results, as there are many different definitions of what makes a source trustworthy. Drawing on the trust issues in requirement 1, multiple filters were discussed as a way to personalise the information presented, making the tool applicable to large numbers of end users even if they prioritise different information.*

The functionality required here is clear and already identified in EU-3 **enabling the user to filter search results - include or exclude results according to a wide range of criteria (F-8).** It is expected that filters will be generally useful in viewing data from the Sense4Us tool, so it is unlikely they should be limited to aspects concerning whether the data source is trusted or not. Given this, it should be possible to add new filtration criteria as they arise.

Requirement UK-3: Transparency of the tool *– for the tool to be trusted, respondents were clear that they would need to understand the methodology the tool was using to present information and outcomes and be able to see all source data, plus any underlying assumptions that had been made. This was discussed primarily in relation to the simulation tool. Researchers in particular were concerned that they needed to be able to understand how the*

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

tool worked and what assumptions it was based on, as they would not be able to have confidence in it otherwise. Without this transparency, there is a significant risk that engagement with the Sense4us tool by end users will be poor.

This requirement is very similar to requirement EU-4, which discussed the need for transparency in the policy modelling component to enable users to understand and therefore have a greater chance of believing the results. Therefore the major functional requirements discussed as a result of EU-4 are relevant here also, namely:

- **easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)** (this includes the policy modelling);
- **transparency in the creation and operation of the policy models (F-13)**;
- **present to the user all available information about the provenance of some data that has been found by the tool (F-9)** for input to the models; and
- **the policy maker should be able to examine a policy model and to add new factors if necessary (F-16)**.

Requirement UK-4: Searching multiple information types *– a tool that could search different types of information, e.g. academic articles, government statistics and open data. Similarly to requirement 2, including a range of different information types that are handled by Sense4us allows the tool to be customisable and applicable to a wider range of users.*

The function required reinforces those from EU-1: to be able to **search different data formats and interface to multiple data sources (F-1)**. The same caveats discussed in EU-1 apply here, mainly that it is impossible to be exhaustive about the different sources within the project so the functional enabling the user to **add new data sources (F-2)** and to **enable the user to add specific data sources and websites as search targets (F-3)** are also applicable here.

It is not mentioned explicitly in this requirement, but it is reasonable to assume that the additional function **to enable searching multiple data sources with the same query and mouse click from the user (F-6),** mentioned in EU-2, would also be beneficial here.

Requirement UK-5: Visualizing raw datasets *– any data that is contained in spreadsheets or is designed to be machine read rather than read by a human will need to be suitably presented (with summaries/commentaries etc.) in order to be accessible to end users. Data from the survey, interviews and focus groups suggest that visualisation was important for all types of end users because without it key data types, such as open data, are impenetrable to most people.*

The specific functionality required here is to be able to **provide summaries or commentaries of raw data sets (F-27)**, but more generally, the requirement requests that **visualisation of data sets (F-28)** be provided to enable policy makers interpret data sets.

Requirement UK-6: Segmenting social media discussions *– understanding who the users of social media are that are involved in discussions, was seen to be lacking in other social media analysis tools and would therefore be an important addition.*

When the requirement says “who the users are”, it is not clear whether the actual identities of the social media users are required. It is suggested that this is not the case due to data protection regulation requirements for handling personal data and to avoid the perception of surveillance of citizens by governments (discussed in requirement UK-7 below). The

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

alternative is to use aggregation methods to elicit characteristics and features of users who comment on policy-related issues. Given this, the functionality required is assumed to be **to discover any publicly available characteristic information about the social media participants who are expressing the opinions (F-19)** (also from EU-6).

Requirement UK-7: Privacy when using social media data *– there was a concern from decision-makers that citizens giving their views on public forums and social media hadn't given their consent for their comments to be used by Parliament or policy-makers. It is therefore important that we make it clear which data has originated from social media so that end users wishing to avoid using this information have the opportunity to do so. Ensuring we provide end users with information about how Sense4us collects social media data is also important. Without understanding how the project collects, uses and stores data, end users are likely to be nervous about using the tool and may not engage with the development of future prototypes.*

The basic requirement is to respect the privacy of citizens, and to avoid the possibility of the policy maker being connected with any perception of surveillance. Therefore, the requirement is about highlighting places where information has been gathered without the consent of the subjects (even though it may be publicly available) so policy makers can make an informed choice whether to use this data or not. The case in point is social media. It is not practical to seek the consent of all social media users whose opinions we wish to consider because of the high numbers of citizens potentially using social media to discuss subjects relevant the policy makers' interest, so therefore social media data is likely to be collected without the subjects' consent.

Given this and the requirement to be able to avoid using social media data, the user needs to be informed about the origin of data collected by the tool. Therefore the functionality required here is to **present to the user all available information about the provenance of some data that has been found by the tool (F-9)** (from EU-3), and to **highlight any data from social media (F-29)**. Additionally, **the data origin should be a filtration criterion (F-30)**, so the user has the option to filter social media-derived data out.

This requirement backs up the requirement to record and display provenance information **(F-9)**. In general, all Sense4Us data collections should record as much source, metadata and provenance information as is supplied, and make it available. This not only helps users to make a choice about protecting the privacy of social media users, it also allows them to make trust decisions based on the origins and processing of the data.

Requirement UK-8: An assessment of the quality and comparability of data *– in order to use different data sets and information to build up a picture of policy consequences, the comparability and underlying quality of these data sets needs to be assessed and certified by the project or the outcomes may be inaccurate or misleading. This relates to requirement 3 that deals with transparency – if end users do not understand how data has been selected and how the quality of the input data has been assessed they are unlikely to use or trust the tool.*

This requirement actually contains two requirements. It is not clear that these two requirements are related or dependent on each other in any way, although both are useful when assessing data sets, either individually or together.

- Comparability of multiple data sets. As a functional requirement, this can be stated as **determination of whether and how multiple data sets can be compared (F-31)**, for

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

example how different elements in the data sets can be regarded as a like-for-like comparison.

- Assessment of the quality of a data set. This is a subjective judgement made by the recipient as to whether the data set is fit for their purposes. This can be dependent on many factors, and only the recipient can determine whether it is fit for their purposes. Sense4Us can assist them in making this judgement by providing as much information as possible about the origins and processing of the data. Having said this, Sense4Us can only provide as much information as already exists in the data, or can be determined by processing within Sense4Us. Therefore the functional requirement is to **present to the user all available information about the provenance of some data that has been found by the tool (F-9)** so the user can make their own judgement about the quality of the data.

2.2.2 Medium priority

Requirement UK-9: Links to global data *– it can be difficult to find and use policy research and knowledge from other jurisdictions. Links to global data sets could provide valuable insight into which policies are successful in other parts of the world. Understanding what initiatives work in other parts of the world is an important part of evidence gathering for policy-makers, therefore a tool that can discover related datasets from outside the UK could be very useful.*

The functional requirement is to **find documents and data sets related to the policy element in question from countries outside the native country of the user (F-32)**. A more specialised version of the requirement is to be able to search for related policies and their consequences in different countries. The user benefit is to learn from others' experience: to find the issues, successes and failures related to their policy issue from abroad.

Requirement UK-10: Training in how to use the tool *– this is to ensure that outcomes are not overstated and limitations of the data are made clear to users. This was discussed primarily in relation to the simulation tool. This is also connected to requirement 3 – those supporting decision-makers were concerned that without training there was a strong risk that the findings of the tool could be overstated and trust in the tool jeopardised.*

The intention here is to enable the users to understand the tools, how to use them, and their limitations. Therefore the function required here is obvious: to provide **easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)**.

Requirement UK- 11: Cross disciplinary search *– a tool capable of searching across issue areas, e.g. health and environment. Similarly to requirement 4, end users felt that evidence used for policy-making or the consequences of policies were rarely limited to one issue area.*

This requirement is similar to EU-2, in that there is a requirement to search across subject discipline areas. Given this, the main requirement from EU-2 applies here: to **enable searching multiple data sources with the same query and mouse click from the user (F-6)**.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

Requirement UK-12: Summaries of results *– although most researchers wanted a brief overview of the information presented to allow them to decide which links might be most relevant very quickly, they also wanted the ability to see key elements upfront e.g. source name and abstract, and have the ability to dig further into the information if they wanted to see more details such as the methodology by which it was generated. The ability to see summarised results of searches and simulations upfront and then be able to go further in to the information would save end users time and increase engagement with the tool.*

The function required is clear from the requirement:

- **present a summary of a search result to the user (F-33),** and
- **present available metadata about a search result to the user (F-34)**.

The user can then use this information to select the results they wish to dig further into. The summary that can be presented will depend on the source data, but could include metadata, any summary information already present in the search result, or summaries could be created by using the Sense4us tools. The key metadata can include provenance information that can be used by the policy maker to help them judge the reliability of the data. Each user may have different preferences about the metadata and summarisation they wish to see, so an additional functional requirement could be to **enable the user to select which metadata they frequently want to see in their search result summaries (F-35)**.

Requirement UK-13: Highlighting related social media discussions *– highlighting related but distinct conversations on social media so new audiences can be found more easily. (See the policy example on electric cars.) Similarly to requirement 6, highlighting related social media discussions would enable end users to find new audiences and expand the number of voices in policy discussions – something that other social listening tools are not currently effective at doing.*

The functional requirement here is to **locate discussions on social media that are related to the policy issue in question (F-36)**. This is similar to the functional requirements of EU-2, with the difference being the target platform is social media. The similarity with EU-2 is to find related subjects to the policy in question, so the functional requirement from EU-2: **to enable finding information for subjects related to those in the policy (F-7)** also applies here.

