REVEAL

FP7-610928

REVEALing hidden concepts in Social Media

Deliverable D1.1
Requirements analysis and specifications

<table>
<thead>
<tr>
<th>Editor(s):</th>
<th>Linda Rath-Wiggins, Ruben Bouwmeester, Jochen Spangenberg, Gabriele Strobel, Marc Dorchain, Eva Jaho, Nikos Sarris</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsible Partner:</td>
<td>ATC</td>
</tr>
<tr>
<td>Status-Version:</td>
<td>Final - v0.8</td>
</tr>
<tr>
<td>Date:</td>
<td>05/05/2014</td>
</tr>
<tr>
<td>EC Distribution:</td>
<td>Public</td>
</tr>
<tr>
<td>Project Number:</td>
<td>FP7-610928</td>
</tr>
<tr>
<td>Project Title:</td>
<td>REVEAL</td>
</tr>
</tbody>
</table>
Title of Deliverable: Requirements analysis and specifications

Date of Delivery to the EC: 05/05/2014

Workpackage responsible for the Deliverable: WP1 – User Requirements and Regulatory Framework

Editor(s): Linda Rath-Wiggins, Ruben Bouwmeester, Jochen Spangenberg, Gabriele Strobel, Marc Dorchain, Eva Jaho, Nikos Sarris

Contributor(s): ALBLF, UKob, CERTH, ITINNO, NCSR'D', SINTEF, KU Leuven

Reviewer(s): SINTEF

Approved by: All Partners

Abstract: This document provides information about the process as well as outcomes of the user requirements analysis that was carried out in the journalism and enterprise sector. It describes the process by which the consortium collected users’ views, needs and obstacles regarding the verification process in Social Media. All this led to formulating the end user requirements of the journalism and enterprise use case respectively. The deliverable is intended to outline user requirements in detail, thereby guiding the development work in subsequent work packages.

Keyword List: Social Media, Journalism, Enterprise, User Requirements
## Document Revision History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Modifications Introduced</th>
<th>Modification Reason</th>
<th>Modified by</th>
</tr>
</thead>
<tbody>
<tr>
<td>v0.1</td>
<td>26/03/2014</td>
<td>Creation of the document structure</td>
<td></td>
<td>ATC</td>
</tr>
<tr>
<td>v0.2</td>
<td>04/04/2014</td>
<td>Inclusion of methodology of user requirements analysis section</td>
<td>DW, SAG, SINTEF</td>
<td></td>
</tr>
<tr>
<td>v0.3</td>
<td>11/04/2014</td>
<td>Inclusion of SAG requirements</td>
<td></td>
<td>SAG</td>
</tr>
<tr>
<td>v0.4</td>
<td>18/04/2014</td>
<td>Inclusion of DW requirements</td>
<td></td>
<td>DW</td>
</tr>
<tr>
<td>v0.5</td>
<td>25/04/2014</td>
<td>Inclusion of technical partners comments/feedback</td>
<td></td>
<td>ATC</td>
</tr>
<tr>
<td>v0.6</td>
<td>29/04/2014</td>
<td>Corrections and editing after comments from the consortium</td>
<td>ATC, DW</td>
<td></td>
</tr>
<tr>
<td>v0.7</td>
<td>30/04/2014</td>
<td>Provision of internal review comments</td>
<td></td>
<td>SINTEF</td>
</tr>
<tr>
<td>v0.8</td>
<td>05/05/2014</td>
<td>Address of received comments and final formatting</td>
<td></td>
<td>ATC</td>
</tr>
</tbody>
</table>
CONTENTS

EXECUTIVE SUMMARY ........................................................................................................... 9

1 INTRODUCTION .................................................................................................................. 10
  1.1 THE REVEAL VISION ........................................................................................................ 10
  1.2 PURPOSE OF THIS DOCUMENT .................................................................................... 12
  1.3 DOCUMENT STRUCTURE ................................................................................................ 12

2 METHODOLOGY FOR USER REQUIREMENTS ELICITATION ........................................... 13
  2.1 METHODS FOR CAPTURING USER REQUIREMENTS ...................................................... 13
    2.1.1 News field .................................................................................................................. 14
      2.1.1.1 Literature study ..................................................................................................... 14
      2.1.1.2 Interviews ............................................................................................................. 18
      2.1.1.3 Market analysis ..................................................................................................... 29
    2.1.2 Enterprise field .......................................................................................................... 33
      2.1.2.1 Observation and brainstorming ............................................................................. 34
      2.1.2.2 Interviews ............................................................................................................. 34
      2.1.2.3 Process of data collection ..................................................................................... 37

3 REVEAL REQUIREMENTS .................................................................................................... 40
  3.1 NEWS REQUIREMENTS .................................................................................................... 40
    3.1.1 Mini scenario 1: Newsgathering support .................................................................... 40
    3.1.2 Mini scenario 2: Revealing the Contributor ................................................................. 40
    3.1.3 Mini-scenario 3: Revealing Multimedia Content ......................................................... 41
    3.1.4 Mini-Scenario 4: Revealing Location ........................................................................... 41
  3.2 ENTERPRISE REQUIREMENTS ....................................................................................... 42
    3.2.1 Preliminary Note: Software AG Specifics ................................................................... 42
    3.2.2 Mini-Scenarios ........................................................................................................... 42
  3.3 DW AND SAG MERGED TABLE OF REVEAL REQUIREMENTS .................................. 46
    3.3.1 Requirements for revealing Contributor ...................................................................... 46
    3.3.2 Requirements for revealing Multimedia Content ......................................................... 53
    3.3.3 Requirements for revealing Context ........................................................................... 56
    3.3.4 General requirements ................................................................................................ 58

4 THE REVEAL APPLICATIONS ............................................................................................. 61
  4.1 NEWS APPLICATION ........................................................................................................ 61
  4.2 ENTERPRISE APPLICATION ........................................................................................... 64

Project Title: REVEAL
Project Coordinator: INTRASOFT International S.A.
Contract No. FP7-610928
www.revealproject.eu
5 REVEAL MODULES............................................................................................................................ 66
6 CONCLUSIONS...................................................................................................................................... 69
7 REFERENCES........................................................................................................................................ 70
8 APPENDIX A: COMPETITIVE TOOLS AND OPPORTUNITIES TO FILL THE MARKET GAPS.......................................................... 72
9 APPENDIX B: QUESTIONNAIRE FOR THE NEWS USE CASE SCENARIO ..... 75
10 APPENDIX C: QUESTIONNAIRE FOR THE ENTERPRISE USE CASE SCENARIO .......................................................... 80
11 APPENDIX D: MATRIX OF FEATURE GROUPS OF ENTERPRISE USE CASE 85
## LIST OF FIGURES

| Figure 1: Some of the concepts we aim to model and reveal | 11 |
| Figure 2: Question: What is your position?, N=20 (multiple answers were possible) | 20 |
| Figure 3: Question: What do you usually work on?, N=20 (multiple answers were possible) | 21 |
| Figure 4: Question: How do you use Social Media platforms professionally?, N=20, multiple answers were possible | 23 |
| Figure 5: Question: When you use Social Media content as a way to collaborate, what do you do?, N=20, multiple answers were possible | 24 |
| Figure 6: Question: Once you find information on Social Media platforms, what do you do?, N=20, multiple answers were possible | 25 |
| Figure 7: Question: How does your verification process look like right now?, N=20, multiple answers were possible | 26 |
| Figure 8: Screenshot Storyful Landing Page, December 2013 | 29 |
| Figure 9: Screenshots of detailed information of one YouTube video | 31 |
| Figure 10: Storyful uses three elements of verification | 32 |
| Figure 11: Identification tree "contributor" | 38 |
| Figure 12: Identification tree "topics" | 40 |
| Figure 13: Process of Requirements Gathering | 43 |
LIST OF TABLES

TABLE 1: USER REQUIREMENTS ANALYSIS TOOLS EMPLOYED .................................................13
TABLE 2: USE OF VARIOUS PLATFORMS, N=20 .................................................................21
TABLE 3: MOTIVATION BEHIND USAGE OF SOCIAL MEDIA PLATFORMS ..........................22
TABLE 4: VERIFICATION PROCESS ON VARIOUS SOCIAL MEDIA ACCOUNTS ..............26
TABLE 5: QUESTION: WHAT TOOLS DO YOU USE IN ORDER TO VERIFY CONTENT ON THE SOCIAL WEB? N=20
................................................................................................................................................27
TABLE 6: QUESTION: WHAT ARE SOME OF THE MAJOR OBSTACLES YOU ENCOUNTER WHILE VERIFYING
content? N=20 ............................................................................................................................28
TABLE 7: SOCIAL MEDIA CHANNELS SOFTWARE AG IS INVOLVED IN .........................33
TABLE 8: ANONYMIZED PARTICIPANTS DATA .........................................................................36
TABLE 9: REVEAL MODULES ....................................................................................................66
TABLE 10: COMPETITORS’ APPLICATIONS ............................................................................72
### DEFINITIONS, ACRONYMS AND ABBREVIATIONS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP</td>
<td>Work Package</td>
</tr>
<tr>
<td>UI</td>
<td>User Interface</td>
</tr>
<tr>
<td>HTTP</td>
<td>Hyper Text Transfer Protocol</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>URI</td>
<td>Uniform Resource Identifier</td>
</tr>
<tr>
<td>RTs</td>
<td>Retweets</td>
</tr>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>CRM</td>
<td>Customer Relation Manager</td>
</tr>
</tbody>
</table>
Executive Summary

This document details the methods used and the result of the process for gathering and analysing the user requirements for the REVEAL scenarios. Moreover, it lists the modules/functionalities identified in this phase by the technical partners that will support the implementation of the derived user requirements.

The analysis of the requirements provides the basis for the technical work in REVEAL, and will play a crucial role during the design, development and evaluation of the REVEAL applications.

A number of Appendices are presented at the end of the document, providing detailed information on different aspects of the requirement analysis process.
1 Introduction

1.1 The REVEAL vision

With regards to the journalistic use case, it can be stated that traditional journalism practices are changing due to disruptive developments posed by Social Media platforms, such as YouTube, Facebook, Twitter, Flickr or Foursquare. Because of the collaborative sharing, exchange and distribution of multimedia content in Social Media, users are increasingly involved in:

- Raising awareness to topics that have not (yet) been on the agenda of “traditional” media organisations,
- Contributing to topics with eyewitness accounts,
- Voicing and formation of opinion concerning any topics discussed,
- Re-contextualising narratives of stories due to the sheer variety of perspectives,
- Creating their own small public spheres, e.g. on YouTube, Facebook or Twitter.

For all these reasons, journalists turn to Social Media to see what is shared, posted and distributed. Furthermore, journalists now have the chance to research topics, curate information and analyse stories published by “the people formerly known as the audience” (Rosen, 2006) that come up in discussions on Social Media platforms. Social Media may also be used for the identification of sources or for finding eyewitnesses in order to interview them as part of the reporting and story production process (Wardle, 2014). Journalists use Social Media to share their experiences, their thoughts and engage in dialogue with their readers (Spangenberg & Heise, 2014).

While working with Social Media content provided by users, journalists must keep in mind the fundamental journalistic norms, such as impartiality, objectivity and accuracy (Golding & Elliott, 1979). Credibility and trustworthiness of journalists or media organisations can only be ensured when these norms are kept at all times. “Trust is something that is hard to earn and easy to lose. And it is a core element of journalism; few other professions are so dependent on trust” (Lorenz et al, 2011). The relevance of the trust factor implies a crucial need for effective verification methods of Social Media sources and content (Schifferes & Newman, 2013). „If telling the truth is our goal, verification must be our standard“ (Buttry, 2014, S. 15). The verification process is not new in the journalistic context: “The medium by which we’re gathering information may change, but the principles of verification always apply” (De Rosa, 2014, S. 18). However, the verification of Social Media sources and content is challenging as Social Media and UGC (user generated content) require journalists to verify as the story unfolds (Buttry, 2014). Multimedia content needs to be challenged „on-the-go“ – especially in emergency situations, journalists “are increasingly presented with an abundance of official sources and can find first-hand sources, the people who actually saw – or even participated – in the events in question. But those accounts can be flawed“ (Buttry, 2014, S. 16). There have been cases in which content has been deliberately manipulated (Brandtzæg et al, 2014). Social Media has changed the “art of verification” profoundly, offering new opportunities, but also presenting those engaged in it with new challenges.

The relevance of verifying content in an easy, transparent and relatively fast way is still relevant, especially when taking into consideration (a) the sheer quantity of content found in social networks, and the (b) fact that a lot of the content can be hoaxes, rumours, or misleading propaganda.

Journalism is a very good example for illustrating the need and usefulness of the described practices, but it is not the only one. The daily exploration of content found in open all closed enterprise communities can benefit tremendously by the analysis proposed. The nature of this content renders itself ideal for extracting many of the REVEAL modalities. Enterprise community
members’ replies, ratings and comments can be utilised to measure modalities like contributor reputation, history, popularity or influence. For example, a community member may provide a very good solution, which can be already evaluated by other community members for its stability and robustness, and can be embedded to a product with little further testing. Cross-linked sources (e.g. Twitter) can also be used to combine the mechanics of other Social Media platforms and enhance the contextual information available. Furthermore, gamification mechanisms can be used to enrich the available information by motivating the members to contribute to the data gathering process. The actors in this case are the online community members, who ask questions, discover problems, software bugs, and provide solutions. These should be aggregated by the REVEAL platform, organised, summarised and evaluated to discover concepts like trustworthiness, completeness, quality, or controversiality. These should be presented to the enterprise in a way they can use it to improve the company’s processes or products.

Our aim therefore is to go much further than content verification. We need to be able to estimate concepts like contributor impact and what extent all this affects reputation or influence. This should allow us to automatically judge more complex concepts like the quality and accuracy of content, and bring us to predicting future trends with greater accuracy.

In the case of journalism, multimedia content provided by users via Social Media is playing an ever-increasing role in the detection, production and distribution of news. Non-professional journalists can be more actively involved in the production and distribution of news, in particular because of the availability of:

- Affordable portable devices (especially smartphones) that allow for the capturing of eyewitness accounts in an audio-visual format;
- Internet access almost anytime and anywhere (mobile and stationary), more and more of it provided at high speeds;
- Social Media platforms with networking capabilities (especially Social Networks) that allow for the sharing of content and the fast spreading of information to potentially millions of people.