2.2.3 Low priority

Requirement UK-14: Data comparisons across different time frames *– as Parliament and government often work on different time frames to other organizations, e.g. the parliamentary session or the electoral cycle rather than calendar year or financial year, it was considered problematic that data was not always available for comparison against these different time scales.*

This requirement is about being able to select time ranges from time-related data. For example unemployment figures may be calculated monthly, and the requirement is to be able to select a time range over which to show the unemployment figures. Given this, the requirement is very similar to EU-7, which is about finding data relevant to a particular time period. Given this, the functional requirement here is the same as EU-7: **to enable the user to specify different time ranges for when the data was (or is) relevant in a search (F-20)**. Correlating the time-related data with events may also be helpful – for example, comparing unemployment in the year before or after a change of government.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

Requirement UK-15: Using the most current data *– one of the opportunities for using open data was the ability to see up to date information and to get a better understanding of current rather than historic trends.*

This requirement may be addressed by the following functional requirements.

- **To enable the user to specify a date that determines the earliest creation point of data in search results (F-37)** – this can be a filtration criterion, so the user can select data created after a certain date.
- **In the event of multiple versions of a data set being available, highlight the most recent (F-38)**.

Requirement UK-16: Showing policy consequences across departments *– as with searching across academic disciplines, there is a need for a tool that will show the policy impacts on a number of different government departments rather than within traditionally defined boundaries. This is particularly important given that cross-cutting macro issues are highlighted as being the most challenging to address.*

This requirement seems to be a consequence of the nature of UK government, which seems to have clearly determined departmental divisions. As long as the things that are being impacted are identified, then it should be of no matter which government department they originate from. The key challenge here is how to identify the factors that the policy in question may impact, and this is addressed by other requirements, e.g. **enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17)** from EU-4. Given this, the main functional requirement here is the ability to **match an impacted factor with a government department (F-39)**, so the respective department can be informed and possibly involved. For example, if the level of unemployment is an impacted factor, it should be clear that in the UK this will be relevant to the Department of Work & Pensions.

Requirement UK-17: Customisability *– in order for either the search or simulation tool to be useful, they would need to be capable of being customized and adapted by the end user in order to cope with changing priorities and new input data.*

The functional requirement is clear: provide the **ability for the policy maker user to customise the tool to their preferences (F-40)**.

2.3 DE Requirements

2.3.1 High Priority

Requirement DE-1

German language *- the policy legislative process in Germany is mainly based on information provided in the German language. Therefore Sense4us tools need to work with German words as input and output as well.*

The functional requirement is clear: that **the tool's UI and the analyses in the research partners' components support German as a language (F-41)**. This is a more specific version of F-26, but F-26 is deemed low priority, and here the need to support German as a language

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

is determined as high priority. Because of its high priority, F-41 it is considered a separate requirement.

Requirement DE-2

***Usability** - tools for policy makers need to be easy to use. Therefore the usability is a 'standard' requirement. If tools are complicated or difficult to understand policy makers won't use them at all.*

This requirement has been discussed at both EU and UK levels, i.e. from EU-4 there is: to provide "**easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)**", and this is equally applicable here. In addition there is the aspect of visualisation. It is difficult to be specific about functional requirements for visualisations, as each type of analysis or result will have different visual aspects, but as a general principle, the visualisations should be clear and easy to understand. This principle should be considered an evaluation criterion that should be tested when the end user partners are evaluating Sense4Us.

Requirement DE-3

***Trusted sources** - the biggest 'gap' explaining why online available information is not being used during the legislative process according to our online survey is the challenge posed in identifying trusted sources.*

This requirement has been discussed at both EC and UK levels, and is again reinforced here. The specific functional requirements are:

- **present to the user all available information about the provenance of some data that has been found by the tool (F-9)**; and
- **enable the user to restrict search results to those that they trust (F-10)**.

Requirement DE-4

***Filter information** - the second biggest 'gap' as to why online available information is not being used during the legislative process is that there is too much information on the Web.*

Filtration is another requirement that appears as high priority at the EC and UK levels. The specific functional requirement is about **enabling the user to filter search results - include or exclude results according to a wide range of criteria (F-8).**

There is a subtlety in the requirement that may mean the requirement is about more than filtration of information – that is that there is too much information on the web, and the most relevant information is difficult to find. To address this, specific searches and filtration can be used, so that the user is shown information relevant to their policy topic. The functional requirements arising from this are therefore:

- **construction of searches based on the policy themes in question (F-42)** and then filtration or
- **ranking of search results using different criteria (F-43)** so that the most relevant to the policy in question are most prominently placed for the user.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

2.3.2 Medium Priority

Requirement DE-5

Extracting opinions from Social Media and blogs – *besides 'fact-based' information from news services (see preliminary interviews), there is a need to access political discussions, which the interviewees currently locate via blogs and on Social Media platforms. Therefore the analysis of social discussions concerning particular topics is relevant for the interviewees. The Senate Chancellery Hamburg has observed a strong increase in the relevance of blogs and Social Media in political communication. Public opinion polling is not being used. But Social Media is the main method for engaging the public online.*

This requirement reinforces retrieving relevant discussions from social media, and the specific functional requirements are:

- **enable opinions from social media and other forums to be gathered (F-5)**; and
- **to be able to acquire information from social media on the topics and themes of the policy in question (F-18)**.

The other functional requirement is the ability to **monitor blogs and comments on them (F-44)**.

Requirement DE-6

Cross-disciplinary search – a tool capable of searching across issue areas, e.g. health and environment. *Decision makers are not limited to policies within their topic areas. They are following and researching topics that are broadcast by the news and media in parallel. Also, information across policy areas and sectors is of interest for policy makers beyond the policy issues for which they are responsible.*

This requirement is a combination of functional requirements that have been expressed at the EC and UK level:

- This requirement reinforces UK-11, which in turn reinforces EU-2. Therefore main requirement from EU-2 applies here: to **enable searching multiple data sources with the same query and mouse click from the user (F-6)**.
- This requirement also discusses searching for information in the broadcast media, which reinforces the requirement **enable traditional online media (e.g. news sites) to be searched (F-4)** from EU-1.
- This requirement also discusses searching for information from areas related to and around the actual policy topic or issue. This reinforces the functional requirements:
    - **enable finding information for subjects related to those in the policy (F-7)** from EU-2 and
    - **enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17)** from EU-4.

2.3.3 Low Priority

Requirement DE-7

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

Searching and analyzing information provided by the Media – *the interviewees mentioned that services from the biggest news websites are used, because these provide trusted and qualitative information. The Senate Chancellery Hamburg mentioned that Internet services that are provided by the media are used during the legislative process. This data could be used as input for the Sense4us tools to support the search and analysis of information provided by the media.*

This reinforces the functional requirement to "**enable traditional online media (e.g. news sites) to be searched (F-4)"** from EU-1, and mentioned above in DE-6. There is also an aspect of how the data sources can be trusted in this requirement, and so the functional requirements for assisting the user in the evaluation of data sources regarding trust also apply – this is mainly **present to the user all available information about the provenance of some data that has been found by the tool (F-9)** from EU-3.

2.4 Verification of Functional Requirements with End User Partners

Once the functional requirements are identified, they were presented to the end user partners (the authors of D2.1) so as to determine whether the functionality interpreted is clear, sensible and correct.

As a result of this, the three end user partners confirmed that the functional requirements were reasonable and sensible. The major comment from the end user partners was that there was some repetition in the functional requirements, mainly where one functional requirement was a more specialised version of another.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

3 Prioritisation of Functional Requirements

The functional requirements were prioritised next, and this is so that the most important are identified and can be dealt with accordingly.

Firstly, the end user partners were also asked to identify any critical functional requirements – this is functionality without which the toolbox is unusable in their opinion. The reason for this is so that we can ensure these requirements are prioritised as essential and are not missed.

Next, the functional requirements were prioritised according to the priority the end user partners assigned their end user requirements. Since functional requirements derive from end user requirements, a functional requirement can inherit the priority of the end user requirements it came from. To account for the high, medium and low priorities assigned to the user requirements in D2.1, we adopted a simple scoring method whereby each of the end user requirements in D2.1 is assigned a numerical value according to its priority. We used:

- high priority = 3,
- medium priority = 2, and
- low priority = 1.

Whenever a functional requirement appears in an end user requirement, it is given the score of the end user requirement. If a functional requirement appears in more than one end user requirement, the scores of the end user requirements are added together. This can be considered as the number of "votes" for each functional requirement coupled with the weight of each vote. For example, if the same functional requirement is in a two high priority requirements and one low priority requirement, its total score is 7 – twice the score of 3 for the high priority requirements, plus 1 for the low priority requirement. All functional requirements were scored in this way, and the overall priority is determined by ranking the functional requirements with the highest scoring ones at the top.

To verify that the prioritisation was sensible, the prioritised list was shown to the end user partners, and they were asked if the prioritisation was sensible. In the event, all end user partners said the prioritisation was sensible.

The result of this step is a set of functional requirements (some of which are identified as critical) and ranked in priority order, and Table 1 on the following pages denotes the final priority order after verification by the end user partners. The table spans multiple pages, and the header rows are repeated at the top of each page. The table is ordered so the highest priority requirements are at the top[3]. The table is ranked by the so-called "Bayesian Average"[4]. This is an average that takes into account how many "votes" there are for a particular functional requirement – the more votes there are, the surer we are that the average computed is correct. It guards against situations like:

- Functional requirement X: average 3, total votes 1
- Functional requirement Y: average 2.93, total votes 20

In effect, there is little difference in the ranking positions when scored by the Bayesian average compared to the sum of all the scores, and also little difference between the ranking positions when the Bayesian average is compared to the standard (mean) average.

[3] For clarity, critical functional requirements are not highlighted in this table. The critical functional requirements are highlighted later in the tables in Section 6.

[4] See http://fulmicoton.com/posts/bayesian_rating/ for full explanation.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

		User Requirement ID																																										
		EU-1	EU-2	EU-3	EU-4	EU-5	EU-6	EU-7	EU-8	EU-9	EU-10	EU-11	EU-12	UK-1	UK-2	UK-3	UK-4	UK-5	UK-6	UK-7	UK-8	UK-9	UK-10	UK-11	UK-12	UK-13	UK-14	UK-15	UK-16	UK-17	DE-1	DE-2	DE-3	DE-4	DE-5	DE-6	DE-7	DE-8				
		H	H	H	H	H	M	M	M	M	M	L	L	H	H	H	H	H	H	H	H	M	M	M	M	M	L	L	L	L	H	H	H	H	M	M	L	L	Priority			
		3	3	3	3	3	2	2	2	2	2	1	1	3	3	3	3	3	3	3	3	2	2	2	2	2	1	1	1	1	3	3	3	3	2	2	1	1	Priority Score	Overall Average Score		4.32
																																							Average Num Votes			1.77
F-ID	Functional Requirement																																						Sum	Average	Num Votes	Bayesian Avg
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)			1	1									1		1				1	1												1				1		22	2.75	8.00	18.79
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)				1											1							1									1							11	2.75	4.00	8.95
6	Enable searching multiple data sources with the same query and mouse click from the user (F-6)		1														1							1												1			10	2.50	4.00	8.26
17	Enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17)				1	1																							1							1			9	2.25	4.00	7.56
8	Enable the user to filter search results - include or exclude results according to a wide range of criteria (F-8)			1											1																			1					9	3.00	3.00	7.26
3	Enable the user to add specific data sources and websites as search targets (F-3)	1									1						1																							8	2.67	3.00	6.63
5	Enable opinions from social media and other forums to be gathered (F-5)	1					1																													1			7	2.33	3.00	6.00
7	Enable finding information for subjects related to those in the policy (F-7)		1																							1										1			7	2.33	3.00	6.00
4	Enable traditional online media (e.g. news sites) to be searched (F-4)	1																																		1	1		6	2.00	3.00	5.38
1	Search different data formats and interface to multiple data sources (F-1)	1															1																							6	3.00	2.00	5.21
2	Add new data sources (F-2)	1															1																							6	3.00	2.00	5.21
10	Enable the user to restrict search results to those that they trust (F-10)			1																													1						6	3.00	2.00	5.21

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

		User Requirement ID																																										
		EU-1	EU-2	EU-3	EU-4	EU-5	EU-6	EU-7	EU-8	EU-9	EU-10	EU-11	EU-12	UK-1	UK-2	UK-3	UK-4	UK-5	UK-6	UK-7	UK-8	UK-9	UK-10	UK-11	UK-12	UK-13	UK-14	UK-15	UK-16	UK-17	DE-1	DE-2	DE-3	DE-4	DE-5	DE-6	DE-7	DE-8				
		H	H	H	H	H	M	M	M	M	M	L	L	H	H	H	H	H	H	H	H	M	M	M	M	M	L	L	L	L	H	H	H	H	M	M	L	L	Priority			
		3	3	3	3	3	2	2	2	2	2	1	1	3	3	3	3	3	3	3	3	2	2	2	2	2	1	1	1	1	3	3	3	3	2	2	1	1	Priority Score	Overall Average Score		4.32
																																							Average Num Votes			1.77
F-ID	Functional Requirement																																						Sum	Average	Num Votes	Bayesian Avg
13	Transparency in the creation and operation of the policy models (F-13)				1											1																							6	3.00	2.00	5.21
16	The policy maker should be able to examine a policy model and to add new factors if necessary (F-16)				1											1																							6	3.00	2.00	5.21
19	Discover any publicly available characteristic information about the social media participants who are expressing the opinions (F-19)						1												1																				5	2.50	2.00	4.68
18	Acquire information from social media on the topics and themes of the policy in question (F-18)						1																												1				4	2.00	2.00	4.15
11	Enable the user to sort data by different criteria (F-11)			1																																			3	3.00	1.00	3.84
14	Enable the searching for scientific information relevant to the policy subjects (F-14)				1																																		3	3.00	1.00	3.84
15	Enable the collection of sentiments / opinions about the policy subjects (F-15)				1																																		3	3.00	1.00	3.84
27	Provide summaries or commentaries of raw data sets (F-27)																	1																					3	3.00	1.00	3.84
28	Visualisation of data sets (F-28)																	1																					3	3.00	1.00	3.84
29	Highlight any data from social media (F-29)																			1																			3	3.00	1.00	3.84
30	The data origin should be a filtration criterion (F-30)																			1																			3	3.00	1.00	3.84
31	Determination of whether and how multiple data sets can be compared (F-31)																				1																		3	3.00	1.00	3.84
41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)																														1								3	3.00	1.00	3.84
42	Construction of searches based on the policy themes in question (F-42)																																	1					3	3.00	1.00	3.84

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

| | | User Requirement ID | | |

| | | User Requirement ID | | |

4 Allocation of Functional Requirements to Sense4Us Research and Development Work

The next step in the process is to allocate the functional requirements to components of Sense4Us. This is for a number of reasons:

- To test our initial approach, described in the project proposal and the Description of Work, and to identify any modifications to it necessary to address the functional (and therefore end user) requirements. Specifically this allocation should highlight gaps: it will determine whether there is any functionality required that is not covered by at least one component.
- To determine a basic system design from the components. This will be input to the main architecture work later on in the project.
- To determine which partners should be responsible for addressing the functional requirements. Some functional requirements will appear in more than one group, and this indicates that the responsibility is divided across the components.

The initial approach broadly follows the project's workpackage structure and partners' responsibilities. From this approach, we defined an initial set of components:

- Search (Semantic, LOD, other etc) – U Koblenz, WP4
- Social Media Analysis – KMI, WP5
- Policy Modelling – U Stockholm, WP6
- Topic Analysis – U Koblenz, WP4
- Presentation / Visualisation of Results – U Koblenz, KMI, U Stockholm, IT Innovation, WP4, WP5, WP6, WP7
- User Environment / Interface – IT Innovation, WP7

It was decided to give each project partner an opportunity to have a say about the assignment of functional requirements to components. The reason for this decision is that the effect of the allocation is to determine the work the research partners need to do in the project, and it was thought democratic and reasonable to give each partner a fair say about the allocations. Also, the project has distinct groups of partners and all groups need to be represented - the research partners have their own research agendas to follow, but the project needs to strike a balance between this, the needs of end users and the project's goals and ambitions. To take all these factors into account, it was decided to ask each partner for their opinion about component allocation. Although it was not a "secret ballot" as such, each partner gave their opinions independently so that one partner's opinion did not influence another's.

To collect partners' votes, a voting form was created containing the ordered list of functional requirements combined with the components. An example of a filled voting form is shown in Figure 1.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

		Categories						
		Search (Semantic, LOD, other etc)	Social Media Analysis	Policy Modelling	Topic Analysis	Presentation / Visualisation of Results	User Environment / Interface	(Other)
	IT Innovation							
F-ID	**Functional Requirement**							
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	x				x		
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)							x
6	Enable searching multiple data sources with the same query and mouse click from the user (F-6)	x						
17	Enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17)	x						
8	Enable the user to filter search results - include or exclude results according to a wide range of criteria (F-8)					x		
3	Enable the user to add specific data sources and websites as search targets (F-3)	x						
5	Enable opinions from social media and other forums to be gathered (F-5)	x	x					
7	Enable finding information for subjects related to those in the policy (F-7)	x			x			
4	Enable traditional online media (e.g. news sites) to be searched (F-4)	x				x		
1	Search different data formats and interface to multiple data sources (F-1)	x				x		

Figure 1: Example Allocation Voting

Each partner was asked to mark which components they thought should address functional requirements. As can be seen from Figure 1, it is possible that a functional requirement can be assigned to more than one component, and so each partner was asked to allocate a requirement to as many components as they saw fit. As an example, there are many cases where a combination of aspects of the research partners' work coupled with visualisation of results is necessary to address a functional requirement.

It can be seen in the figure that there is an "other" column. The intention was to give the partners an opportunity to indicate another component if they felt that the components specified already were not sufficient to address the functional requirements. In the event, the

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

only use for this "other" column was as a space for occasional comments, and to indicate that documentation was required.

Once all partners had cast votes, the results were collated to determine the actual allocation of functional requirements to components. An illustration of the collation's output is shown in Figure 2.

						Categories					
						Search (Semantic, LOD, other	Social Media Analysis	Policy Modelling	Topic Analysis	Presentation / Visualisation of	User Environment / Interface
END USER REQUIREMENT PRIORITY											
Average	Num Votes	Bayesian Avg	CRITICAL ?	F-ID	Functional Requirement						
2.75	8.00	18.79	Y (HS)	9	Present to the user all available information about the provenance of some data that has been found by the	5	4	4	2	5	5
2.75	4.00	8.95	Y (HS)	12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of	4	4	5	4	5	4
2.50	4.00	8.26	Y (HS)	6	Enable searching multiple data sources with the same query and mouse click from the user (F-6)	6	2	1	1	3	5
2.25	4.00	7.56	Y (HS)	17	Enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17)	6	2	2	2	2	2
3.00	3.00	7.26	Y (HS)	8	Enable the user to filter search results - include or exclude results according to a wide range of criteria (F-8)	4	2	1	1	6	5
2.67	3.00	6.63		3	Enable the user to add specific data sources and websites as search targets (F-3)	5	3	1	3	2	4

F-ID	Votes Distribution
9	
12	
6	
17	
8	
3	

Figure 2: Collation of Allocations

The figure shows the functional requirements in rows (the yellow ones are marked as critical, with an indication of the partner indicating criticality), and the components (entitled "Categories") in columns. The number of votes for allocation of each component to functional requirement is given in the intersection between these rows and columns, and the intersecting cells are colour-coded to show the number of votes – grey is the lowest and red is the highest.