News organisations can now take this user-generated multimedia content and use that for researching purposes, for embedding it into their own body work (e.g. as part of their online articles etc.) or for finding experts and eyewitnesses. In that sense, their former “gatekeeper” functions have been challenged profoundly.

In the context of enterprises, it is challenging to discover what people are saying about a product or service. Knowing how influential these people are and how much this can affect the reputation and marketability of the product or service is challenging. We are entering a stage that allows us to automatically judge the quality of content, and even predict future trends.

Figure 1: Some of the concepts we aim to model and reveal
In both contexts, REVEAL aims to *discover higher level concepts hidden within information* (graphically illustrated in Figure 1). In Social Media we do not only have bare content; we also have to deal with interconnected sources. Furthermore, we have to handle and interpret interactions between sources. And there is much more information available that indicates the context within which the content is being used and the context in which these interactions are taking place.

### 1.2 Purpose of this document

This deliverable aims to define the user requirements that will form the basis for the REVEAL architecture (modules and inter-connections) design that is considered necessary to fulfil the identified requirements. Thus, this document facilitates the communication between partners from research, development, and from use case analysis activities. It describes concisely the process the consortium undertook for the collection and analysis of user requirements and the envisaged application scenarios for the news and enterprise use cases respectively.

This document will provide subsequent Work Packages with user requirements and design objectives with regards to the news and enterprise domain. In this context, it is of major importance, due to the fact that subsequent Work Packages rely on the establishment of clearly formulated user requirements.

### 1.3 Document structure

In order to achieve the aforementioned objective, we followed an approach consisting of several steps. In that respect, the deliverable is structured as follows:

Section 2 presents the methodology and material that has been used in order to derive the user requirements. It also elaborates on the methods followed by WP1 partners engaged in the user requirements elicitation process. Different methods were chosen, based on the needs of the two different fields that REVEAL will focus: news and enterprise.

The analysis of the findings of the methods for deriving the user requirements, result in a number of mini scenarios that are presented in Section 3. These mini scenarios are further analysed in detailed requirements or user stories.

Section 4 presents the envisioned application scenario for the news use case with sketches that can help future work on visualising the relevant requirements derived. For the enterprise use case, there is some supporting visualisation based on the methodology followed for the derivation of the requirements, but some user interface determinants are presented. The results from both use cases are presented and make sure that the gathered requirements and knowledge is transferred to the developer partners in an understandable and meaningful way.

Section 5 provides a list of the REVEAL modules/functionalities as discussed with the technical partners for satisfying the user requirements identified. We do not consider the list as final at this stage, as we intend to update and enrich this during the project lifetime.

Finally, Section 6 provides some concluding remarks.
2 Methodology for User Requirements Elicitation

2.1 Methods for capturing user requirements

Methods such as observation, interview, document analysis, focus groups and analysis, checklists or questionnaires can be used for the elicitation of user requirements. Scenarios and use cases have also become a popular technique for task analysis that can be used to assist users in developing a clear view of the forthcoming applications, and thus provide more accurate feedback. Different requirements analysis methods can be applied in parallel to complement each other in order to yield more effective results.

For carrying out the process of requirements identification and analysis for REVEAL a variety of strategies and approaches were used in a complementary way. They are listed together with their benefits and drawbacks in the following table. A mix of all below mentioned methods was employed in the REVEAL user requirements survey in order to collect data from various target groups with different backgrounds and from different localities. Implementation details and results are described in the next sections.

Table 1: User requirements analysis tools employed

<table>
<thead>
<tr>
<th>METHOD/TECHNIQUE</th>
<th>DESCRIPTION</th>
<th>BENEFIT</th>
<th>DRAWBACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example Scenario / Usage Scenario</td>
<td>Detailed realistic examples of how users may carry out their tasks in a specified context with the targeted system</td>
<td>They can bring user needs to life and explain their vision about the future application</td>
<td>They may raise expectations too much and over-simplify the application</td>
</tr>
<tr>
<td>User Surveys/Questionnaires</td>
<td>A set of questions for a sample of users. Surveys can help determine needs, current work practices and attitudes towards the new system ideas</td>
<td>Relatively quick method of determining preferences of large user groups. It also allows statistical analysis</td>
<td>This method may not capture in depth comments and may not permit follow-up with the respondents to clarify vague responses</td>
</tr>
<tr>
<td>Interviewing</td>
<td>A series of fixed questions with scope for the end user to expand on his response</td>
<td>Interviews allow quick elicitation of ideas &amp; concepts, and follow-ups to clarify subject opinions</td>
<td>Time-consuming. More difficult coding and analysis of results</td>
</tr>
<tr>
<td>Review of Existing Systems / Services (i.e. direct or indirect review)</td>
<td>Comparison of expected product (or components thereof) with existing systems and services</td>
<td>Effective in identifying current problems and promoting new features</td>
<td>This method may lead to including too many new functions or make the system too similar to a competitor's</td>
</tr>
</tbody>
</table>
2.1.1 News field
In order to gather user requirements, Deutsche Welle, in co-operation with Sintef, decided to divide the process of the gathering of requirements into three different parts. Firstly, a literature study to outline known facts of: (1) already existing fact-checking practices in newsrooms, (2) helpful verification features that have been proven and tested and (3) experts who have done research on verification practices over a period of time. Secondly, a market analysis will assist in being able to depict successful platforms and/or tools that have proven to be helpful for journalists. A detailed list of functionalities can assist in understanding the variety of needs addressed in the context of the tools. Finally, expert interviews with journalists, editors and people working in media organisations make it possible for the requirements gathering to understand current practices, obstacles and wishes concerning the verification of Social Media content. Findings of these three parts are consolidated at the end of the paragraph and form the foundation of the requirements depicted in the requirements section.

2.1.1.1 Literature study
In this paragraph, the most relevant findings are described from sources such as blogs, books and use cases concerning verification in Social Media. Verification as part of the journalistic workflow is increasingly covered and discussed in blogs and other publications. As Buzzfeed declared 2014 as the “Year of the Viral Debunk”, Warzel (2014) states that “(t)ruth telling and debunking are fundamental journalistic acts, online or otherwise”. Especially with regards to Social Media platforms, more and more people are using networks in order to share, gather or aggregate information. A lot of communication has shifted to Social Media and journalists use these platforms, among others, as radar to understand what is going on and what users talk about or are interested in. In order to include Social Media content in a news story requires journalists to make sure that facts are stated and that the content published is true. Verification and accuracy were recently identified as key ethical challenges by the Online News Association (ONA). The ONA social newsgathering working group raised questions that need to be considered when discussing verification approaches by media organisations (Online News Association, 2014): „Will you trust verification done by other news organizations or journalists? How do you determine that a social account is reliable? (…) Are there tools or technology that can help? How do you even define “verification,” and what’s your threshold for deciding that something is accurate? (…) How much of the process do you want to share with your readers and viewers?” These questions undermine the complexity of verification methods in a journalistic workflow. Mark Little (2014), in turn, CEO of Storyful, published the “ten principles that power social journalism”, in which he declares: “Ladies and Gentlemen, we have unanimous agreement on the need to verify Social Media content. Let’s move on. Time to build and scale the collaborative tools needed to do the job” (Little, 2014).

In early 2014, the Verification Handbook was published, the first of its kind as far as the REVEAL consortium is aware. It provides guidelines for verifying UGC (user-generated content). Edited by Craig Silverman (2014), the book offers examples of best practices on how to verify content provided by the crowd / non-professionals.

The two examples depicted below show the verification process of Social Media content and the different approaches by journalists.

The first example is a tweet with an image posted on Twitter by the respected CNN journalist Hala Gorani1. Gorani shows the 4 year-old boy named Marwan who is crossing the Syrian Desert alone after being separated from his family while fleeing from Syria. Allegedly, the boy was found by UN

---

1 https://storify.com/OpenNewsroom/marwan
staff. The story raises a lot of questions, such as: How did this kid survive? How did the UN staff find the boy? Many of Hala Gorani’s followers commented on the story and re-tweeted it. Soon it became global news.

Even though Hala Gorani is a trusted source from a trusted broadcaster, a member of the Open Newsroom community decided to try and verify the story. Also, Shiv Malik from the Guardian started researching to find out if the Guardian could use the story for publication.

By contacting the UNHCR crew, Malik found out that Marwan’s family was only twenty steps ahead of him.

The situation seemed not to be as distressing as it was presented by Hala Gorani. The conversation on Twitter between Hala Gorani and Shiv Malik shows the various verification processes they both undertook.
The story is not fake in the true sense of the word, but rather leaving out important parts to get a full picture. It seems the truth has been “bent” in order to enlarge the impact of the reporting on the situation in Syria. The verification approach by Shiv Malik has made this clear.

The second example is this image sent to the Guardian’s open journalism platform “GuardianWitness” (Law & Bannock, 2014). It was part of UGC (User Generated Content) about the storms that raged over the U.K. in October 2013.

The image depicts a giant beach ball, twice the size of a double decker bus, on the loose in the centre of London. At The Guardian, the journalists were not sure whether they could publish it because they suspected it could be a hoax. Therefore, they needed to verify the authenticity of the image. Via Google reverse image search and the image-checking tool TinEye, the Guardian team checked whether the image had been published elsewhere online. Also, they tried to contact the person who had sent them the image by the email address that was connected to the respective GuardianWitness account. Moreover, the Guardian staff checked the location via Google Street View (e.g. they tried to answer questions, such as “Do the buildings match? “Is such a perspective/angle possible?” etc.). They searched online for additional mentions of a giant beach ball around that location and they found it had previously been tethered to a building nearby. Via research on Twitter, the journalists were able to confirm other sightings around the same time as claimed in this image.
They even found a later image of a different user that showed that the ball had deflated. After all the research they finally got into contact with the person who had sent them the image and were able to confirm that he had taken the photo himself.

In other words, the image turned out to be real and the photo was published on The Guardian’s live-blog and was shared intensively in Social Media.

Among others, these examples show that witnesses and experts on the ground can provide valuable information in places where journalists have no access or have not been able to arrive yet. However, verification of Social Media content is still of utmost importance when dealing with user generated content. These examples show different ways of how to make sure the content is authentic and real. At present, there are multiple online tools (see Appendix A for a detailed list), such as Storyful or TinEye that journalists may use for Social Media verification. However, verification processes are usually a combination of manually conducted iterations (such as actually calling a source, talking on the phone asking questions etc.) and computer-assisted executions. Wardle (2014) points out that „(v)erification is a key skill, made possible through free online tools and old-fashioned journalism techniques. No technology can automatically verify a piece of UGC with 100 percent certainty” (Wardle, 2014, S. 26). In other words, verifying content can be done in a combination of (1) using tools as well as (2) applying already existing journalistic approaches. However, “many are increasingly looking to ‘advanced computing’ to accelerate and possibly automate the process of verification. As with any other technique, using advanced computing to verify Social Media content in near real time has promises and pitfalls” (Meier, 2014, S. 77). Applying advanced computing to verify user-generated content is still relatively new (Meier, 2014, S. 77), so it remains to be seen what this field of research will show as we go along in this project.

The Verification Handbook also discusses platforms and tools that are helpful for the journalistic verification process. For example, Tweetdeck is highlighted as a platform that helps to understand what the audience is talking about and due to the creation of lists, it can provide an overview of the discussions posted by a trusted network. „If Social Media is a police scanner, TweetDeck is your radio” (De Rosa, 2014, S. 19). Also in the book, Wardle (2014) highlights the following elements as key factors that need to be checked:

- **Provenance**
  - Confirming authenticity (e.g. checking on Google’s reverse image search tool or TinEye, whether it is an original picture or not; cross-referencing with exif data)

- **Source**
o Researching history of the uploader (e.g. on Twitter: linked websites, previous pictures/videos, previous status updates, who are friends/followers?)

- Date
  o Check other sources such as newspapers or other date / time stamps (peculiarities need to be noted, e.g. the fact that YouTube shows timestamps in Pacific Standard Time), Weather information

- Location
  o Check on Google Street Maps (e.g. for buildings)

Many news organisations have published guidelines that outline how content from Social Media should be handled. In some news organisations, specialised journalists have particular roles, such as Social Media editor. In other organisations, groups of people deal with Social Media organised around Social Media desks, as is the case with the BBC’s User Generated Content and Social Media Hub (Turner, 2012; Schifferes & Newman, 2013). These specialised journalists are responsible for verifying content found in Social Media.

Debunking hoaxes and misinformation are increasingly featured as columns, e.g. on the Gawker blog\(^2\). The articles primarily focus on “calling out fake images”\(^3\). Also, the Washington Post has a category devoted to “What was fake on the internet this week”\(^4\). Another interesting find is the work of Matt Novak\(^5\) and his blog Paleofuture, which is part of Gizmodo\(^6\). Novak’s aim is to debunk hoaxes and misinformation.

Also, more and more media organisations make their verification approaches transparent by showing exactly how journalists work in specific scenarios, e.g. in the case of real-time debunking efforts at The Atlantic during Hurricane Sandy\(^7\). The article shows very hands-on what journalists look for when verifying content on Social Media platforms. Also, the New Yorker still employs full-time staff checkers to verify every assertion in each piece before it is published\(^8\).

2.1.1.2 Interviews

In order to gather requirements, and as outlined previously, interviews were conducted by Deutsche Welle and by Sintef.

- Deutsche Welle conducted 20 structured interviews with journalists between January and April 2014. They lasted between 50 to 70 minutes and included interviews with journalists working in various European countries: Germany (14), France (1), Denmark (1), and the UK (1). Journalists had varying nationalities (especially due to the fact that Deutsche Welle employs people from all over the world to provide its services in 30 different languages).