Given that most intersections had at least one vote, the question now was how to determine the actual allocation? It is not sensible to have all components assigned to all functional requirements, but the most important components must be assigned to a functional requirement. It was decided that two key principles should determine the actual allocation:

1. Where the distribution of votes is similar across components, most components should address the requirement. (Examples are F-9 and F-12 in the figure – the histogram at the right shows fairly evenly distributed votes.)
2. Where the distribution of votes is different across components, only those components with the most votes should address the requirement. (An example in the figure is F-17 – Semantic Search is the clear winner here.)

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

The main reasons for these principles are that we need to be inclusive when the vote is evenly split because the decision is that most components should be partially responsible for addressing a requirement. Conversely, when there is a clear decision, we should only include the components with the most votes (the clear winners).

The mechanism chosen to achieve the allocation is very simple – to determine an acceptance criterion for components' votes in a requirement that is based on a percentage of the maximum across all components in a requirement. If the votes for a component are greater than the percentage, then the component should address the requirement. An example of this is shown in Figure 3.

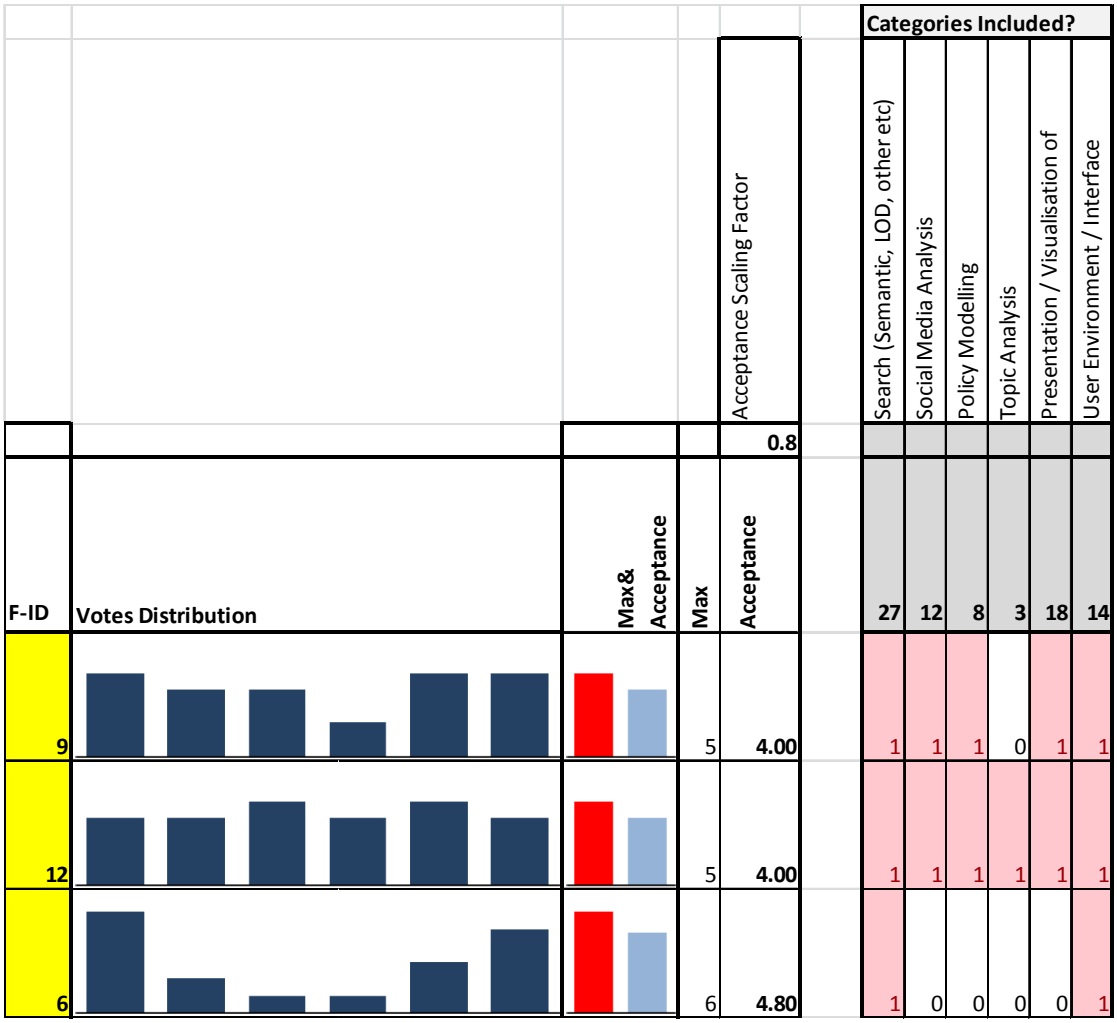

Figure 3: Example acceptance

This figure shows the votes distribution of the top three functional requirements (reprising the right hand side of Figure 2). The maximum number of votes is shown in red, and the acceptance level is shown in light blue. The columns to the right show which components are allocated to a functional requirement as a result of the acceptance criterion (a filled-in pink box with a "1" inside indicates allocation).

In the example, we can see that by choosing an acceptance scaling factor of 80% (0.8 in the figure), functional requirements with evenly distributed votes (e.g. F-9 and F-12 above) have most (or all) of the components included, while those with a greater variance in the voting (e.g. F-6 above) have the components with the most votes included.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

The scaling factor was chosen by visual experimentation. Different values of the scaling were tried (working from 1.0 downwards) and when the components chosen to address a functional requirement were consistent with the two key allocation principles above, the scaling factor was deemed correct. Based on this, a factor of 0.8 (illustrated above) was chosen. This technique gave an allocation of components to functional requirements as shown in Table 2 below, with a pink box indicating that the component is in scope for the requirement.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

			Category	Search (Semantic, LOD, other etc)	How to address	Social Media Analysis	How to address	Policy Modelling	How to address	Topic Analysis	How to address	Presentation / Visualisation of Results	How to address	User Environment / Interface	How to address
			Partner Responsible	UKOB		KMI		U STOCK		UKOB		ITI		ITI	
			(Checksum)	0		0		0		0		0		0	
END USER REQUIREMENT PRIORITY															
Bayesian Avg	CRITICAL ?	F-ID	Functional Requirement	27		12		8		3		18		14	
18.79	Y (HS)	9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	1		1		1		0		1		1	
8.95	Y (HS)	12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)	1		1		1		1		1		1	
8.26	Y (HS)	6	Enable searching multiple data sources with the same query and mouse click from the user (F-6)	1		0		0		0		0		1	
7.56	Y (HS)	17	Enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17)	1		0		0		0		0		0	
7.26	Y (HS)	8	Enable the user to filter search results - include or exclude results according to a wide range of criteria (F-8)	0		0		0		0		1		1	
6.63		3	Enable the user to add specific data sources and websites as search targets (F-3)	1		0		0		0		0		1	
6.00	Y (HS)	5	Enable opinions from social media and other forums to be gathered (F-5)	0		1		0		0		0		0	
6.00		7	Enable finding information for subjects related to those in the policy (F-7)	1		1		0		0		0		0	
5.38		4	Enable traditional online media (e.g. news sites) to be searched (F-4)	1		0		0		0		1		0	
5.21	Y (HS)	1	Search different data formats and interface to multiple data sources (F-1)	1		0		0		0		0		0	
5.21	Y (HS)	2	Add new data sources (F-2)	1		0		1		0		0		1	
5.21	Y (HS)	10	Enable the user to restrict search results to those that they trust (F-10)	1		0		0		0		1		1	
5.21	Y (HS)	13	Transparency in the creation and operation of the policy models (F-13)	0		0		1		0		0		0	
5.21	Y (HS)	16	The policy maker should be able to examine a policy model and to add new factors if necessary (F-16)	0		0		1		0		0		0	
4.68	Y (HS)	19	Discover any publicly available characteristic information about the social media participants who are expressing the opinions (F-19)	0		1		0		0		0		0	
4.15		18	Acquire information from social media on the topics and themes of the policy in question (F-18)	0		1		0		0		0		0	

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

			Category	Search (Semantic, LOD, other etc)	How to address	Social Media Analysis	How to address	Policy Modelling	How to address	Topic Analysis	How to address	Presentation / Visualisation of Results	How to address	User Environment / Interface	How to address
			Partner Responsible	UKOB		KMI		U STOCK		UKOB		ITI		ITI	
			(Checksum)	0		0		0		0		0		0	
END USER REQUIREMENT PRIORITY															
3.84		11	Enable the user to sort data by different criteria (F-11)	0		0		0		0		1		0	
3.84		14	Enable the searching for scientific information relevant to the policy subjects (F-14)	1		0		0		0		0		0	
3.84	Y (HS)	15	Enable the collection of sentiments / opinions about the policy subjects (F-15)	0		1		0		0		0		0	
3.84	Y (HS)	27	Provide summaries or commentaries of raw data sets (F-27)	1		0		0		0		1		0	
3.84	Y (HS)	28	Visualisation of data sets (F-28)	0		0		0		0		1		0	
3.84		29	Highlight any data from social media (F-29)	0		1		0		0		1		1	
3.84		30	The data origin should be a filtration criterion (F-30)	1		0		0		0		1		0	
3.84	Y (HS)	31	Determination of whether and how multiple data sets can be compared (F-31)	1		0		0		0		1		0	
3.84	Y (GESI	41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)	1		1		1		1		1		1	
3.84		42	Construction of searches based on the policy themes in question (F-42)	1		0		0		1		0		0	
3.84		43	Ranking of search results using different criteria (F-43)	1		0		0		0		1		1	
3.62		20	Enable the user to specify different time ranges for when the data was (or is) relevant in a search (F-20)	1		1		0		0		1		0	
3.48	Y (HS)	21	Provide user accounts, with storage and retrieval of data and workflows all within a user's account (F-21)	0		0		0		0		0		1	
3.48	Y (HS)	22	Locate information related to the topic in question, but restricted to specified countries (F-22)	1		0		0		0		0		0	
3.48		23	Provide a facility to assess the impact of a policy decision (F-23)	0		0		1		0		0		0	
3.48	Y (Gov2	24	Ability to use data from both the European Commission and the European Parliament (F-24)	1		0		0		0		0		0	
3.48	Y (HS)	32	Find documents and data sets related to the policy element in question from countries outside the native country of the user (F-32)	1		0		0		0		0		0	
3.48		33	Present a summary of a search result to the user (F-33)	1		0		0		0		1		0	
3.48		34	Present available metadata about a search result to the user (F-34)	1		0		0		0		1		0	
3.48		35	Enable the user to select which metadata they frequently want to see in their search result summaries (F-35)	0		0		0		0		1		0	
3.48		36	Locate discussions on social media that are related to the policy issue in question (F-36)	0		1		0		0		0		0	