---

\(^2\) http://gawker.com/antiviral-heres-whats-bullshit-on-the-internet-this-1553596302


\(^4\) http://www.washingtonpost.com/newssearch/search.html?st=%22what+was+fake+on+the+internet%22

\(^5\) https://twitter.com/paleofuture

\(^6\) http://www.paleofuture.com/

\(^7\) http://www.theatlantic.com/technology/archive/2012/10/sorting-the-real-sandy-photos-from-the-fakes/264243/

\(^8\) http://www.cjr.org/critical_eye/fact-checking_at_the_new_yorker.php?page=all
• Sintef conducted 7 semi-structured interviews with Norwegian journalists, which lasted 35 to 50 minutes. The interviews were audio-recorded and transcribed.

Regarding the interviews conducted by Deutsche Welle

• Journalists were recruited through the individual networks of the involved DW research team. Interviews took place either face-to-face or via telephone (for the complete questionnaire that was used see Appendix B);
• Objectives included:
  o Understanding the current verification process of journalists who have first-hand experience,
  o Acquiring knowledge of the most crucial needs highlighted by interviewees,
  o Gaining insights of functionalities wanted and needed.
• Because of the combination of a qualitative and a quantitative questionnaire, it was possible to
  o Let interviewees answer in their own ways and receive new insights the interviewers might not have thought of originally
  o Obtain comparable data by asking repeated questions
• Interviews were divided into three different parts:
  o Online usage,
  o Current use of verification tools including challenges and obstacles
  o Socio-demographic information (for the complete questionnaire, see Appendix B).

Regarding the interviews conducted by Sintef

• Seven journalists from Norway were recruited from four major news organizations in Oslo;
• Objectives included:
  o To explore how journalists currently include Social Media in their working practices, with a special focus on verification.
  o This included the following guiding questions:
    • How is Social Media used for professional purposes?
    • How do journalists identify and use Social Media content and contributors?
    • How do journalists verify Social Media content (workflows, processes, aids, etc.)?
    • Which tools are used for verification (e.g. external tools, internal tools and customized tools)?
    • What obstacles are there to using Social Media content (in terms of reliability, difficulties, shortcomings, etc.)?
    • What would make Social Media content verification less difficult for journalists?
• For detailed analysis of the questionnaire and its outcome, see also Braendtzaeg et al (2014)

The following analysis focuses on the interviews conducted by Deutsche Welle. Out of the 20 interviewees, 12 of the interviewees work for Deutsche Welle, 2 work for NDR (another German public broadcaster), 1 works for the BBC, 1 for Agence France Press, 1 for Denmark’s Radio, 1 for Qantara, and two of the interviewees did not want to disclose their employers. Their age ranged between 21 and 50.

The interviewees work in different positions, they were mostly editors (12) and journalists (12) (see Figure 2). Besides the positions listed below, they were also able to click on “Other”, which they clarified as Social Media Trainer and Strategist, TV Coach, Social Media Editor, Editorial Manager of
R&D Media and Technology Department, and Presenter. This shows that people from a range of positions were included in the interviews. We did not ask developers, people from the graphics department nor did we talk to various heads of departments, as we believe they are not at the forefront of working towards verifying Social Media content.

In order to be able to ask people working on various projects, we wanted to make sure that we not only ask journalists working for breaking news. Therefore, we made sure to conduct interviews with people also working on feature stories (see Figure 3).

![Figure 2: Question: What is your position? , N=20 (multiple answers were possible)](image)

Figure 2: Question: What is your position? , N=20 (multiple answers were possible)
Figure 3: Question: What do you usually work on?, N=20 (multiple answers were possible)

After having looked at some of the sociodemographic information, the following paragraphs will provide an overview and highlights of the journalists’ use of Social Media and their current process of verification. For practical purposes, we will call all interviewees “journalists” in the following sections.

2.1.1.2.1 Usage of Social Media
All of the journalists work with Social Media platforms in their professional life, 18 out of the 20 interviewees use Social Media both professionally as well as privately and they all said that verification is “very important” for journalistic work.

Ask about the Social Media platforms they use, it became clear that most journalists focus on just a few Social Media platforms, mainly on Facebook, Twitter and YouTube (see Table 2). These are platforms that are used either daily or at least several times a week and for different purposes, including research, distribution of their own multimedia content and contacting the audience or experts regarding specific topics. Also, Google Plus is a platform that journalists use more frequently. The same can be said about blogs. Regarding blogs, the journalists mentioned that their selection of blogs they follow is quite dynamic as well as individual depending on articles they are working on, or their field of expertise. Other platforms they frequently visit include Rivva, Reddit, Tame.it, Newswhip, Virato, and Social Bookmarking Tools like Delicious, Readers like Feedly, Storify, Tweetdeck and Snapchat. The variety of the platforms depict that there are no common practices shared by all journalists. Rather, an individual choice of platforms and frequency of use highlight the various demands of journalists.

Table 2: Use of various platforms, N=20

Project Title: REVEAL
Project Coordinator: INTRASOF International S.A.
Concerning the motivation of using Social Media platforms (see Table 3), individual practices seem to be dominant here as well. What becomes obvious is that journalists use Social Media platforms for many different purposes and they tend to use them on a daily basis. Finding useful information, posting information, newsgathering and research, as well as staying in contact with others is mostly done on a daily basis in journalists’ professional use.

### Table 3: Motivation behind usage of Social Media platforms

<table>
<thead>
<tr>
<th>Platforms/Use</th>
<th>Never/Aprin almost never</th>
<th>Once a month</th>
<th>Several times a month</th>
<th>Once a week</th>
<th>Several times a week</th>
<th>Daily</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Twitter</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>YouTube</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Pinterest</td>
<td>15</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Instagram</td>
<td>12</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Flickr</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>14</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Tumblr</td>
<td>13</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Path</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Foursquare</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Blogs</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Google+</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Professionally, journalists use Social Media platforms as a

- Source to obtain information in the newsgathering process,
- Distribution platform for already existing content,
- Way to speak to or get in touch with the audience,
- Means to get in contact with experts/eyewitnesses (see Figure 4).
We also asked about further activities regarding their professional Social Media use and they answered by pointing out that they also detect new topics and trends that are being discussed on Twitter or Facebook, and they also push their own articles by distributing them and cross-promoting them via several Social Media platforms. Furthermore, the journalists establish conversations around particular topics, contact other people for further information and they monitor their own Social Media success of articles and/or tweets. One journalist stated: “I start the day reading my Twitter timeline and, in doing so, I feel extremely well informed about what is going on, what requires further attention and what I follow up”. Another journalist uses Social Media platforms completely differently: “We try to initiate debates that are relevant to us. Sometimes, we pick a topic, sometimes we see topics already discussed in communities. We try to contribute to both and use Facebook like a platform. We want to inform but we also want to be a discussion platform”. That means that journalists are also highly interested in collaborating with people outside of their own media organisations (also see Figure 5). Mostly, they ask for topics to cover or for user generated content.
When you use social media content as a way to collaborate, what do you do?

![Bar chart showing the responses to the question](image)

N= 20, multiple answers were possible

Figure 5: Question: When you use Social Media content as a way to collaborate, what do you do?, N=20, multiple answers were possible

Furthermore, journalists highlighted that they often ask for opinions and feedback regarding an already published article or question. Comments can then serve as a starting point for a discussion. Journalists ask their audience for opinions, for contact information in case the journalists want to use experts' takes on stories. As a tool to gather news, one journalist states that “I use Social Media mainly for monitoring topics and find out trends, stories and what matters to my audience. I let myself be inspired by my audience and get food for thought”.

Once the journalists have found content in Social Media, they integrate it into their own articles, they contact the people they received the information from, they double check with other sources or they contact witnesses (see Figure 6).
Further activities include staying informed, using the content as a lead to sources, finding out what matters most to the audience, checking out links/topics via most successful hashtags, checking out reactions concerning a specific topic, sharing information with followers, asking for further contributions, picking up ideas, angles etc. which can be researched.

All of these findings are relevant in order to understand how the journalists are using Social Media in their everyday working process. What can be concluded is that journalists use Social Media in many different ways. Some prefer the distributional aspect of platforms, such as Twitter. Other journalists have clearly stated to prefer Social Media as an initial platform for discussions regarding specific topics, e.g. on Facebook. Despite all the different ways of handling/using Social Media, it seems that most journalists prefer a handful of tools which they use daily. There seems to be a focus on specifically YouTube, Twitter, and Facebook. These various uses of Social Media have to be taken into consideration when considering the verification process as the individual uses might also apply to verifying Social Media content.

2.1.1.2 Verification Process

After having looked at Social Media use, this next paragraph focuses on the verification process of Social Media content. Regarding the journalists’ current process of verification, the practices over all the media organisations seem to not follow one specific route. Instead, many journalists seem to have found their own ways of how to verify content in Social Media (see Figure 7). Some of the media organisations already have specific guidelines regarding the verification process of Social media content. Others do not, and it somehow is part of overall editorial guidelines.

Our questions revealed that some journalists check whether content had been used by others (e.g. via retweets or shares), while others put greater emphasis on checking whether content comes from
verified accounts (indicated by e.g. verified badges on Twitter or on Facebook). The figure below provides an overview.

Figure 7: Question: How does your verification process look like right now? N=20, multiple answers were possible

What is interesting about the verification process is that journalists use different parameters when trying to verify content on each of the Social Media platforms (see Table 4). On Twitter, journalists check the contributor of content primarily by looking for a verified badge on their profile, they also look at the number of followers and the relation of the number of followers and the number of following. They also look at the timeline and the quality of the tweets over a certain amount of time. The content is checked by looking at the quality of the link posted, the picture or also the number of retweets. On YouTube, the parameters are a bit different. Here, the journalists focus on the videos posted by a contributor but they also look at the profile of the person posting videos of interest. On Facebook, journalists like to check out the “About” section to learn more about a contributor.

Table 4: Verification Process on various Social Media accounts

<table>
<thead>
<tr>
<th>Platform</th>
<th>What journalists look at (Verification Parameters) - Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter</td>
<td>Verified Badge on Profile, Number of followers, Who is following exactly (trustworthy people?), quality of timeline, content of tweets over longer period of time, links used in tweets, number retweets, Twitter lists</td>
</tr>
<tr>
<td>YouTube</td>
<td>Other videos being uploaded, profile of producer, number of subscriptions</td>
</tr>
<tr>
<td>Facebook</td>
<td>‘About’ section, number of friends, verified badge</td>
</tr>
</tbody>
</table>
Every platform is used differently in the verification process. Currently, there are no standard ways of verification. Since verification is general is not new for journalistic processes, however, the journalists seem to apply what they know from the traditional ways into the Social Media context.

- They look closely at the person who provides the information (=contributor). It is important to “have a really close look at the contributor”, one interviewee points out. The primary thing journalists look closer at is WHO is providing the information by e.g. contacting the people via phone or trying to find out more about the sources, e.g. on their websites or networks such as LinkedIn.
- The journalists look closely at the content itself (=content) by cross-checking on other platforms or looking at potential manipulations.
- They check additional information about the topic (context) by cross-checking with other platforms and reviewing multiple sources as part of their verification process. Some of the journalists pointed out that „using the „good-old” Google“ is a typical way of finding out more about the context of content they find in the social web.

It seems like the journalists follow the three steps to verify content in Social Media: contributor, content, context. Sometimes, journalists also use tools or platforms that offer assistance in verifying content (see Table 5).

**Table 5: Question: What tools do you use in order to verify content on the social web? N=20**

<table>
<thead>
<tr>
<th>Photos</th>
<th>Videos</th>
<th>Contributors</th>
<th>Texts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Image Search, Tin Eye, Exif, Topsy, Tungstene</td>
<td>Google Maps, Streetview, similar YouTube videos?, Storyful</td>
<td>Phone call, e-mail, Google, checking profiles (followers, tweets, lists), other SN accounts?, Namechecker and the like, Google, Xing etc, People directories (to get contact details and check on contributors) such as peekyou.com</td>
<td>Google text search, phrase search, manually contacting them: second source (not via Twitter)</td>
</tr>
</tbody>
</table>

Journalists point out that especially photos pose a number of challenges when trying to verify images, as tools like Tin Eye or Google Image Search only tell you if a picture has been manipulated. They do not, however, tell you when a picture has NOT been manipulated, but is used in another context or time / period. For textual information, journalists tend to use Google search for already existing phrases and when it comes to videos, journalists check whether they can confirm places via Google Maps’ Streetview or they check YouTube for similar content. Also, some of the journalists mentioned Storyful as a way to outsource their verification of Social Media content (for more information, go to the market analysis paragraph of this Deliverable).

Some of the journalists have stated that they are unhappy about the intransparency of algorithms (e.g. that work for Newswhip, Trendsmap, Tweetdeck, Topsy or Flumes) and that they would like to know how the algorithms work exactly in order to understand displayed results such as lists (e.g. trending lists).

What is most useful in the verification process seems to be the “journalistic common sense”, one journalist claims. Tools can be helpful but cannot replace traditional journalistic workflows and...
procedures. Manual work needs to be considered as well. Sometimes, all the parameters on various Social Media platforms do not help in the verification process because at times, the quantity of e.g. tweets or followers does not matter. Rather, a quick phone call can answer more questions than a sophisticated algorithm.

Asking journalists about what they would like to have at their disposal with regards to verification support, the quote that stood out the most was: „I would like to have the ultimate truth-machine“. In addition to that, journalists would like to work with an integrated platform that displays content from as many Social Media platforms as possible, providing information of the various content types that are contributed by a single contributor on individual stories (e.g. show the video, tweets and photo a user has posted on various social networks). Journalists also expressed a need to be able to personalise this integrated platform depending on topics they are interested in, people they want to follow, regions they want to focus on etc. Furthermore, journalists stated that they would like to see instantly how a story has spread, for example who posted about an event first and then how this story spread via all Social Media platforms. Also a credibility score about contributors (such as Klout, but with a clear and open algorithm) would be helpful in order to understand the relevance of a contributor. Moreover, they would like some easy way to contact people / profiles for verification purposes (i.e. get quick and direct access to contact details such as phone number, mail address etc.). Some of the obstacles that make the verification process a little harder (see Table 6) are primarily: (1) there is too much information out there and it becomes difficult to understand what the audience is currently talking about. Also, it is more difficult to find out what the experts are saying. (2) Many of the platforms the journalists are working with are not really geared towards journalistic needs. While Twitter seems to be the most important platform for the newsgathering process at the moment, there are also others like Facebook which are not necessarily geared towards specific journalists’ needs.