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

			Category	Search (Semantic, LOD, other etc)	How to address	Social Media Analysis	How to address	Policy Modelling	How to address	Topic Analysis	How to address	Presentation / Visualisation of Results	How to address	User Environment / Interface	How to address
			Partner Responsible	UKOB		KMI		U STOCK		UKOB		ITI		ITI	
			(Checksum)	0		0		0		0		0		0	
END USER REQUIREMENT PRIORITY															
3.48		44	Monitor blogs and comments on them (F-44)	0		1		0		0		0		0	
3.12		25	Augment search terms the user has specified and perform additional and related searches to the ones the user is running (F-25)	1		0		0		0		0		0	
3.12		26	Support multiple languages, in the user interface labelling, searching and analyses (F-26)	0		0		0		0		0		1	
3.12		37	Enable the user to specify a date that determines the earliest creation point of data in search results (F-37)	1		0		0		0		0		1	
3.12		38	In the event of multiple versions of a data set being available, highlight the most recent (F-38)	1		0		0		0		1		0	
3.12		39	Match an impacted factor with a government department (F-39)	1		0		1		0		0		0	
3.12	Y (HS)	40	Ability for the policy maker user to customise the tool to their preferences (F-40)	0		0		0		0		0		1	

Table 2: Allocation of Functional Requirements to Components

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

5 Determine If and How Functional Requirements Can be Addressed

Once the allocation to components was complete, the allocated requirements for each component were distributed to the research partners (U Stockholm, OU and U Koblenz) and the integrating partner (IT Innovation) to determine if and how they can be addressed. An individual Excel sheet based on Table 2 was created for each of the research and integrating partners, and the components and functional requirements applicable were highlighted. These were distributed to the research and integrating partners, who filled in all the highlighted sections. For each highlighted section, the responsible partner briefly described how they can address the requirements assigned to their component, or if they cannot address a requirement, why. An example of a response is shown in Figure 4.

	Category	Presentation / Visualisation of Results	How to address
	Partner Responsible	**ITI**	
		0	
F-ID	**Functional Requirement**	**18**	
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	1	All available information regarding the source / origins / publisher / processing of any data will presented to the user. Some of this may be generated outside Sense4us, and in other cases may be generated within Sense4us. Where the provenance information comes from outside Sense4us, this will be shown, but obviously Sense4us is limited by the information it is given. When any data is processed within Sense4us, the entire processing chain will be shown the user, and the user can see the data at each stage of Sense4us processing. All "intermediate" data will be stored in the Sense4us database, and will be accessible to the user.
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)	1	Sense4us will provide user documentation, as well as introductory presentations and tutorials.

Figure 4: Example "how to address functional requirement" response

The result of this was a statement for each component, by its main responsible partner, about how the component is able to address the functional requirements applicable to the component.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

6 Collate Functional Specification

The final step is to collate the responses from the research and integrating partners about how to address requirements into a functional specification for each component in the system. The answers from the project partners on how to address functional requirements were inspected, and this section contains a set of tables for each of the main components[5].

The tables show the functional requirements, how the research partners will address the requirements, a summary of the functionality, and (where applicable) the input and output components and data. There is also a classification, which is a first attempt at identifying common functionality across components. It is not yet known whether this will be useful, and this will be investigated in the work leading up to D3.2, where the initial system architecture is specified.

Some rows in the tables are marked with colours, and these correspond to whether and how much the requirement will be addressed. This is described in the “Address?” column, and which has the values:

- "N" = No (Red) - reasons for why the requirement will not be addressed are given,
- "N/A" = Not Applicable (grey italics),
- “D” = Deferred (purple) - more information or discussion is required to address the requirement.
- "R" = Redundant (pink) - repeats another functional requirement,
- "P" = Partially (grey),
- “New” = New Requirement (orange) - this is a requirement that is new in a section (this is usually the result of moving the requirement from another section marked “N/A”, “N” etc).

If there is no classification in the “Address” column, the row is coloured white, and requirement will be addressed

These tables form a functional specification of each major component in the Sense4us system, and are input to the System Architecture Design task, T3.2, where the detailed system design will take place.

[5] Some tables span multiple pages, and a single component is regarded as one table, even though it may span more than one page.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

6.1 Policy Modelling

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Policy Modelling											
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	Y (HS)	18.79		1	Visualisation of the causal diagram(s) for the Policy issue; producing data reports with a checklist for the problem elements including key actors, stakeholders, variables with the available statistical data; producing impact assessment reports for the policy options; Visualisations of the simulation results using simple and accessible graphics.	Provenance information in showing how results were derived: evidence backing up variables' values, user can see model's construction and view intermediate data in the model	Policy Modelling		UI	Evidence + Visualisation of model	Transparency / Visualisation / Provenance
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)	Y (HS)	8.95		1	A documentation for the simulation tool, including guiding examples, Video tutorials and written manual (Step-wise use) of the tool. End-user engagement and training as a part of the tool testing and validation stages	Documentation	N/A				Documentation
13	Transparency in the creation and operation of the policy models (F-13)	Y (HS)	5.21		1	Engaging end user in the early stages of building a model for the policy problem, by structuring the problem in terms of the key variables and parameters identified by the user. Allowing the user to define and modify links, policy objectives in terms of targeted changes in outcome indicators. Providing a documentation of the methodology used for the simulation tool	User participation in model building: user can see model's construction, user can update model's links and variables	UI	All input information to model - parameters, values, links	UI	Visualisation of model	Model creation / Transparency / Visualisation
16	The policy maker should be able to examine a policy model and to add new factors if necessary (F-16)	Y (HS)	5.21		1	Visualisation of simulation graph, emphasising underlying assumptions made with respect to causal relationships, flexible modelling GUI.	User can see model's construction, user can update model's links and variables (see also F-13)	UI	Policy model	UI	Updated policy model	Model update / Transparency / Visualisation
2	Add new data sources (F-2)	Y (HS)	5.21	*D*	*1*	*Not quite sure what is meant with "data source".*						
41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)	Y (GESIS)	3.84	P	1	The policy modelling tools are in general language agnostic - language support is mainly concerned with the UI functionality for the simulation tool	German language support in UI for policy modelling	UI				German language support

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Policy Modelling											
23	Provide a facility to assess the impact of a policy decision (F-23)		3.48		1	Well this simply is what the causal mapping simulation aims to do. So implement a tool exploiting the causal mapping method, simulating the impact on a set of goal variables when changing controllable variables (subject to policy decisions)	Assess the impact of policy changes: showing the effects on dependent variables when input variables change	UI	Values for input variables corresponding to policy options	UI	Consequences on dependent variables	Impact Assessment
39	Match an impacted factor with a government department (F-39)		3.12	*N*	*1*	*This is unlikely to be addressed - different governments have different structures and responsibilities*						

Table 3: Policy Modelling Functional Specification

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

6.2 Presentation / Visualisation of Results

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Presentation / Visualisation of Results											
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	Y (HS)	18.79		1	All available information regarding the source / origins / publisher / processing of any data will presented to the user. Some of this may be generated outside Sense4us, and in other cases may be generated within Sense4us. Where the provenance information comes from outside Sense4us, this will be shown, but obviously Sense4us is limited by the information it is given. When any data is processed within Sense4us, the entire processing chain will be shown the user, and the user can see the data at each stage of Sense4us processing. All "intermediate" data will be stored in the Sense4us database, and will be accessible to the user.	Present provenance metadata to the user: e.g. source / origins / publisher / processing of data	Any	Any data	UI	Provenance metadata	Metadata Presentation /Provenance
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)	Y (HS)	8.95		1	Sense4us will provide user documentation, as well as introductory presentations and tutorials.	Documentation	N/A				Documentation
8	Enable the user to filter search results - include or exclude results according to a wide range of criteria (F-8)	Y (HS)	7.26		1	The user will be able to select filters from a list or add keywords to filter any results coming from Sense4us. Filter examples could be data sources, source types, date of creation, subject keywords etc. Once a filter is selected, the data is displayed filtered by that filter. It is to be decided what happens when a user selects multiple filters - should the data be filtered to show the data with all the filtration criteria, or show data with any of the filter criteria? It is of course possible to have this user-selectable.	Filtration of output using criteria selected by user	Any	Any result type + Filtration criteria	UI	Filtered results	Result Filtration