Table 6: Question: What are some of the major obstacles you encounter while verifying content? N=20

<table>
<thead>
<tr>
<th>Obstacle</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too time-consuming</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Too much information</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Too little verified information</td>
<td>4</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Difficult to separate marketing from real information</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>No one platform that suits all needs</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Many platforms not geared towards journalistic needs</td>
<td>2</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>External verification tools not allowed in my organization</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>
In conclusion, quite a few factors have to be taken into consideration when looking at the verification process. Journalists point out that especially photos pose a number of challenges when trying to verify content. External services like Tin Eye or Google Image Search retrieve similar images on the web and then the journalists still have to decide whether the pictures are manipulated or not. The tools do not help in finding out whether a picture has been manipulated or not. There are not just a few parameters that have to be checked, the verification process differs from platform to platform. On YouTube, journalists look at specific factors that might only play a minor role on Facebook or Twitter and vice versa. Finally, all journalists stated that verification is “very important” but some obstacles seem to hinder a smooth verification procedure, such as the time it takes to verify content, the abundance of information that needs to be verified or the fact that there is no single integrated platform that solely focuses on verifying content coming from all sorts of Social Media platforms. Taking all this into account, a market analysis of two of the platforms mentioned the most is intended to provide further information of some of the functionalities that seem to be valued highly by journalists.

2.1.1.3 Market analysis

Based on the literature study and the conducted interviews, it was possible to narrow down some of the platforms in the market that stand out: namely the ones that provide many of the features expressed by journalists as well as the ones that are widely used already. For a list of further competitors, please see Appendix A. Here, we want to focus on a selection of platforms of growing popularity, as our research has shown.

2.1.1.3.1 Storyful

Storyful combines automatic and human-intervened content verification. Journalists working for Storyful use a combination of technology and traditional journalism skills to be able to verify Social Media content. Storyful claims to be “the world’s first Social News Agency”. The company was founded by Mark Little in 2010 and has since then partnered up with media organisations, such as ABC News, Reuters and the New York Times. A team of reporters monitors and engages with communities on Twitter, YouTube, Google+ and any social platform that a partner is interested in (see screenshot of the Storyful landing page).

![Figure 8: Screenshot Storyful Landing Page, December 2013](image-url)
In the process of verification, Storyful has listed some of the journalistic questions they raise when trying to verify YouTube videos\(^9\):

- Where is this account registered and where has the uploader been based, judging by their history?
- Are there other accounts – Twitter, Facebook, a blog or website – affiliated with this uploader? What information do they bear to indicate recent location, activity, reliability, bias, agenda?
- How long have these accounts been in existence? How active are they?
- Do they write in slang or dialect that is identifiable in the video’s narration?
- Can we find WHOIS information for an affiliated website?
- Is the person listed in local directories? Do their online social circles indicate they are close to this story/location?
- Does the uploader ‘scrape’ videos from news organisations and other YouTube accounts, or do they upload solely user-generated content?
- Are the videos on this account of a consistent quality?
- Are video descriptions consistent and mostly from a specific location? Are they dated? Do they have file extensions such as .AVI or .MP4 in the video title?
- Are we familiar with this account – has their content and reportage been reliable in the past?

Tools and technologies aiding in answering the above questions accurately and quickly would be much desired by any organisation involved in the process of verification and thus constitutes a huge market gap and potential.

When they discover compelling content – videos, photos, maps or tweets - Storyful journalists search for key details, such as the contributor, the content and the context that will help clients to determine whether they can use the content in their own offerings (e.g. broadcasts, websites).

Storyful clients get to see right away which information Storyful journalists have been able to accumulate. Storyful journalists assign statuses to the content (such as “cleared”, “licensed”, “pending”). They show all the information they have found and specifically focus on three different elements when verifying the content, namely source, location, and date (see Figure 10).

This process includes some of the following workflows:

- Engagement with the member of the community who shared the video in an effort to establish the identity of the original uploader.
- Translation of every word that comes with a video for additional context.
- Review of the uploader’s history to see whether he/she has shared useful and credible content in the past, or if he/she is simply ‘scraping’ other people’s content (e.g. videos).
- Use of Google Street View/Maps/satellite imagery to help verify the locations in a video.

\(^9\) source: http://blog.storyful.com/2012/04/24/inside-storyful-storyfuls-verification-process/#.U2C9uCSABBL
- Consultation with other news sources to confirm whether events in a video happened as they were described.
- Examination of key features in a video such as weather and background landscape to see if they match known facts on the ground.

![Figure 9: Screenshots of detailed information of one YouTube video](image-url)
Figure 10: Storyful uses three elements of verification

2.1.1.3.2 TinEye/Google reverse images
TinEye and Google reverse images assist journalists in finding similar images on the web. However, these services do not help in analysing whether an image is an original or if an image has been modified. TinEye also does not recognise outlines of objects or perform facial recognition, but recognises the entire image, and some altered versions of that image. Another tool, Fotoforensics, does show data about the images, but the tool cannot help in verifying the content, it can only be used as a microspace for looking at some of the metadata provided.
2.1.2 Enterprise field
Software AG is hosting their own social media communities. The overall entry point of those social media communities is the TECHCommunity (http://techcommunity.softwareag.com). The largest stand-alone community is the ARIS Community with currently more than 300.000 registered users (http://www.ariscommunity.com/). In addition, Software AG is using and linked to several external hosted social media channels:

Table 7: Social media channels Software AG is involved in

<table>
<thead>
<tr>
<th>Channel</th>
<th>Facebook</th>
<th>Twitter</th>
<th>YouTube</th>
<th>LinkedIn</th>
<th>Communities</th>
<th>Blogs</th>
<th>Slide-Share</th>
<th>Flickr</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIS Community</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Adabas Community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>B2B.com</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CentraSite Community</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ehcache</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Community</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reality Check</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Argentina</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Brasil</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG France</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Germany</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Software AG Global</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Software AG Investor Relations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Isreal – SPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Italy</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Latin America</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG North America</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Spain</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG Türkiye</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG USA, Inc. Recruitment</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software AG University Relations</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TECHcommunity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terracotta</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>webMethods Community</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

So the methods described below that have been used in the requirements process and the interviews had to be treated in the sense that all social media activities of Software AG are included, especially not excluding the external channels (like Twitter).

As mentioned above, there are many different methods and techniques in the requirements process. It is often described as a series of activities such as elicitation, modelling, triage, specification, and verification (Hickey & Davis, 2002).
The method chosen by Software AG has derived from the process Software AG is practising in its long-lasting history as a software company. The so-called “Feature Management Program”, a set of various elicitation techniques, turned out to be its best approach to, in accordance with (Zowghi & Coulin, 2005), learn and understand the needs of users with the ultimate aim of communicating these needs to the system developers. Software AG’s requirements elicitation involves activities like gathering customer inputs through close contact, market or competitor analyses. Interviews in the internal department of “User Experience” are a very effective module to get helpful feedback by end-users generating a range of criteria on the basis of empirical data using an explorative factor analysis.

This “maturity” as a well-experienced software company leads to identify the adequate requirements methods and to an improved understanding of end-user needs.

Following this briefly described techniques and with the deep insight into the internal work processes, Software AG developed a requirements process comprising the following techniques:

2.1.2.1 Observation and brainstorming

With this technique the analyst observes the actual execution of existing processes by users without direct interference (Zowghi & Coulin, 2005). The observer finds out how users are retrieving, maintaining and compiling the information during the daily work.

To include such an observer in the analyst’s team, Software AG arranged a brainstorming session as a second technique with a colleague from the ARIS Community. In an informal discussion and with the knowledge of the ARIS Community expert as an observer, this technique allowed discovering existing needs and developing preliminary research questions for the project and targeting system. Finally as a result of observation and brainstorming, the definition of stakeholder groups was generated (see Section 2.1.2.2).

2.1.2.2 Interviews

In conjunction with and after observation and brainstorming techniques, Software AG started with the most traditional and commonly used techniques for requirements elicitation (Zowghi & Coulin, 2005): the interviews.

Intervies belong to the type of the direct approach with the purpose is to enhance the understanding of the problems of a system that is currently in use (Qureshi, Asim, & Sahar, 2011). In literature interviews are divided into three categories:

- Unstructured interviews are conversational in nature where the interviewer enforces only limited control over the direction of discussions (Zowghi & Coulin, 2005).
- Structured interviews are conducted using a pre-determined set of questions to gather specific information (ibid).
- Semi-structured interviews are conducted with a fairly open framework which allows for focused, conversational, two-way communication. They can be used both to give and receive information.

Software AG preferred the direct approach with the method of the semi-structured interview with open-end questions. In an open-end question the interviewer poses a question, and then allows subjects to answer as they wish (Goguen & Linde, 1993). Using the description of the interview process in the following section, the reasons for this method will be pointed out.

Nature of Interview Framework
After several brainstorming sessions in the research team together with some colleagues of the community team in Saarbruecken and in Darmstadt, Software AG developed a semi-structured interview framework. The framework was commented on, completed and finally improved by SINTEF.

It includes the elements of both structured and unstructured questions. Some questions were predetermined to gather short and specific information. For example: “Which specific people, organizations, offers do you follow in the media?” Some questions were created as an impetus to allow both the interviewer and the person being interviewed the flexibility to probe for details or discuss issues. For example: “If you could wish a method/tool that supports users in verification processes of information what would be interesting for you.” And the interviewer adds: “Below you see a list of probably useful information. There is no need to go through every single point separately. They only serve as inspiration…”

Some of the questions were created during the interview for the purpose of being open for a narrative spontaneous conversation.

The following benefits made it easy to decide for the semi-structured interview:

- All participants are Software AG colleagues. This allows a relaxed conversation in a familiar atmosphere, with this being less intrusive.
- The information obtained will provide not just answers, but the reasons for the answers.
- Sensitive issues may be more easily discussed.
- The interviewer is better able to respond to individual characters, whether it’s an open-minded, friendly and supportive or a rather taciturn respondent.

Central Research Questions

The underlying research questions can be characterized by the following:

- How do our colleagues behave in social media, especially when it comes to product development, marketing activities, issue or thought leader identification?
- What would make their work easier in the context of the verification process?

Firstly the active and passive behaviour of different groups in social media is interesting. Secondly, the meaning of verification processes for the interviewees was to be investigated.

The overall central questions were further specified. This list below shows the questions which are divided in six parts:

1. General User Behaviour
   a. Please tell me how you use social media for your daily work?
   b. Please describe your activities and the frequency in the communities.
   c. Which specific people/organizations/offers do you follow in the media?
   d. Do you specifically verify content that comes from social media platforms, and, if so, how?
2. Estimation of the Significance
   a. In your view, what role do social media play for the development and the distribution of our products
   b. Did you make your own experiences (real-life examples)?
3. Requirements
4. Handling
   a. Imagine you're starting up your computer and a mashup with brandnew data of the
described modalities is appearing on the screen (e.g. divided by industries).
      i. Would you welcome it?
      ii. How do you think such a mashup should look?
      iii. Could you imagine other kinds of reporting forms, like dashboards, reports,
alerts…?

5. Real-life Scenarios
   a. Imagine a new product version is releasing and there is an overwhelming number of
negative contributions.
      i. What do you think of a tool revealing backgrounds and data of contributors?
      ii. Which kind of information would be helpful?

6. Crowdsourcing and Gamification
   a. Are you participating in crowdsourcing and gamification mechanisms in general? If
so, please tell us some examples.
   b. Do you think it’s useful to reward contributors with reductions or license renewals of
SAG products?
   c. Could gamification elements provide additional information of higher-level modalities
like trustworthiness, completeness, quality, or controversiality?

Participants

The eleven participants have been selected on the basis of different professional groups or
noticeable activities in social media. They are divided into the following groups:

- Community Management
- Product Management
- Product Marketing
- Sales
- Social Media Activists/Experts

The consideration that these groups will derive benefit from REVEAL played a part as selection
criterion. The interviewees are spread across different countries and subsidiaries and from diverse
cultural backgrounds, like Austria, Bulgaria, Canada, Germany, and the USA. They were not paid for
their participation.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Function/role in SAG</th>
<th>Subsidiary</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Product Marketing</td>
<td>Saarbruecken/Germany</td>
</tr>
<tr>
<td>B</td>
<td>ARIS Community</td>
<td>Saarbruecken/Germany</td>
</tr>
<tr>
<td>C</td>
<td>TECHcommunity</td>
<td>Darmstadt/Germany</td>
</tr>
<tr>
<td>D</td>
<td>Sales</td>
<td>Vienna/Austria</td>
</tr>
<tr>
<td>E</td>
<td>Product Management</td>
<td>Saarbruecken/Germany</td>
</tr>
<tr>
<td>F</td>
<td>TECHcommunity</td>
<td>Sofia/Bulgaria</td>
</tr>
<tr>
<td>G</td>
<td>Product Marketing</td>
<td>Reston/USA</td>
</tr>
<tr>
<td>H</td>
<td>Product Management</td>
<td>Saarbruecken/Germany</td>
</tr>
<tr>
<td>I</td>
<td>Product Management</td>
<td>Berlin/Germany</td>
</tr>
</tbody>
</table>
Interview preparation and procedure

Once the questions have been defined and the specific stakeholders identified, then the actual elicitation of requirements began with the interviews. The interviews were conducted by two persons of the research team. They arranged an appointment for the interview which should ideally be in form of a face-to-face-meeting to come together in an open conversation. The main advantage of face-to-face or direct interviews is that the researcher can adapt the questions as necessary, clarify doubts and ensure that the responses are properly understood, by repeating or rephrasing the questions. The researcher can also pick up nonverbal cues from the respondent. And it sometimes provides opportunities for meeting informally to chat.

Unfortunately because of the distance it wasn’t always possible to meet the participant in the same room, so a phone call had to replace a direct encounter in some cases.

The interview framework were designed so that it took no more than one hour. Particular importance was placed on extensively briefing the interview partners before the interview in the form of background information of the REVEAL project. The questions were made available for the participants in advance so they had the chance to consider the questions. It turned out to be that this was not an obstacle for spontaneous narratives in the interview situation.

One interviewee who is involved in the design of the interview framework served as a test person. It offered the possibility to make modifications before involving other participants.

Each session was conducted as follows:
1. The participant was greeted.
2. The moderator introduced (if necessary) him- or herself.
3. The participant was given an overview of the project and the interview framework.
4. The participant was then asked to answer the described questions.

2.1.2.3 Process of data collection

After conducting the interviews the interviewer summarized the most essential quotes in a list of requirements. Accordingly, these requirements were tabulated in a comprehensive categorized matrix (see Appendix D) which shows feature groups sorted by three categories:
- requirements concerning the contributor
- requirements concerning the content
- technical requirements and general handling.

And all these quotes in turn are related to the topics „thought leadership“, “feature requests/bug messages/comments on the product” and “contributor and “information research”.

After creating the matrix, Software AG developed a symbol which reflects the coherences between contributors, content and the identification of their backgrounds: the identification tree (Figure 10 and 11). It gives answers to the following questions:
- Which contributors need to be determined in detail?
- For which reason?
- Between which attributes of contributors has to be distinguished?
- Which roles do contributors play in social media?
- Where else are they connected?
- What influence/effect do contributors have on social media?
- Are there cross-connections between contributors?
- Which contributor initiates a discussion?
- Can causes, courses and consequences of discussions be clarified?

The second identification tree “topic” (Figure 11) concentrates more closely on the various topics and their coherences between contributors and communication processes.

The tree in this context can therefore be concluded as an efficient symbol for rootedness of influences, ramification of connections and growth of networks.

Finally, from these trees, we derived a set of mini-scenarios which will be described in the section 3.2 as an essential result of the requirements elicitation.

Figure 11: Identification tree “contributor”
3 REVEAL Requirements

3.1 News requirements

Based on the findings of the literature study, the interviews, and the market analysis, we have gathered a number of requirements for the journalistic use case. Before we present the requirements as a table, we present some mini scenarios that are tightly focused on specific end user challenges and that can be seen as typical journalistic workflows. By going through all the different mini-scenarios, the requirements become clear. All the requirements will be displayed as user stories at the end of this section.

3.1.1 Mini scenario 1: Newsgathering support

Especially in the newsgathering process, verifying content is one of the key essentials of the journalistic workflow. In order to find relevant topics, journalists have to be able to search for content in multiple social networks. The subsequent example is to illustrate this: “As the search for Malaysia Airline’s missing Boeing 777 moves into its 11th day, a multitude of theories about the plane’s fate are circulating on forums and Social Media”, the BBC reports\(^\text{10}\). The journalist starts up a new REVEAL “task” to monitor Social Media for this news story. The journalist provides REVEAL an initial set of keywords (e.g. #MH370, #aviation, #aircraft etc.) and the initial temporal periods for the real-time crawl (12 days from now) and historical search (the past 11 days) operations. The temporal period can be extended dynamically if the news story runs longer than expected.

3.1.2 Mini scenario 2: Revealing the Contributor

One of the key findings of the interviews conducted encompasses the relevance of the contributors of information in social networks. Knowing the source of a video, a picture or a tweet is essential for journalists in the verification process. Once the source is found, journalists might want to contact the person, or if there is no time, find out whether the contributor is trustworthy or influential in their respective Social Media community (by looking at features or attributes). Based on this knowledge, journalists can determine next steps.

As Malaysia Airlines flight MH370 disappears with 239 people on board, the search area has been extended and various other countries committed equipment and personnel in order to assist in finding the plane. The journalist is currently faced with information from various sources in social networks, such as family members of the missing people on board threatening to go on a hunger strike, airlines releasing more information about the flight, news organizations, politicians and aviation administrations publishing their latest news on possible scenarios while waiting for the discovery of the plane.

In order to (1) keep up-to-date but also to (2) be able to determine which information or video/tweet/photo can be viewed as verified or not, the journalist starts to search for relevant contributors, including

- Opinion-leaders, who provide the most trustworthy information in forms of links, pictures, and videos
- Social graphs for people related to news story keywords (Twitter, YouTube, Flickr)

\(^{10}\) http://www.bbc.com/news/magazine-26609687
### 3.1.3 Mini-scenario 3: Revealing Multimedia Content

Based on the interviews, journalists find numerous media items in social networks that have been posted by different users and need verification. Photos, videos, as well as text pose different challenges for journalists in the verification process. Currently, different tools can help to assist. Journalists use Google Images, Tineye but also platforms such as Storyful in order to be able to say whether an image or a video is true or false.

#### 3.1.3.1 Mini-scenario | Picture

Last pictures of the pilot, worried family members, satellite pictures – as the search for the plane and passengers continue, countless pictures from different sources have been published. First of all, the journalist will launch a crawling job searching for photos linked to the news as they are breaking. The search result will include photos based on the search terms #mh370 from Twitter, Instagram and Flickr. Secondly, the journalist will use the pictures to send for provenance checking. Suspect items will be analysed. The results will include sets of similar content items, with links to source content and authors, and reports on any detected image manipulation.

#### 3.1.3.2 Mini-scenario | Text

Journalists will use the text documents (web pages, PDF documents, Social Media textual content) for provenance checking. The results will include sets of documents with a similar writing style.

### 3.1.4 Mini-Scenario 4: Revealing Location

From a geo-spatial perspective, the missing flight MH370 is particularly interesting, as many countries are involved in providing information. On the one hand, there is a Malaysian airline assisting in information coming from Kuala Lumpur. Then there are Chinese politicians and the administration that are involved due to many Chinese passengers missing and the arrival airport being Beijing. Also, we have other nationalities missing (e.g. French). Countries, such as the US and Australia are involved as they assist in finding the missing plane. And then there are the families, waiting for news. Journalists flag this event as an emerging news story and want to maintain a situation assessment picture to help them with frequent news reports as the event unfolds.

#### 3.1.4.1 Mini-scenario | Mapping workbench

The journalist uses a ‘mapping workbench’ to display real-time content for news related to the missing jet on a map display. The journalist specifies the spatial area (e.g. Indian Ocean, as well as China, Malaysia, Pakistan, and Australia) and news story keywords (#missing jet #MH370). A display is then available at which journalists can see clusters of relevant content items, set up filter criteria to adjust displayed clusters and drill down to see specific content and author details. Easy links to off-the-shelf tools are provided such as Google Translate and Google Image Search.

#### 3.1.4.2 Mini-scenario | Trending workbench

The Journalist uses a ‘trending workbench’ to display global trending topics and identify a breaking news story. Besides the MH370 story, other breaking news need to be taken into consideration, such as the Budget Day 2014. Journalists on the newsdesk start to

- Track other news constantly. A timeline display shows journalists the global trends and news trends, along with short summaries of content,
3.2 Enterprise requirements

3.2.1 Preliminary Note: Software AG Specifics
Before presenting the findings of the requirements elicitation there is a need to note the specifics of Software AG, based on the interviews, underlining the character of a business software company in contrast to Deutsche Welle, an organisation involved in media production and distribution.

- Software AG is extremely customer- and market-oriented and is therefore primarily focused on the objective to continually deliver and improve the best products and services to its customers. This also means that all activities aim to grow in revenue and profit. In our case, therefore all activities in social media pursue this goal.
- In contrast to Deutsche Welle, information retrieval from Social Media doesn’t belong to the core task of Software AG employees and for this reason they are not particularly active in the general purpose social media channels (Facebook, Twitter, etc).
- Software AG develops and sells complex products which require explanation (complex sales). For that reason, social media don’t usually result in final purchasing decisions.
- Our customers are hesitant and restrained handling social media. This especially regards decision-makers more than real product users.
- Software AG is moving in the B2B, not in the B2C area, where social media play a minor role. There are for example not masses of positive or negative reactions expected.
- In a global company like Software AG, the request for revealing backgrounds and data of contributors might be a special country-specific bias. In some regions, employees accept people as experts without detecting their expertise in social media channels.
- For some of the following scenarios it is not possible to concentrate on one specific social media channel only. Often, information is mixed between the own social media communities and external channels. So we have to think of “virtual” communities that do exist in various social media channels but do share similar content and have similar purpose.

Therefore, the focus in the case of SAG is in the enterprise community networks and not so much in the general purpose social networks, although there is the expectation that useful material will be discovered in these social networks as well.

3.2.2 Mini-Scenarios
In Section 2.1.2 we extensively described the derivation of requirements elicitation and the data collection. Just to summarize and as a quick reminder, we depict the process of requirements gathering in the following Figure 13.
Deriving from the described methods, in this section we will define some mini-scenarios that are tightly focussed on specific end user challenges in the enterprise community area.

From an end-user perspective defining mini-scenarios will allow the REVEAL project to create a variety of small demonstrators each focussed on a specific end user domain challenge. These demonstrators provide a way for end user partners to understand concretely the business value that could be added by different combinations of technology components.

The mini-scenarios are defined from the Software AG perspective which means that they are focused on the business value for Software AG in general and could probably be used as reference models for other B2B companies (cf. Section 2.2.2), especially in the IT industry.

After the interviews it became clear respondents and customers aren’t only active in the TECHcommunity (the main SAG online community network) but also in various social media in other communities. For that reason it is necessary to consider the entire social media sphere, with particular emphasis on an integrated TECHcommunity.

Each of the following mini-scenarios is described by means of the parameters end-user purpose, identified role, profile, activity, and outlook.
3.2.2.1  Mini-scenario | “Newbies”

**End user purpose:** Each new contributor in the TECHcommunity or in other social media in the context of Software AG and its products or services could be a prospect. For that reason REVEAL has to create a clear profile of the contact. Is he really interested in our products or is he probably a competitor? Could he be a journalist or analyst?

**Identified role:** new arrival

**Profile:** profession, company, region, industry, function in company, business interests, connection to social media, connection to and influence in special interest groups, user of similar products

**Activity:** creating profiles and interests

**Outlook:** input for other scenarios

3.2.2.2  Mini-scenario | Customer relation

**End user purpose:** Assumed, an existing customer is searching for a solution for a specific problem in his enterprise processes. Regarding this, he is posting a respective comment in the TECHcommunity or in other social media. Probably SAG products would provide a solution. Other customers could benefit from it.

**Identified role:** customer

**Profile:** profession, company, region, industry, function in company

**Activity:** finding sales or product management

**Outlook:** connecting to internal management system (optional) or internal social media tools like Yammer

3.2.2.3  Mini-scenario | Support

**End user purpose:** Assumed, an existing customer detected a bug in a product and posted it in the TECHcommunity. People from support contact the contributor and forward it to the development where a new patch is provided in the end. Therefore the product “Process Live” is a good example. It’s the first SaaS only product of SAG. Instead of offering customers traditional support it is linking to the ARIS Community, the customers shall leverage the knowledge of the community members for their projects. As an additional source of information there is a twitter channel (@processlive) available, that should be followed by customers.

**Identified role:** requester

**Profile:** contact to support

**Activity:** finding similar postings and users/similar problems

**Outlook:** product improvement

3.2.2.4  Mini-scenario | Analysts

**End user purpose:** Analysts (e.g. from Gartner) have always played an important role in the history of Software AG. Product Management and Development have to orient to the publications and the points of view of analysts. Purpose of this scenario is to identify the analyst in all media to keep in contact to him and early trends.

**Identified role:** analysts
Profile: influence, trustworthiness

Activity: finding trends (of the analysts), current interests

Outlook: get a good position in analysts “magic quadrant”, contributor monitoring, product improvement

3.2.2.5 Mini-scenario | Negative or positive discussions

End user purpose: Assumed, somebody initiates a negative discussion about a new release. To keep control of the activities the tool has to reveal the reason and probably the motivation behind the contributor behaviour. Alternatively: a new product has been released and marketing needs to assemble positive feedback.

Identified role: has to be discovered

Profile: profession, company, region, industry, function in company, connection to and influence in social media, connection to and influence in special interest groups, previous contacts to support or sales

Activity: finding people behind the discussion/story, finding duplicates

Outlook: moderating, intervene

3.2.2.6 Mini-scenario | Event: Innovation World

End user purpose: “Innovation World” is the biggest company-owned event taking place once a year in the USA. Through targeted actions in the social network, Software AG intend to draw attention in the run-up to the event attracting as many visitors as possible to the event and making people aware of the topics and highlights during the Innovation World.

Identified role: has to find out

Profile: profession (journalist, analyst, customer…), company, region, industry, function in company, connection to and influence in social media, connection to and influence in special interest groups, previous contacts to sales …

Activity: finding people who inform about the event, finding out reactions to postings

Outlook: following up postings, bringing people to the event

3.2.2.7 Mini-scenario | Innovation gathering

End user purpose: For the development of products and services it is essential to keep up-to-date in the business areas of Software AG. Not only the brand-new topics but also the initiator of trends and the thought leaders has to be observed in the social media. Influence and trustworthiness are essential in this scenario.

Identified role: thought leader (score), expert, competitor


Activity: finding/creating and connecting to TECHcomunity; extracting innovative ideas/posts; finding other people (to follow them), identify trustworthiness

Outlook: product improvement
3.3 DW and SAG merged Table of REVEAL requirements

After the requirements elicitation and data collection and after creating mini-scenarios, we present here a detailed list of user requirements, where Deutsche Welle and Software AG have consolidated their findings.

At this step, each requirement gathered, has been examined with respect to the following:

- **Clarity:** The requirement should be unambiguously interpretable and straightforward to map to a system property.
- **Completeness:** The requirement should provide sufficient details and background in order to be of value to the system design.
- **Conflicts:** The requirement should not create conflicts with other requirements.

We categorise the requirements based on:

- requirements related to revealing more about the Contributor;
- requirements related to revealing more about the Content;
- requirements related to revealing more about the Context; and
- general requirements.

We map them to:

- the related mini-scenario(s) as derived previously;
- the role(s) identified;
- the use case(s): news or/and enterprise;
- the WP(s) responsible for their implementation, and
- the priority (1 being the highest).

For the majority of the requirements, the technical partners provided their comments/clarification with respect to their envisioned implementation.