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Presentation / Visualisation of Results											
4	Enable traditional online media (e.g. news sites) to be searched (F-4)		5.38	*N*	*1*	*It may not be possible to address this requirement, especially if there are copyright restrictions on the content of the news sites to be searched. In addition, this requirement can be addressed by existing Google searches. What may be of interest is to incorporate Google search results with LOD and open data searches conducted by Sense4us.*	*Search traditional media websites*	*UI*	*URLs of websites and search terms*	*UI*	*Search results*	*Standard searching*
10	Enable the user to restrict search results to those that they trust (F-10)	Y (HS)	5.21		1	This is a special case of filtration (see responses to F-8), coupled with provenance information (see responses to F-9). Trusted sources should be a filter category, and the user can add data sources or criteria to this as they wish.	Filtration by trusted sites (See F-8 and F-9)	UI or User profile	Any result type + Filtration criteria (here trusted sources / processors, etc)	UI	Filtered results	Result Filtration / Trust / Provenance
41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)	Y (GESIS)	3.84	*N/A*	*1*	*Not applicable to result presentation: any translation of results is down to the components that generate them.*	*German language support*	*Any*	*Any result data*	*UI*	*Results in German language*	*German language support*
27	Provide summaries or commentaries of raw data sets (F-27)	Y (HS)	3.84	p	1	Where possible, the tool will display summaries of data sets, but it is not possible to guarantee that all data sets will be able to be summarised. Some data sets will come with summaries, and these can be shown to the user (if they are correctly identified). In other cases, Sense4us may be able to provide automatically generated summaries of data (e.g. from documents). In general, each data type and source will need to be examined on a case by case basis to determine whether a summary is present and how it can be identified, or if a summary can be generated.	Present summaries from result metadata	Any	Any result data + Summary metadata	UI	Summary Metadata	Metadata Presentation

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Presentation / Visualisation of Results											
29	Highlight any data from social media (F-29)		3.84		1	This is simply a case of showing the data source to the user and pointing out that it is social media.	Highlight data source if it is social media	Any	Any result data	UI	Highlighted results	Metadata Presentation
31	Determination of whether and how multiple data sets can be compared (F-31)	Y (HS)	3.84	P	1	It is expected that if this requirement is addressed at all, it will be addressed minimally and under tighly constrained conditions. Whether it is possible to address this requirement depends strongly on the data being compared and its compatibility. It is unlikely that automatic detection of compatible data sets will be possible, due to the multitudes of different data formats and semantics that may be encountered. Having said this, some fields may be possible to compare ("date" for example).	Comparison of data sets	Any	Any result data	UI	Compared Results	Result Comparison
30	The data origin should be a filtration criterion (F-30)		3.84		1	this is a special case of filtration (see responses to F-8), coupled with provenance information (see responses to F-9).	Filter data by origin	Semantic Search	Any result data + Source origin metadata	UI	Results filtered by origin	Result Filtration / Metadata Presentation
11	Enable the user to sort data by different criteria (F-11)		3.84		1	Results should be able to be sorted, and the criteria is likely to be many and varied. Results will be sort-able where possible, and this depends on the criteria. It is easy to sort numerically, alphabetically or chronogically, but other types of criteria will be examined on a case by case basis to determine whether it is possible. For example, discrete data sets of categories might only be sorted based on their names' alphabetical ordering, and this may be meaningless to their semantics. Where it is not possible to sort, the UI will indicate this by greying out any "sort" buttons or tools.	Sorting & ranking results	UI	Any result data + Sort Criteria	UI	Sorted results	Result Sorting

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Presentation / Visualisation of Results											
43	Ranking of search results using different criteria (F-43)		3.84	*R*	*1*	*This is the same response as to F-11.*	*Sorting & ranking results*	*UI*	*Any result data + Sort Criteria*	*UI*	*Sorted results*	*Result Sorting*
28	Visualisation of data sets (F-28)	Y (HS)	3.84	P	1	This requirement is very general, as each data set is different and will have different visualisation criteria. In addition, there are many different visualisation types, and some are more suited to some data sets than others. Also, different users prefer to see things in different ways. Each data type encountered will be examined on a case by case basis to determine if it can be visualised, and how it can be visualised. This is potentially a vast task and needs to be made tractable, so it would be a good idea to evolve a standard set of visualisations, based on what is popular with end users, what is possible given the tools available, and what is compatible with the data.	Dataset visualisation (strongly dependent on data type)	Any	Any result data	UI	Visualised results	Visualisation
20	Enable the user to specify different time ranges for when the data was (or is) relevant in a search (F-20)		3.62		1	Restricting search results by date is only possible when the creation / modification date of data sources is known, so therefore some searches may not be possible to date-restrict. This version is in filtration or ranking of search results by date, and this will occur as any other ranking / sort / filtration (see responses to F-11 and F-8). There is another function in the UI section, also addressing F-20.	Restrict search using publication date	UI	Search terms + Date range	Semantic search	Date-restricted search query	Restrict Searches by Metadata
34	Present available metadata about a search result to the user (F-34)		3.48		1	Where available, standard metadata will be presented to the user in search results. See response to F-33.	Present all available metadata to user. Superset of F-9 & F-27	Semantic Search	Result data	UI	Result metatata	Metadata Presentation

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Presentation / Visualisation of Results											
35	Enable the user to select which metadata they frequently want to see in their search result summaries (F-35)		3.48		1	The user profile will contain a section for favourite metadata, and this allows the user to select metadata that will be presented in search results.	User-selectable favourite metadata	UI (user profile)	Metadata favourites	UI	Result metatata	Metadata Presentation
33	Present a summary of a search result to the user (F-33)		4.00		1	The design idea for the result presentation is to show a snippet of a result, alongside standard metata where available (e.g. date created / modified, the source url / analysis, any processing that has taken place, etc). The user can click through any of the search result snippets to see a full-page version of the search result behind the snippet. The type of snippet will depend on the type of search result - e.g. if is a document, a headline and some lines of text can be shown. Where possible sort criteria will also be available to the user so they can sort or rank the results (see F-11).	Result snippet: (summary if possible), provenance, dates, source etc	Any	Any result data	UI	Summary and metadata of results	Summary of results
38	In the event of multiple versions of a data set being available, highlight the most recent (F-38)		3.12	*R*	*1*	*This is only possible is creation / modification dates are available, and if so, the results can be ranked into date order and the most recent selected for presentation to the user.*	*Similar to F-20, ranking by most recent*	*any*	*result data*	*UI*	*ranked results*	*Restrict Searches by Metadata*

Table 4: Functional Specification for Presentation / Visualisation of Results

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

6.3 Search

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Search (Semantic, LOD, other etc)											
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	Y (HS)	18.79		1	VoID descriptions (when available) will be used to access the general metadata of the dataset. Metadata includes the name of dataset, its textual description, publisher, date of publishing, size etc.	Display provenance metadata extracted from data	Semantic Search	Search result data	UI	Provenance metadata	Metadata Presentation
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)	Y (HS)	8.95		1		Documentation	N/A				Documentation
6	Enable searching multiple data sources with the same query and mouse click from the user (F-6)	Y (HS)	8.26		1	User may be given an option to select the data sources from a list of data sources which are preselected for the use case scenario OR finding owl:sameAs links discovered from some registry (e.g., sameas.org). Usually when accessing data about an entity owl:sameAs toward other data sources are provided. For instance, the entity dbpedia:Germany (in the DBpedia sata source) has an owl:sameAs with fb:Germany (in the Freebase data source).	User can select data sources from list. Also auto discovery of data sources.	UI	Options for data sources and user can select the ones they want	Semantic search	Data souces to search	Data Source Selection
17	Enable searching for factors related to their policy subjects using Sense4Us from different data sources (F-17)	Y (HS)	7.56		1	Topic modeling tool box will provide the search terms representing the various factors of the policy and user can use those search terms to find relevant information from different data sources.	Search using terms generated by topic analysis and related terms found by searching relationships in the LOD graph	Topic Analysis	Topic keywords	UI	Related keywords & entities	Find related keywords & entities
3	Enable the user to add specific data sources and websites as search targets (F-3)		6.63		1	The user will be provided with an option to add links to target data sources that she may want to add to search.	User can manually add links to search data sources	UI	Links from user	Semantic search	Links from user	Data Source Selection

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Search (Semantic, LOD, other etc)											
7	Enable finding information for subjects related to those in the policy (F-7)		6.00	*R*	*1*	*Topic modeling tool box will provide the search terms representing the various subjects of the policy. Entity lookup service will be used to find entities from the topics terms. Those entities will be searched in the LOD and a sub-graph containing and connecting those entities will be extracted from the LOD.*	*Topic analysis terms are input, corresponding entities found and LOD searched to find connecting entities. The result is a network graph showing how the input terms are related to the other entities.*	*Topic Analysis*	*topic keywords*	*UI*	*graph with data links related to topic keywords*	
4	Enable traditional online media (e.g. news sites) to be searched (F-4)		5.38	*N*	*1*	*(SJT): whether this is possible depends on if there is an API to the specified media source. We cannot screen scrape for copyright infringement reasons. If there is a supplied API, then we must also abide by the terms and conditions.*						
2	Add new data sources (F-2)	Y (HS)	5.21	*R*	*1*	*Links to the new datasource will be added*	*Similar to F-3*	*UI*				*Data Source Selection*
10	Enable the user to restrict search results to those that they trust (F-10)	Y (HS)	5.21	*R*	*1*	*This feature will be made available by leveraging information about the dataset retrieved from the VoID description. The user can then select only data sources for which she wants to see the results in the final resultset.*	*Similar functionality to F-9, plus the ability to filter*	*Any*	*Provenance metadata + filtration criteria*	*UI*	*filtered data*	*Result Filtration*
1	Search different data formats and interface to multiple data sources (F-1)	Y (HS)	5.21		1	The crawler will look for CSV and Excel files for relevant data. If available the metadata of these types of file will be retrieved for relevance analysis to the policy subject. The relevant files will be retrieved and will be converted to RDF format using tools from GESIS.	Search for CSV and Excel data files and convert to RDF	UI	Search terms + Flag denoting specfic search for CSV and excel files	UI	Ranked RDF representatio ns of data retrieved	Excel + CSV Search
27	Provide summaries or commentaries of raw data sets (F-27)	Y (HS)	3.84	P	1	Metadata information of dataset if available will be used as a summary of the dataset and will be presented to the user.	Present summaries of data using metadata (if available)	Semantic Search	search results	UI	Summaries from result metatata	Metadata Presentaton