3.3.1 Requirements for revealing Contributor

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Revealing the</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Contributor</td>
<td>Thought leader (score), expert, competitor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Innovation gathering</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description:** I can search for contributors based on search terms (e.g. #MH370).

**Comments:** We will maintain information with related keywords for each contributor. WP2 will allow search. WP6 UI will provide a search interface and display results.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>Revealing the</td>
<td>Contributor detector</td>
<td>News</td>
<td>WP2/6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Contributor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description:** I can obtain trusted networks on a story by story basis (based on search terms).

**Comments:** WP2 will allow search. WP6 UI will provide a search interface and display results.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3</td>
<td>Revealing the</td>
<td>Contributor detector</td>
<td>News</td>
<td>WP4/5/6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Contributor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description:** I can search for contributors based on geo-location and see them displayed on a map.

**Comments:** Related to Geospatial Context extraction, annotation of content items with a
geolocation. WP5 HTTP will allow assessments to be specified with one-off search terms. WP4 geo-context storm processes. WP5 map visualization of situation picture (i.e. search results) and filter options. Note for search this picture is static. For streaming picture is real-time updated. WP6 UI will allow geo-search terms to be entered and WP5 HTTP endpoint called.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/3/4</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can see whether the person is a first-hand source (on sight) and whether it is true what he/she posts (video/audio/text) (e.g., location). Results also linking to REVEAL-external verification-tools (e.g. if Storyful has uncovered a story/photo as false).

**Comments:** First hand source is interpreted as: Is it an original post (e.g. tweet) or a re-post (retweet)? WP4 geo-classifier will identify in-situ/third-party type reports. WP2/3 text & image forensics could help for provenance questions.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A5</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/4</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can receive information about the contributors I want to know more about and I can see a trust score.

**Comments:** Websites, Bio, picture, further activities on other social networks, if possible phone numbers. This is particularly important in conflicts, for revealing vested interests or false profiles. And concerning SAG: A score with a certain number of points, which is visible close to the user data, would be significant. You could define a threshold value. Is the user positioned above this level than the product manager gets an alert: Follow this user!

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A6</td>
<td>Revealing the Contributor</td>
<td>Community detector</td>
<td>News &amp; Enterprise</td>
<td>WP2</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can receive information about the community the contributor is located in (topics discussed in that community).

**Comments:** We will keep a profile of the community.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A7</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/4/5</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can see who posted the most retweeted tweet based on search terms (e.g. #MH370).

**Comments:** Once relevant tweet is found. Simple Query to Twitter API. WP4 & WP5 will count aggregated statistics and rank most frequent hashtags/URLs/topics.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A8</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/4/5</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can see who posted most about a topic based on search terms (e.g. #MH370).

**Comments:** Most influential users will be based on a variation of topic based page rank. WP4/5 could probably also count user frequency statistics.
<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A9</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Innovation gathering</td>
<td>Thought leader (score), expert, competitor</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description:** I can determine the most influential contributors (opinion-leaders). E.g. Followers, Friends, Mentions, RTs, Shares on Facebook.

**Comments:** Most influential users will be based on a variation of topic based page rank.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A10</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News</td>
<td>WP2</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can find the most relevant links posted by the most influential contributor.

**Comments:** Rank users on influence, pop the first one and select all the URLs that were found in their texts.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A11</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News</td>
<td>WP2/3</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can identify the most relevant photos posted by the most influential contributor.

**Comments:** The multimedia indexing and search module in WP3 will support restricting the search to given contributors, of whom the identification of the most influential contributors will be provided by WP2.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A12</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News</td>
<td>WP2/3</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can identify the most relevant videos posted by the most influential contributor.

**Comments:** The multimedia indexing and search module in WP3 will support restricting the search to given contributors, of whom the identification of the most influential contributors will be provided by WP2.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A13</td>
<td>Revealing the Contributor</td>
<td>Content detector</td>
<td>News</td>
<td>WP2/3</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can find the most relevant topics (e.g. other #malaysia #airplane) posted by the most influential contributor.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A14</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/3/4/6</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I can see all the information about an influential person (e.g., website, profile picture, phone number etc.).

**Comments:** We interpret this to mean: “Let me see the complete profile of the user”. WP2/3/4 should keep source ID so can provide web URI's to content/people. There are also privacy issues to take into consideration.
<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A15</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description**: I can filter influential contributors based on # of followers, # of tweets.

**Comments**: It can be done by consulting with PServer (WP2 module) that maintains profiles and/or twitter API.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A16</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/6</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description**: I can see what my trusted network is talking about/what information is being shared.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A17</td>
<td>Revealing the Contributor</td>
<td>Community detector</td>
<td>News</td>
<td>WP2/3/4</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description**: I can find various communities discussing different perspectives of one story (e.g. family members of missing flight vs. journalists vs pilots).

**Comments**: Perspective will be based on Topic/Subtopic detection per community. Sub-Topics will be aligned with user groups. It is possible to find the number (if more than one) of significant communities related to a topic (combination of community detection and topic analysis). This refers to community detection and profiling in WP2. This scenario may be further extended by e.g. searching for popular and/or controversial material posted by members of these communities, thus relating it to Task 4.2.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A18</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP3</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description**: I can identify the people (sources (accounts/webpages)) who posted the original posts (audio/image/video) as opposed to just “retweets”.

**Comments**: WP3 forensics.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A19</td>
<td>Revealing the Contributor</td>
<td>Community detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/5/6</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description**: I can create a list of trusted users for a specific story/topic (e.g. Malaysia Airline flight MH370).

**Comments**: WP6 UI. WP2 will support trusted lists in social analysis. WP5 will support trusted lists in trust models.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>A20</td>
<td>Revealing the Contributor</td>
<td>Contributor detector</td>
<td>News &amp; Enterprise</td>
<td>WP2/5/6</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description**: I can manually define who is trustworthy/influential (e.g. after looking up their tweets or uploaded videos).

**Comments**: WP6 UI. WP2 will support trusted lists in social analysis. WP5 will support trusted lists in trust models.
### Requirement 1: Reveal the Contributor

**Scenario(s) name:** Revealing the Contributor  
**Role(s):** Contributor detector  
**Use case(s):** News  
**WP(s):** WP2  
**Priority:** 1

**Description:** I can ask my trusted network for help / be able to get in touch with my trusted network (e.g., “For verification – have you seen it? Can you go and check it?”).

### Requirement 2: Newbies

**Scenario(s) name:** Newbies  
**Role(s):** CRM / Community Manager  
**Use case(s):** Enterprise  
**WP(s):** WP2, WP4  
**Priority:** 1

**Description:** I have access to an integrated tool that allows me to identify easily the profile of new community user when I don’t know anything about him/her.

**Comments:** Cross-community interlinking and data fusion.

### Requirement 3: Newbies

**Scenario(s) name:** Newbies  
**Role(s):** CRM / Community Manager  
**Use case(s):** Enterprise  
**WP(s):** WP2  
**Priority:** 1

**Description:** I can easily see if a user has access not only to TECHcommunity but is also active on other social media platforms.

**Comments:** Social Media Platforms relevant for the enterprise scenario: TECHcommunity & ARIScommunity, Twitter, LinkedIn, Facebook, YouTube, Blogs, SlideShare, Flickr, Instagram

### Requirement 4: Newbies

**Scenario(s) name:** Newbies  
**Role(s):** CRM / Community Manager  
**Use case(s):** Enterprise  
**WP(s):** WP2  
**Priority:** 1

**Description:** I can keep track of the history of his/her community entrance (How did he/she find us?).

### Requirement 5: Newbies

**Scenario(s) name:** Newbies  
**Role(s):** CRM / Community Manager  
**Use case(s):** Enterprise  
**WP(s):** WP2  
**Priority:** 1

**Description:** The profile of a community user shall at least contain the fields: profession, company, region, industry, function/role in company, business interests, business skills, connection to social media, connection to and influence in special interest groups, usage of similar products.

**Comments:** The above field will be maintained in PServer (WP2 module). The users might provide some of this information (e.g., profession, company) during sign up in TECHCommunity. It may be possible to reveal e.g., interests given that there exist users who have provided some information, based on user similarities.

### Requirement 6: Customer Relations

**Scenario(s) name:** Customer Relations  
**Role(s):** CRM / Product Manager  
**Use case(s):** Enterprise  
**WP(s):** WP2  
**Priority:** 1

**Description:** If someone is posting a comment in social media to find a solution for a specific product problem I shall be able to find the responsible expert to take care of a specific topic / keyword.

**Comments:** Expert discovery considered as equivalent to influential user discovery. Role analysis in online communities.

### Requirement 7: Customer Relations

**Scenario(s) name:** Customer Relations  
**Role(s):** CRM  
**Use case(s):** Enterprise  
**WP(s):** WP2  
**Priority:** 3

**Description:** If someone is posting a comment in social media to find a solution for a specific product problem I shall be able to find the responsible expert to take care of a specific topic / keyword. To support this I should be able to connect its profile to an internal management system.
## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
A28 | Customer Relation | CRM | Enterprise | WP2 | 1

**Description:** If someone is posting a comment in social media to find a solution for a specific product problem I shall be able to find the responsible expert to take care of a specific topic / keyword. To support this I should be able to address the problem to internal social media community like Yammer.

## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
A29 | Support | CRM | Enterprise | WP2 | 2

**Description:** If someone is looking for an answer, if one of our products is able to fulfill specific requirements I also shall be able to find the responsible expert to take care of this question.

## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
A30 | Support | Technical support | Enterprise | WP2 | 2

**Description:** Someone is looking for technical support through social media for a SAG product. I shall be able to find similar postings and problems to lead to a solution. E.g. the support for a SAG SaaS only product is organized through communities only and there is a dedicated twitter channel and followers that might be used. E.g., follow @processlive and the followers trough social media channels.

## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
A31 | Support | Technical support | Enterprise | WP2 | 1

**Description:** Someone detected a bug in a product and posted it in the TECHcommunity. I shall be able to find out a probably existing contact to support. If not, people from support contact the contributor and forward it to the development where a new patch is provided in the end.

## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
A32 | Analysts | Innovation Manager | Enterprise | WP2 | 1

**Description:** I shall identify people of very high interest, e.g. business analysts of Gartner to keep contact with them or ideally lead them to the social media channel.

**Comments:** Those are people with an extremely high influence and trustworthiness using the same keywords SAG is also focusing on. They have their own blogs and online publications as well.

## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
A33 | Analysts | Innovation Manager | Enterprise | WP2 | 1

**Description:** I shall identify people of very high interest, e.g. business analysts of Gartner to keep finding new trends and assign a relevant profile to them (topic, region, analyst group, contacts to contributors).

**Comments:** Topic modelling in combination of authors and time. It might be possible to characterize the topics related to a user and try to extract and characterize the community they are a part of in terms of topic.

## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
A34 | Analysts | Innovation Manager | Enterprise | WP2/6 | 1

**Description:** I can define a list of trusted users for a specific topic and I can make some personal changes to their tool generated profiles.
### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
--- | --- | --- | --- | ---
A35 | Negative or positive discussions | Marketing | Enterprise | WP2 | 1

**Description:** I shall be able to find negative or positive discussions in all known social media groups belonging to SAG on behalf of the tool. I need to know as much as possible about the person and his background and find possible duplicate discussions to be able to intervene and moderate the situation. Has he influence, is he trustworthy, is he a customer, competitor? E.g., search in LinkedIn SAG groups.

### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
--- | --- | --- | --- | ---
A36 | Event: Innovation World | Marketing | Enterprise | WP2 | 1

**Description:** An important event will happen and I like to address the most important people. The tool will support me to find relevant people.

**Comments:** "Relevant" for the enterprise scenario does include: is the person active in public forums/blogs, is he/she often invited as a speaker in conferences, the position in the company, politics, science? Topic modelling in combination of authors and time.

### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
--- | --- | --- | --- | ---
A37 | Event: Innovation World | Marketing | Enterprise | WP2 | 2

**Description:** An important event will happen and I like to address the most important people. The tool will support me to find reactions of the announcement in social media.

**Comments:** Search for reactions on invitations/announcements. Popular and/or controversial discussions made by relevant users may be identified. This highly depends on the nature/availability of relevant data.

### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
--- | --- | --- | --- | ---
A38 | Event: Innovation World | Marketing | Enterprise | WP2 | 1

**Description:** An important event will happen and I like to address the most important people. The tool will support me to motivate people to join the event.

### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
--- | --- | --- | --- | ---
A39 | Event: Innovation World | Marketing | Enterprise | WP2 | 1

**Description:** An important event will happen and I like to address the most important people. The tool will support me to motivate people to give feedback.

### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
--- | --- | --- | --- | ---
A40 | Innovation gathering | Innovation Manager | Enterprise | WP2 | 1

**Description:** I need to be up to date and I am gathering for innovation in a specific business area. The tool shall lead me to thought leaders. They have a specific profile, or may have a specific score in the TECHcommunity.

### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
--- | --- | --- | --- | ---
A41 | Innovation gathering | Innovation Manager | Enterprise | WP2 | 1

**Description:** I need to be up to date and I am gathering for innovation. The tool shall help me to find a new trend, maybe make use of crowd sourcing. The country should be identified. Either many people OR interesting people are talking about a topic (e.g., search for #IBO).