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Search (Semantic, LOD, other etc)											
30	The data origin should be a filtration criterion (F-30)		3.84		1	The user will be able to filter datasets based on the origin/publisher.	Filter data sets by publisher	Semantic Search	Search results + publisher filter cirtierion	UI	Results filtered by publisher	Result Filtration
43	Ranking of search results using different criteria (F-43)		3.84		1	User will have the option the rank the results based on information available about each dataset like publisher/source, time of publishing, last update etc..	Ranking / sorting search results	Semantic Search	Search results + ranking criteria	UI	Ranked results	Result Sorting
14	Enable the searching for scientific information relevant to the policy subjects (F-14)		3.84	P	1	Research papers can be found using dblp (http://www.informatik.uni-trier.de/~Ley/db/) for retrieving relevant papers.	Find research papers using DBLP	Topic Analysis	Topic keywords	UI	relevant research papers	
31	Determination of whether and how multiple data sets can be compared (F-31)	Y (HS)	3.84	*N*	*1*	*This requirement needs clarification to decide whether or how to address it.*						
41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)	Y (GESIS)	3.84	P	1	Searches can be conducted using German data sources and with German language keywords.	German language Searches	Topic Analysis / UI	Topic keywords in German Language	UI	Search Results in German Language	Language Support
42	Construction of searches based on the policy themes in question (F-42)		3.84		1	The terms from the topic will be used to find entities on the Web.	Topic analysis output terms used as search terms	Topic Analysis	Topic keywords	Semantic search	Search Results	Topic Analysis / Semantic Search
20	Enable the user to specify different time ranges for when the data was (or is) relevant in a search (F-20)		3.62		1	If the time of publishing of a dataset will be available in its VoID description, the user will have the option to specify time information for searching the data.	Restrict search using publication date	UI	Date range	Semantic search	Date-restricted search query	Restrict Searches by Metadata
33	Present a summary of a search result to the user (F-33)		3.48	*R*	*1*	*Metadata information about the dataset if available will be used as a summary of the dataset and will be presented to the user.*	*Same as F-27*					*Metadata presentaton*

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Search (Semantic, LOD, other etc)											
34	Present available metadata about a search result to the user (F-34)		3.48		1	Metadata information if available will be provided to the user.	Present metata from search results to user. Superset of F-9 & F-27	Semantic Search	result data	UI	Result metatata	Metadata Presentaton
22	Locate information related to the topic in question, but restricted to specified countries (F-22)	Y (HS)	3.48		1	It needs to be defined whether the requirement refers to the publisher of the data or data itself. In some cases the publisher may be not from within the same country as the data. In principle, however, if this information is available in the metadata of the dataset then this requirement will be addressed accordingly.	Restrict search to use or exlude country of publication (incorporating F-32)	UI	Countries to include (or exclude)	Semantic search	Country-restricted search query	Restrict Searches by Metadata
24	Ability to use data from both the European Commission and the European Parliament (F-24)	Y (Gov2u)	3.48		1	If the data is available in the form of linked open data or as open data and the publisher has provided APIs to access the data, the tool will be able to look into those data repositories for information relevant to policy themes.	Restrict data sources to those provided by the European Commission & Pailiament	UI	Flag: select EC & EP only data sources	Semantic search	Flag: search only EC & EP data sources	Restrict Searches by Metadata
32	Find documents and data sets related to the policy element in question from countries outside the native country of the user (F-32)	Y (HS)	3.48	*R*	*1*	*Links to those other datasets may be added as search target. As an example, Dbpedia provides owl:sameAs links for several languages.*	*restrict search excluding a country (negation of F22?) - incorporated into F-22*	*UI*	*country to exclude (possibly automatic from user profile)*	*Semantic search*	*search excluding data soruces from specified country*	*Restrict Searches by Metadata*
36	Locate discussions on social media that are related to the policy issue in question (F-36)		3.48	New	1	The location of information will be based on the selection of keywords representing the policy. These keywords can be voluntarily provided by the users or automatically extracted from a policy document. (SJT: Not applicable to Social Media Analysis - part of the social media data gathering)	Search restricted to specified countries	UI	Search terms	Social Media Analysis	Social media data	Social Media Data Gathering
44	Monitor blogs and comments on them (F-44)		3.48	new	1	Social media tools will monitor particular social media conversations (posts and comments) by considering the users' demands (SJT: Not applicable to Social Media Analysis - part of the social media data gathering)	Sentiment & opinions from blogs & comments	UI	Search terms	Social Media Analysis	Social media data	Social Media Data Gathering

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Search (Semantic, LOD, other etc)											
25	Augment search terms the user has specified and perform additional and related searches to the ones the user is running (F-25)		3.12		1	The search terms extracted in the form of topics from policy terms will be enhanced using synonyms and by using the entity lookup service	Find related terms using multiple instances of LOD graph search	Topic Analysis	Topic keywords	UI	Related keywords & entities	Find Related entities / keywords
37	Enable the user to specify a date that determines the earliest creation point of data in search results (F-37)		3.12	*R*	*1*	*If time of publishing of the dataset will be available, the user will have the option to set the date beyond which she is willing to see results.*	*Similar to F-20*					*Restrict Searches by Metadata*
38	In the event of multiple versions of a data set being available, highlight the most recent (F-38)		3.12	*R*	*1*	*The most recent versions will be highlighted by leveraging updated data and information available in the metadata.*	*Similar to F-20, ranking by most recent*	*any*	*Result data*	*UI*	*Ranked results*	*Restrict Searches by Metadata*
39	Match an impacted factor with a government department (F-39)		3.12	*N*	*1*	*Further clarification is needed to be able to determine whether this requriement can be addressed and how.*						

Table 5: Functional Specification for Search

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

6.4 Topic Analysis

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Topic Analysis											
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)	Y (HS)	8.95		1		Documentation	N/A				Documentation
41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)	Y (GESIS)	3.84		1		German language support	UI	Policy document		German language topics	Topic Analysis
42	Construction of searches based on the policy themes in question (F-42)		3.84		1	Topic analysis tool will provide the topics and top terms in each topic will be used as search terms for further processing and will be used as query terms.	Topic analysis to provide terms	UI	policy document	Semantic Search	search results	Topic Analysis

Table 6: Functional Specification for Topic Analysis

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

6.5 User Environment / Interface

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	User Environment / Interface											
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	Y (HS)	18.79		1	All provenance information will be presented to the user in the user interface. What forms this will take will depend on what information is available for externally-sourced data, but for data generated by Sense4us (running analyses to generate derived data), the chain of processing will be shown to the user, and the user can see what Sense4us tools were run on what data and with what settings.	Present provenance metadata to the user	Any	Any Result Type	UI	Provenance Metadata	Metadata Presentation
12	Easy to understand documentation and training on what the tools are capable of, how to use the tools, and the types of result to expect (F-12)	Y (HS)	8.95		1		Documentation	N/A				Documentation
6	Enable searching multiple data sources with the same query and mouse click from the user (F-6)	Y (HS)	8.26		1	Users will be able to specify search terms, or use the output of another Sense4us tool as search terms, and then be able to specify the sources of data to search (e.g social media, open data etc). All searches conducted as a result of this group of searches will have their results stored separately in the Sense4us database under the user's profile, but they will be also accessible as a group corresponding to the group of searches.	Specification, execution and display of search results where searches are conducted across multiple locations	UI	URLs of locations to search	UI	Combined search results	Search Specification / Result Presentation
8	Enable the user to filter search results - include or exclude results according to a wide range of criteria (F-8)	Y (HS)	7.26		1	It is likely that filters will be displayed in categories so the user can find them easily, and users should be able to build up a set of filtration terms they use most, and have these associated with their user profile (e.g. "favourite filters").	Filtration of output from user-selected criteria	any	Any Result Type + Filtration Criteria	UI	Filtered Results	Result Filtration

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	User Environment / Interface											
3	Enable the user to add specific data sources and websites as search targets (F-3)		6.63	P	1	A box will be provided to allow URLs of data sources and websites to be added manually by the user, to be included in a group of searches (see response to F-6). Data sources can take many forms, from a URL of a linked open data site, to the twitter account of a social media user. Ideally it should be possible to specify these different types of data source, but the project will need to investigate which types are most effective to support. It is unlikely every type of data source will be supported, and it is unknown whether a data source will be supported from simply a URL - whether a data source can be supported will only become apparent once a search is attempted upon it, so the user should be warned to expect that some sources will not produce any results. There are likely to be classes of data source (e.g. straightforward web page, linked open data, social media), and some of these can be preset into lists from which the user can select. Users may have favourite data source starting points, so it should also be possible to store these in a user's profile.	Specification of search locations	UI	URLs of locations to search	Semantic Search	Search targets	Search Specification
10	Enable the user to restrict search results to those that they trust (F-10)	Y (HS)	5.21	*R*	*1*	*coupled with provenance information (see responses to F-9). It should be possible for a user to specify a "white list" of trusted data sources and have this stored within their profile so that the result display may use this as filtration criteria, or the searches themselves could be run only on the trusted sources. See note in the response to F-3 about different types of data source - this is applicable here also. There is an additional means of achieving the same end here - to specify trusted data sources before a search is run, so the search only runs over the specified data sources. This could be more economical than a general search, as the data sources are limited to a finite set.*	*Filtration by trusted sites (See F-8 and F-9)*	*UI or user profile*	*Filtration Criteria - trusted places*	*UI*	*Filtered results*	*Result Filtration*