**Comments:** Topic modelling in combination of authors and time.
### 3.3.2 Requirements for revealing Multimedia Content

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>Revealing Multimedia Content: Picture/video</td>
<td>Content detector</td>
<td>News</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can search for most relevant photos and videos based on keywords AND/OR based on trusted source (e.g. #MH370).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> The indexing framework in WP3 will support query by keyword and source. The trusted source will have been identified in WP2 (as input parameter for this requirement). Initial results will be provided by Task 3.1 and improved by Task 3.2. Similar to the requirements #A11 and #A12.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>Revealing Multimedia Content: Picture/video</td>
<td>Content detector</td>
<td>News</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can see a summary of the most important metadata related to a video or photo (e.g. exif data such as date and location).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> It is important to note that most of the social networks (e.g. Facebook, Flickr) strip out exif metadata from images. However, if the metadata is available, it is a trivial task to provide them to the end-user. Regarding metadata related to a video, please see the relevant comment in the requirement #B8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>B3</td>
<td>Revealing Multimedia Content: Picture/video</td>
<td>Content detector</td>
<td>News</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can search for similar images/videos (e.g. I see images on a timeline of publication – which of the images have been published first).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> As a benchmark we could use TinEye and Google Images. Similarity-based video search is less effective, since it can only rely on videos whose available key frames are similar.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>B4</td>
<td>Revealing Multimedia Content: Picture</td>
<td>Image detector</td>
<td>News</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I see more pictures from the same photographer (e.g. could be the same user).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> The issue here is where we can get the photographer name or id. If the name is stored in the exif data, this is a trivial task, but the majority of users don’t store their names and lots of social networks strip out exif metadata anyway. For cases where the photographer name has to be extracted from accompanying text or inferred in any way, this information has to be provided by another module (e.g., from WP2).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>B5</td>
<td>Revealing Multimedia Content: Picture</td>
<td>Image detector</td>
<td>News</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can search for provenance of image.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> Assuming that the original image has been collected by the system.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>B6</td>
<td>Revealing Multimedia Content</td>
<td>Content detector</td>
<td>News</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I receive a list of results of similar content items, with links to source content and authors, and reports on any detected image manipulation.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> Image manipulation detection and clustering in WP3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>Scenario(s) name</td>
<td>Role(s)</td>
<td>Use case(s)</td>
<td>WP(s)</td>
<td>Priority</td>
</tr>
<tr>
<td>---</td>
<td>-----------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>B7</td>
<td>Revealing Multimedia Content: Picture</td>
<td>Image detector</td>
<td>News</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I see the changes that have been made to a picture (e.g. cropping, blurring, splicing, Tone adjustments, contrast adjustment, pasting, patching, etc.).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comments:</strong></td>
<td>The type of manipulations that we can detect must be pre-defined. Detecting tone adjustments for instance, is probably not very helpful. We should concentrate on manipulations that fundamentally change the underlying image structure like cropping and splicing.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B8</td>
<td>Revealing Multimedia Content: Video</td>
<td>Video detector</td>
<td>News</td>
<td>WP3</td>
<td>2</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I receive a list of metadata for videos (e.g. comparable to exif data of images). I need to receive a list of as many metadata related to videos</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comments:</strong></td>
<td>There is not exactly an exif equivalent for video files. When shooting a video with a camera, some metadata is stored in the .thm file, which is a small jpeg, a thumbnail of the first frame of the video accompanying the .mov file. Acquiring this file in case of online multimedia is probably not possible, as YouTube for example re-encodes the uploaded videos. There are generally some metadata stored in some form for all video formats but this diversity in combination with the fact that very few videos in their original form can be found online prohibit us from fulfilling this requirement for the vast majority of online videos.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B9</td>
<td>Revealing Multimedia Content: Picture</td>
<td>Image detector</td>
<td>News &amp; Enterprise</td>
<td>WP3</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I can also use existing image search engines to verify a photo. I have access to external tools to help me if needed.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comments:</strong></td>
<td>Most of the available services are not free for the amount of queries that the REVEAL system would demand. We will consider buying subscriptions for one or several services later in the project.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B10</td>
<td>Revealing Multimedia Content: Text Support</td>
<td>Text detector Requester</td>
<td>News &amp; Enterprise</td>
<td>WP3/4</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I can search for most relevant text based on keywords AND/OR based on trusted user.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comments:</strong></td>
<td>Also based on Instagram comments or tweets or posts on Facebook (entities per platform)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B11</td>
<td>Revealing Multimedia Content: Text Support</td>
<td>Text detector Requester</td>
<td>News &amp; Enterprise</td>
<td>WP3/4/5/6</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I can see all the relevant metadata related to a text.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comments:</strong></td>
<td>WP3/4 will keep metadata. WP5/6 will display metadata when we have it.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority
---|---|---|---|---
B12  
Revealing Multimedia Content: Text Support  
Text detector  
Requester  
News & Enterprise | | WP3 | 1 |

**Description:** I can search for similar text.

**Comments:** Clustering and relation discovery will aid towards search for similar text.

| # | Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority |
---|---|---|---|---|---|
B13  
Newsgathering support  
Content detector  
News  
WP2/3/4/6 | | 1 |

**Comments:** We need to define the set of target concepts. WP6 UI allows configuration options to be specified. WP2/3/4 support configuration options.

| # | Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority |
---|---|---|---|---|---|
B14  
Newsgathering support  
Content detector  
News  
WP2/3/4 | | 1 |

**Description:** I can find the most relevant text based on my search terms from multiple social media sources.

**Comments:** I can choose whether I want the most recent or the most relevant text based from multiple social media sources. WP2/3/4 identifies relevant items.

| # | Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority |
---|---|---|---|---|---|
B15  
Newsgathering support  
Content detector  
News  
WP3 | | 1 |

**Description:** I can find the most relevant photo published based on my search terms from multiple social media sources.

**Comments:** Relevance can be interpreted in many ways. The multimedia indexing framework in WP3 will provide a flexible search facility for retrieving indexed content in different ways. Initial implementation will be provided by Task 3.1 and later improved by Task 3.2.

| # | Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority |
---|---|---|---|---|---|
B16  
Newsgathering support  
Content detector  
News  
WP3 | | 1 |

**Description:** I can find the most relevant video published based on my search terms from multiple social media sources.

**Comments:** The retrieval will be based mostly on the textual metadata of videos. Initial results will be provided by Task 3.1 and improved by Task 3.2.
### 3.3.3 Requirements for revealing Context

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Revealing location: Mapping workbench</td>
<td>Location detector</td>
<td>News</td>
<td>WP4/5/6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can define a spatial location on a map that displays social media topic of interest.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> Related to Geospatial Context extraction, annotation of content items with a geolocation. WP6 UI for configuration. WP5 HTTP endpoint called by WP6. WP4/5 for map data and visualization.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>Revealing location: Mapping workbench</td>
<td>Location detector</td>
<td>News</td>
<td>WP4/5/6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can filter the map over a location, including the social networks I want to display &lt;YouTube, Flickr, Twitter, Foursquare, Instagram&gt;, the time interval of content and the search term.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> Related to Geospatial Context extraction, annotation of content items with a geolocation. WP6 UI for configuration. WP5 HTTP endpoint called by WP6. WP4/5 for map data and visualization.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>Revealing location: Mapping workbench</td>
<td>Location detector</td>
<td>News</td>
<td>WP4 &amp; WP5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can click on the mapped content, which will automatically lead me to the original tweet, Flickr photo, YouTube video.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> Tracking provenance information in the content items. WP4/5 for map data and visualization. Source URI's will be provided for content items displayed.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>Revealing location: Mapping workbench</td>
<td>Location detector</td>
<td>News</td>
<td>WP3 &amp; WP4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can manually upload content (image, video, text) and search for the original location.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>Revealing location: Trending workbench</td>
<td>Trend detector</td>
<td>News</td>
<td>WP4 &amp; WP5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can receive an overview of trending topics on a map.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong> Related to Geospatial Context extraction, annotation of content items with a geolocation. Trending topics seems related to Topical Context extraction and time based analysis. WP4/5 map display showing admin regions (e.g., city) and the ranked by frequency hashtag/URL/topic in each.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td>Revealing location: Trending workbench</td>
<td>Trend detector</td>
<td>News</td>
<td>WP4 &amp; WP5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td><strong>Description:</strong> I can set filters to a map, including the social networks I want to display &lt;YouTube, Flickr, Twitter, Foursquare, Instagram&gt;, the time interval of content and the search term.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>Scenario(s) name</td>
<td>Role(s)</td>
<td>Use case(s)</td>
<td>WP(s)</td>
<td>Priority</td>
</tr>
<tr>
<td>---</td>
<td>-----------------</td>
<td>---------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>C7</td>
<td>Revealing location: Trending workbench</td>
<td>Trend detector</td>
<td>News</td>
<td>WP4 &amp; WP5</td>
<td>3</td>
</tr>
</tbody>
</table>

**Description:** I can click on the mapped content, which will automatically lead me to the original tweet, Flickr photo, YouTube video.

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>C8</td>
<td>Revealing location: Trending workbench</td>
<td>Trend detector</td>
<td>News</td>
<td>WP4/5/6</td>
<td>1</td>
</tr>
</tbody>
</table>

**Description:** I filter in/out specific users (e.g. content from verified users with tweets).

**Comments:** WP4/5 for map data and visualisation filter options will be provided (for map display) OR keywords & configuration can be changed via WP6 UI and effects will ripple through to real-time map once next update is received from WP4.
### 3.3.4 General requirements

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>All</td>
<td>REVEAL end user</td>
<td>News &amp; Enterprise</td>
<td>WP5 &amp; WP6</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I can access an easy-to-use, efficient and fast REVEAL user interface from which I can configure the social media networks and users and media items I am interested in. <strong>Comments:</strong> WP5 (situation assessment web UI). WP6 (pilot UI embedding WP5 web pages).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2</td>
<td>All</td>
<td>REVEAL end user</td>
<td>News &amp; Enterprise</td>
<td>WP6</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I want to be able to register/log-in to REVEAL and to have easy access to all REVEAL functionalities. <strong>Comments:</strong> Regarding the integrated platform, end users should be able to combine all user stories.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3</td>
<td>All</td>
<td>REVEAL end user</td>
<td>Enterprise</td>
<td>WP5 &amp; WP6</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I want to be able to individualise my REVEAL platform interface (personal source settings, widgets, etc.). <strong>Comments:</strong> Via web interfaces (end user UIs).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>D4</td>
<td>All</td>
<td>REVEAL end user</td>
<td>News &amp; Enterprise</td>
<td>WP5 &amp; WP6</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I can work through all mini-scenarios via one interface. <strong>Comments:</strong> WP6 (pilot UI showing results &amp; embedding WP5 web pages).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>D5</td>
<td>All</td>
<td>REVEAL end user</td>
<td>News &amp; Enterprise</td>
<td>WP6</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I can access the tool from multiple devices: desktop, mobile, tablet. <strong>Comments:</strong> The option of using HTML5 could be a generic solution for supporting both web-based and mobile interfaces.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>D6</td>
<td>• Newsgathering support</td>
<td>• Content detector</td>
<td>Customer</td>
<td>News &amp; Enterprise</td>
<td>WP2/3/4/5/6</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I can find content (in near-real-time) from multiple social media sources based on search terms (e.g., #MH370). <strong>Comments:</strong> We can deal with the provision of wrappers for social sources, as well as with the provision of retrieval mechanisms for content from the local platform given an input query. WP2/3/4 (web crawlers). WP4/5 will provide an HTTP endpoint to allow WP6 UI to start a new assessment (including search keywords).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Scenario(s) name</th>
<th>Role(s)</th>
<th>Use case(s)</th>
<th>WP(s)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>D7</td>
<td>Newsgathering support</td>
<td>Content detector</td>
<td>News</td>
<td>WP2/3/4/5/6</td>
<td>1</td>
</tr>
<tr>
<td><strong>Description:</strong></td>
<td>I stay up-to-date through this breaking news scenario based on pre-defined search terms (e.g., #MH370). <strong>Comments:</strong> WP2/3/4 incremental &amp; real-time processing. WP5 situation assessments updated in real-time (displayed via web page). WP6 supports for off the shelf tools like TweetDeck.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### D8
- **Scenario(s) name**: Newsgathering support
- **Role(s)**: Content detector, Customer
- **Use case(s)**: News & Enterprise
- **WP(s)**: WP2/3/4/5/6
- **Priority**: 1

**Description**: I can personally select social media platforms as sources for finding content (e.g., I only want to search on Twitter AND/OR Weibo AND/OR YouTube AND/OR Flickr).

**Comments**: WP2/3/4 (web crawlers). WP4/5 will provide an HTTP endpoint to allow WP6 UI to start a new assessment (including search keywords). WP5 web page will provide some sort of filtered layers. WP6 UI to provide easy way to configure this.

### D9
- **Scenario(s) name**: Newsgathering support
- **Role(s)**: Content detector, Customer
- **Use case(s)**: News & Enterprise
- **WP(s)**: WP5 & WP6
- **Priority**: 1

**Description**: I can select various search parameters for finding content. This is also important because it helps me understand how the REVEAL process is taking place (how the algorithms work). For instance, I just use a hashtag AND/OR select a social network platform AND/OR specify the media item (just video AND/OR audio etc.) and users (e.g. only verified users).

**Comments**: Search operators will be provided by the multimedia search framework of WP3. WP6 UI will provide easy way to configure this. WP5 endpoint will allow WP6 to pass on parameters to assessment criteria.

### D10
- **Scenario(s) name**: Newsgathering support
- **Role(s)**: Content detector
- **Use case(s)**: News
- **WP(s)**: WP3/5/6
- **Priority**: 2

**Description**: I can select various search parameters for finding content and I will receive a list of related topics based on my search terms. E.g. I type in #MH370 and I receive “do you also want to crawl #malaysia, #aviation” etc.

**Comments**: WP6 UI will provide easy way to configure this. WP5 endpoint will allow WP6 to pass on parameters to assessment criteria.

### D11
- **Scenario(s) name**: Newsgathering support
- **Role(s)**: Expert detector / list creator, Thought leader (score), expert, competitor
- **Use case(s)**: News & Enterprise
- **WP(s)**: WP2
- **Priority**: 1

**Description**: I can receive a system-generated collection of users derived from the networks of users I have defined as trustworthy (e.g. plane crash will result in list of trusted experts (such as pilots etc.).

**Comments**: Journalists trust news organizations such as AP, Reuters, BBC. Users these trusted sources follow and retweet are regarded as more trustworthy. Regarding the rest sources, we can work on this within the context WP2. This can NOT be a real-time service, as it will need time to crawl a non-trivial neighbourhood of users around the pre-defined trustworthy set and also to perform the calculations.

### D12
- **Scenario(s) name**: Newsgathering support
- **Role(s)**: Expert detector / list creator, Thought leader (score), expert, competitor
- **Use case(s)**: News & Enterprise
- **WP(s)**: WP2/5/6
- **Priority**: 1

**Description**: I can manually define a network of trusted expertise based on search terms by simply adding them to the list of trusted users. E.g. add people I know are experts on aviation, psychologists, technicians, pilots, AND/OR add a Twitter list(s) AND/OR LinkedIn contacts.