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	User Environment / Interface											
2	Add new data sources (F-2)	Y (HS)	5.21	*R*	*1*	*This is addressed in the response to F-3.*	*(Same as F-3) Specification of search locations*	*UI*	*URLs of locations to search*	*Semantic Search*	*Search targets*	*Search Specification*
43	Ranking of search results using different criteria (F-43)		3.84		1	The user will be given different ranking criteria to choose from (assuming it is possible to rank the results), e.g. they can sort by different columns in the result list, or if other crtieria are available that are not result columns, they can choose from a list of criteria to sort by.	Sorting & ranking results (Same response as F-11 in Result Presentation)	UI	Sort Criteria (also could be automatic so no input needed)	UI	Ranked Results	Result Sorting
29	Highlight any data from social media (F-29)		3.84	*R*	*1*	*This is simply a case of showing the data source to the user and pointing out that it is social media.*	*Addressed by result presentation*					
41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)	Y (GESIS)	3.84		1	Here we consider the UI aspects alone. The UI will be developed so that it may be skinned with different languages and a translation file supplied for German. (We will need help from GESIS for this.) This is a special case of F-26.	UI elements in German			UI	German Language version of UI	Language support
21	Provide user accounts, with storage and retrieval of data and workflows all within a user's account (F-21)	Y (HS)	3.48		1	User accounts / profiles will be provided and will contain a number of different customisation functions, some of which have been mentioned in responses to other functional requirements e.g. user's favourite filtration criteria (F-8) and the sources the user trusts (F-10). In addition, all results from searches and analyses are stored in a back-end database, and are stored connected with the user account that executed the search or analysis. Therefore a user can access data from the history of their actions.	User accounts & profiles, user-based data storage etc	UI	Setting preferences	multiple		User accounts & profiles
20	Enable the user to specify different time ranges for when the data was (or is) relevant in a search (F-20)		3.62		1	Restricting searches to a specified date (NEW FEATURE added because F-20 can be addressed by restricting the search criteria, or by filtering results - this version is restricting search criteria)	Restrict search using publication date	UI	Search terms + Date range	Semantic search	date-restricted search query	Restrict Searches by Metadata

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	User Environment / Interface											
26	Support multiple languages, in the user interface labelling, searching and analyses (F-26)		3.12	P	1	This is a more general case of F-41. The UI will be developed so that it may be skinned with different languages and translation files can be supplied for different languages.	Multiple language support			UI	UI elements in multiple languages	Language support
37	Enable the user to specify a date that determines the earliest creation point of data in search results (F-37)		3.12		1	This is only possible when dates are provided by data sources. The earliest date can be a filtration criteria, e.g. "creation date > 2014-09-18" meaning created after this date.	Filtering data by date	UI	Date range critieria	UI	date-filtered results	Result Filtration
40	Ability for the policy maker user to customise the tool to their preferences (F-40)	Y (HS)	3.12		1	User profiles with customisation functions will be provided. There are many other functional requirement responses that contain customisation and user profile aspects (e.g. favourite filters and trusted source lists), and these will be gathered together to provide the specification for the user profile section of the Sense4us toolkit.	Customisation within user profiles	UI	Setting preferences			User accounts & profiles

Table 7: Functional Specification for User Environment / Interface

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

6.6 Social Media Analysis

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Social Media Analysis											
9	Present to the user all available information about the provenance of some data that has been found by the tool (F-9)	Y (HS)	18.79		1	We will provide information to the user about: (1) which platform the data comes from (Twitter/Facebook, etc.), (2) which social media user accounts the information of interest comes from and (3) when available, we will provide information of whether the social media user account is verified (https://support.twitter.com/articles/119135-faqs-about-verified-accounts)	Show provenance data for social media data	Social Media Analysis	Social Media Analysis Data	UI	Social media provenance data	Provenance
5	Enable opinions from social media and other forums to be gathered (F-5)	Y (HS)	6.00		1	The social media analysis tools developed in Sense4us will compute sentiment and opinions from user generated content and provide these insights to the policy makers	Sentiment & opinions from social media	Search	Social media data (e.g. tweets or posts)	UI	Sentiment & opinions about the policy	Social Media Analysis
7	Enable finding information for subjects related to those in the policy (F-7)		6.00	*N/A*	*1*	*(SJT: Not applicable - this is more to do with searches than analysis of what people are saying and sentiment analysis)*						
19	Discover any publicly available characteristic information about the social media participants who are expressing the opinions (F-19)	Y (HS)	4.68		1	To address this requirement the tool will provide all user data publicly available in the social media platforms (e.g., name of the user, location, description, time in the platform, etc.) as well as additional insights of those users based on their activities in the platform (post rate, size of their social network, etc.)	Social media metadata about participants	Search	Social media data (e.g. tweets or posts)	UI	Social media metadata	
18	Acquire information from social media on the topics and themes of the policy in question (F-18)		4.15		1	In addition to polarity information, the social media analysis tools will extract arguments in favour and against a particular policy based on the monitored conversations in social media	Arguments about the policy subjects	Search	Social media data (e.g. tweets or posts)	UI	Arguments for and against the policy	Policy arguments

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Social Media Analysis											
15	Enable the collection of sentiments / opinions about the policy subjects (F-15)	Y (HS)	3.84		1	Every policy subject will be represented by a subset of social media posts. Sentiments and opinions about a policy subject will be computed by extracting relevant polarity information from the previous filtered subset of policy-relevant data. This filtering process will be performed by allowing the user to select some keywords related to the policy, or by automatically extracting these keywords from a policy document	Sentiment & opinions from social media	Search	Social media data (e.g. tweets or posts)	UI	Sentiment & opinions about the policy	Sentiment analysis
29	Highlight any data from social media (F-29)		3.84		1	To visualise large datasets of data several intuitive metrics will be provided such as, size of the dataset, time period when the dataset is produced, origin, etc.	Highlight social media data	Search	Social media data (e.g. tweets or posts)	UI	Any data from social media highlighted plus metadata about the dataset	Provenance & Metadata
41	The tool's UI and the analyses in the research partners' components support German as a language (F-41)	Y (GES	3.84		1	English Natural Language Processing (NLP) libraries will be replaced by German NLP libraries to enable analysing data in both languages. It is important to consider that English NLP tools are more researched and refined so it is possible to loose certain level of accuracy in the analysis when replacing English NLP libraries by the equivalent German ones	German language support in sentiment & opinion analysis	Search	Social media data (e.g. tweets or posts)	UI	Sentiment analysis in German language	German Language Support
20	Enable the user to specify different time ranges for when the data was (or is) relevant in a search (F-20)		3.62		1	The user will be able to provide the time period and the granularity for the analysis (initial date / final date and time granularity, minute, hour, week, etc.) The tool will recommend a specific time-granularity depending on the volume of the available data	Time-based sentiment and opinion analysis	Search + UI	Social media data + Time range & granularity	UI	Sentiment analysis given data over this time	Sentiment analysis
22	Locate information related to the topic in question, but restricted to specified countries (F-22)	Y (HS)	3.48		1	When location information is available for users and posts (like in Twitter), the analyses will be able to restrict data to specific countries	Location-based sentiment & opinion analysis	Search + UI	Social media data + specified country	UI	Sentiment analysis on country-restricted data set	Sentiment analysis

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

F-ID	Functional Requirement	Critical ?	Priority (Bayesian Avg)	Address?	In Scope?	How to address	Summary	Input Compo-nent	Input Data	Output Compo-nent	Output Data	Classification
	Social Media Analysis											
36	Locate discussions on social media that are related to the policy issue in question (F-36)		3.48	*N/A*	*1*	*The location of information will be based on the selection of keywords representing the policy. These keywords can be voluntarily provided by the users or automatically extracted from a policy document. (SJT: Not applicable to Social Media Analysis - part of the social media data gathering)*	*Search restricted to specified countries*	*UI*	*Search terms*	*Social Media Analysis*	*Social media data*	*Social Media Data Gathering*
44	Monitor blogs and comments on them (F-44)		3.48	*N/A*	*1*	*Social media tools will monitor particular social media conversations (posts and comments) by considering the users' demands (SJT: Not applicable to Social Media Analysis - part of the social media data gathering)*	*Sentiment & opinions from blogs & comments*	*UI*	*Search terms*	*Social Media Analysis*	*Social media data*	*Social Media Data Gathering*

Table 8: Functional Specification for Social Media Analysis

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project

7 Conclusions

This deliverable has described the process and results in the work to derive the functional specification for the coarse grained components in Sense4us.

The major result of this deliverable is a set of tables (in Section 6) that describe the functional requirements that have been derived from the end user requirements, and the functional specification of components, which is how the project will address the functional requirements. This set of tables will be used as input to the architecture task T3.2, and the first result of this work will be in the deliverable D3.2, due PM18. Also, because they contain specifications for each of the components in the research and integration work in the project, the tables in Section 6 will be used as a benchmark by which to evaluate the research and integrating partners' work – we can use the tables as a checklist of functionality that different components need, and evaluate the components as to whether they provide that functionality.

Another exploitable result of the work described in this deliverable is also the design and testing of the functional specification process itself. This process has provided a formal procedure by which to transform the end user requirements into the specification of components in the project. It is believed that the process can be useful and applicable to a wide range of other projects.

The major requirement for the functional specification process was that there should be a clear and obvious path from the end user requirements through to the specification of the components that are meant to address these requirements. The process has addressed this through:

- involvement of the other project partners at all steps to validate intermediate results (e.g. the functional requirements and their priorities), and
- provision of records (shown in this deliverable) that relate the end user requirements to the functional requirements, and the functional requirements to the component specification needed to address them.

From the start, the process was designed to be open and inclusive for the project and all project partners were consulted during the process in one way or another. We believe the result is more justifiable (and fair) than if one partner had conducted a functional requirements analysis unilaterally and presented it as a *fait accompli* to the other partners.

The Sense4us project has a number of iterative requirements-specification-implementation cycles and the functional specification process will be also iterated later in the project. This will enable the incorporation of new end user requirements that arise from initial demonstrations and evaluation, as well as new discoveries being made by the research conducted by the academic partners in the project. As a result of this process, the functional specification will be updated – new functionality added, and it is possible other functionality may be removed, if it proves not useful.

© Copyright 2014 IT Innovation Centre and other partners of the Sense4Us project