**Comments**: WP6 UI to provide an easy way to define trusted network. WP2 allow trusted lists of people WP5 allow trusted lists of people for trust analysis.
### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority |
|------------------|---------|-------------|-------|----------|
| **D13** | - Newsgathering support  
- Innovation gathering | - Expert detector / list creator  
- Thought leader (score), expert, competitor | News & Enterprise | WP2 & WP5 | 1 |

**Description:** I can follow my trusted network / what they tweet or what images/videos they upload (e.g. based either automatically added experts or manually added experts).

**Comments:** WP2 allows trusted lists of people, WP5 allows trusted lists of people for trust analysis.

### Scenario(s) name | Role(s) | Use case(s) | WP(s) | Priority |
|------------------|---------|-------------|-------|----------|
| **D14** | - Newsgathering support  
- Innovation gathering | - Content detector  
- Thought leader (score), expert, competitor | News & Enterprise | WP2/3/4/6 | 1 |

**Description:** I can define parameters for relevance/influence that will have impact on search – personalisation.

**Comments:** Relevant contributors might not be influential, i.e. they might contribute very relevant content, but might not be well connected with many followers and retweets. E.g., sometimes activity of poster is more important than number of retweets. These are some of the factors we need to take into consideration (e.g. when it comes to Twitter: Number of tweets, followers, lists, activity). WP6 UI allows configuration options to be specified. WP2/3/4 support configuration options.
4 The REVEAL applications

4.1 News application

In order to be more visual and help understand how a news application and respective features could look like, DW drafted a number of sketches. These sketches are supposed to help translate the requirements presented in previous sections into visible and possible layouts for further work.

The first sketch is supposed to display an overview of an integrated Reveal platform. The functionalities can be divided into three different phases: the (1) “Source and Select” phase, (2) a “Compare and Combine” phase, and a (3) “Reveal and Report” phase. In the Source and Select phase, a user profile can be created and the respective user can create a new story. The user can also create a team in which further users can work on the same story. In addition to that, end-users can search for content in various Social Media platforms (see mini-scenario 1 and mini-scenario 4). During the second phase, the content that was found in Social Media can now be checked/verified (see mini-scenario 2 and mini-scenario 3). The content that has been verified should now be available for publication in the third phase, which is the Reveal and Report phase. Via an HTML snippet, the story (with the verified content) should be embeddable and available for other users that can comment on the story or give feedback or can be used by other end-users (i.e. journalists).

The following sketches show the further development of this initial sketch and display that over the past months, further thoughts went into very specific parts of sketching various functionalities. The sketches will be developed further.
4.2 Enterprise application

At present, and based on the interview results, Software AG is not yet able to envision a more concrete enterprise application. However, this will be pursued further as the project progresses. The interviews conducted highlighted that most of the participants did not want to focus on a specific user interface. But they mentioned the following specific desirable determinants:

- Up-to-date data should be explicitly visualized.
- Existing tools like Hootsuite or Lithium could provide useful approaches.
- There should be a user interface view for different professional groups like product managers, sales managers or marketing managers.
- Open interfaces to other tools for linking them together effectively are essential.
- The information preparation should be very simple and interactive which means that all essentials have to be visualized at a glance and little text should be accompanied by many illustrations and photos to assess something quickly.
- There should be an interconnection between REVEAL and the internal SAG tool Yammer, to share and get information and to follow specific people or groups as a combination of both tools.
- A score with a certain number of points should be visible close to the user data. For example: If the user is positioned above a special threshold, the product manager gets an alert such as “Follow this user”!
- Negative or positive feedback could lead to an automated reaction like “Did you mean...?” or “Do you already know...?”. Proposed solutions, personal contact could be initiated in a further step.
5 REVEAL modules

We present here the foreseen REVEAL modules/functionalities\(^{11}\) that will help us reveal the hidden modalities around contributor, content and context. These modules are intended to satisfy the user requirements identified in Section 3.

Table 9: REVEAL modules

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contributor modules</strong></td>
<td></td>
</tr>
<tr>
<td>Community detection</td>
<td>Discovers communities of people that have something in common in two steps: a) builds communities based on links (any kind of virtual connection between people, e.g. retweets) and b) characterises communities based on the most frequently used keywords by the users</td>
</tr>
<tr>
<td>Influence calculation</td>
<td>Calculates the user’s influence on a specific topic based on e.g., number of retweets or mentions, user activity, characteristics of the message’s content</td>
</tr>
<tr>
<td>Information cascades</td>
<td>Discovers through which paths the information (e.g., tweets) is diffusing, what the characteristics of those paths that may explain the diffusion are, and which users have a pivotal role in information diffusion</td>
</tr>
<tr>
<td>User network profile classifier</td>
<td>Derives information about a user, based on the network structure around a user, e.g., associates the user with subjects (a keyword) and possibly other behaviours</td>
</tr>
<tr>
<td>Community extractor profiler</td>
<td>Extracts communities (groups) of online users and profile information per community (such as subjects prevalent in the community, its temporal evolution)</td>
</tr>
<tr>
<td>Role analysis</td>
<td>Assigns roles to users based upon their interaction in a specific community (e.g., Elitist, Grunt, Joining Conversationalist, Popular Initiator, Popular Participant, Supporter, Taciturn, Ignored)</td>
</tr>
<tr>
<td><strong>Content modules</strong></td>
<td></td>
</tr>
<tr>
<td>Semantic segmentation</td>
<td>Cleans HTML pages (e.g. side banners, advertisements and other irrelevant texts) and identifies the interesting sections</td>
</tr>
<tr>
<td>Linguistic analysis</td>
<td>Identifies named entities in text. The entities will be grounded with Wikipedia (if such a page exists). The entities can be persons, places, organizations etc.</td>
</tr>
<tr>
<td>Text indexing</td>
<td>Provides a search interface to texts and retrieves the documents that match a specific term (e.g. Olympic Games)</td>
</tr>
</tbody>
</table>

\(^{11}\) In the future, it is likely that the modules listed here will be updated, and/or new modules introduced.
<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relation discovery</td>
<td>Clusters texts into near-duplicates, or sets of similar documents, based on defined relations</td>
</tr>
<tr>
<td>Stylometry</td>
<td>Discovers the writing style of a contributor. It finds the important features that define a style and identifies if a tweet origins from a journalist for example, rather than a simple user</td>
</tr>
<tr>
<td>Multimedia geolocator</td>
<td>Infers location information from multimedia items and makes use of the visual content of images (metadata)</td>
</tr>
<tr>
<td>Multimedia similarity retriever</td>
<td>Retrieves images (out of the set of indexed ones) that are very similar (near-duplicates) to the input image</td>
</tr>
<tr>
<td>Multimedia clusterer</td>
<td>Extracts clusters (groups) of images that depict the same scene/story</td>
</tr>
<tr>
<td>Multimedia manipulation detector</td>
<td>Detects a number of manipulations/transformations (e.g. cropping, blurring, splicing) between two images</td>
</tr>
<tr>
<td>Multimedia manipulation history builder</td>
<td>Creates a social interaction graph around a given popular online media item</td>
</tr>
<tr>
<td>Multimedia concept detector</td>
<td>Assigns concepts/class labels (e.g. Person, Sky, Nature, etc.) to images</td>
</tr>
<tr>
<td>Context modules</td>
<td></td>
</tr>
<tr>
<td>Geospatial context modelling</td>
<td>Provides for a given content item a geospatial context in the form of geo-coordinates as well as spatially related entities in a semantic format</td>
</tr>
<tr>
<td>Social context model</td>
<td>Takes a source (user, account) and models the social context of interactions and behaviour</td>
</tr>
<tr>
<td>Multimedia social context builder</td>
<td>Creates the social interaction graph around a given popular online media item.</td>
</tr>
<tr>
<td>Multimedia social modality extractor</td>
<td>Extracts social modalities (popularity, influence, controversiality) for online media items.</td>
</tr>
<tr>
<td>Topical context model</td>
<td>Provides for a given content item a topical context in the form of categories from a classification taxonomy, as well as topically related entities in a semantic format</td>
</tr>
<tr>
<td>Semantic context model</td>
<td>Provides for various input items a semantic context by interlinking the items with semantic relations. This semantic network can incorporate further entities or concepts which form central points in the obtained network and can help to understand the context</td>
</tr>
<tr>
<td>Social Media client framework</td>
<td>Crawls real-time Social Media (Twitter, YouTube, 4Square, Flickr, Instagram, etc.) for a specified reporting period (e.g. 12 days)</td>
</tr>
<tr>
<td>Geospatial context extraction pre-</td>
<td>Downloads OpenStreetMap and Gazetteer information for each spatial context in the form of geo-coordinates as well as spatially related entities in a semantic format</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>processing</td>
<td>region of interest</td>
</tr>
<tr>
<td>Geospatial context</td>
<td>Geoparses all textual content (e.g. Twitter, YouTube titles)</td>
</tr>
<tr>
<td>extraction processing</td>
<td></td>
</tr>
</tbody>
</table>
6 Conclusions

This deliverable presented the process followed by the REVEAL project to capture, analyze and record user requirements for the development of the two use cases: news and enterprise. The tools and processes used to achieve these objectives have been described in detail, and the major conclusions have been presented.

From the methodology followed different mini scenarios have been derived, which were tightly focused on specific end user challenges and that can be seen as typical journalistic or enterprise workflows. These mini scenarios were further analysed in specific requirements or user stories and categorised based on what modalities they are trying to reveal. We also mapped these requirements with the WP responsible for their implementation and have given a priority based for each of them.

From an end user perspective, defining mini-scenarios allows the REVEAL project to create a variety of small demonstrators (e.g. in the form of tools, platform(s) or components) each focused on a specific end user domain challenge.

From a technology partner perspective, defining mini-scenarios reduces the complexity, and therefore the risk of failure/delay of integration work. This allows the effective use of an agile software methodology to integrate and test combinations of partner components against test cases based on each of the scenario’s narrative.

Although this document describes a baseline of the user requirements, it is expected that subsequent refinements will be applied throughout the design and implementation phases. During the internal and pilot testing, users will evaluate the tools developed for each use case and report any feedback to the implementation team.

This deliverable is of vital importance, as subsequent development work will largely be based on the results of the requirements analyses conducted in this deliverable. Moreover, the outcome of this deliverable will be used by the technical partners to make their final system specifications and architecture design.
7 References


8 APPENDIX A: Competitive tools and opportunities to fill the market gaps

In the table below we show what kind of verification is performed by a selection of different competitor applications. As stated, this is only a snapshot of related tools and services that are being monitored and analysed by the REVEAL consortium. For an analysis and portrayal of additional tools, see especially the REVEAL website on http://revealproject.eu/ and especially our regular column “Verify This Week”.

Table 10: Competitors’ applications

<table>
<thead>
<tr>
<th>Tool with verification capabilities</th>
<th>Kind of verification performed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sulia</strong></td>
<td>It uses a combination of crowdsourcing, human editorial insight, and machine learning. Sulia creates topic channels, somewhat akin to news verticals, whose content is curated from experts in particular subjects. It builds a web-of-trust between news sources and connects trusted sources and enthusiasts on shared interests across thousands of subjects, including breaking news and events.</td>
</tr>
<tr>
<td><strong>Reporter Kit</strong></td>
<td>It contains a trust index that permits to see how many submissions a source has made, how many were not published, and even for what reasons. It also logs geolocation history for every member and every submission to quickly compare multiple pinpoints on a Google map, and easily authenticate the provenance of the files. It allows looking at the Meta Data stored in the files to corroborate the member’s description with the actual facts of the files. For example, when were the images captured, with what device, and if they were edited in Photoshop before being submitted.</td>
</tr>
</tbody>
</table>
| **Storyful**                       | It combines automatic and human-intervened content verification. Storyful journalists use a combination of technology and journalism to find the most authentic sources on the key stories and topics. A team of reporters monitors and engages with communities on Twitter, YouTube, Google+ and any social platform with the power to change a broader conversation. When they discover compelling content – video, photos, maps or tweets - Storyful journalists search for the key details, data and context that will help clients safely use the content in their broadcasts or on their websites. Storyful delivers a range of hand-curated and automated content feeds through a web dashboard or directly into a client’s content management system (CMS). In addition, it has developed a verification process for high-value video content. This process includes some of the following workflows:  
  1. Engagement with the member of the community who shared the video in an effort to establish the identity of the original uploader.  
  2. Translation of every word that comes with a video for additional context.  
  3. Review of the uploader’s history to see whether he/she has shared useful and credible content in the past, or if he/she is simply ‘scraping’ other people’s videos.  
  4. Use of Google street view/maps/satellite imagery to help verify the locations in a video.  
  5. Consultation with other news sources to confirm events in a video happened as they were described.  
  6. Examination of key features in a video such as weather and background landscape to see if they match known facts on the ground. |
| **Flumes**                         | It tries to reveal trustworthiness providing a verification process, using a trust |
scoring algorithm to verify content and sources. Users can decide if a source is trustworthy by checking locations, similar images and influence scores. It also performs automatic curation, aiming to order the content by recency and popularity, and produces a list of key influencers specific to a topic of interest. Furthermore, it performs sentiment analysis, showing mood trends over time on any subject and assigning a strength of sentiment to a message.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ushahidi</td>
<td>It relies on crowdsourcing for content verification, based on a verification guide addressed to the users of the application. Ushahidi support for verification has until now been limited to a fairly simple backend categorization system by which administrators tag reports as &quot;verified&quot; or &quot;unverified&quot;. Another tool developed by the Ushahidi team is &quot;SwiftRiver&quot;, which has content curation, verification and filtering capabilities. Using citizen editors, it provides them with tools to identify relevant sources and rate them, identify and rate authoritative users, and curate discussions and media around any topic or event.</td>
</tr>
<tr>
<td>PolitiFact</td>
<td>It is concentrated on political issues and tries to reveal truth based on the editors' manual review. It monitors the factual accuracy of what is said by major U.S. political players in the form of TV ads, debates, speeches, interviews and news releases. Its goal is to apply the best practices of both journalism and scholarship, and to increase public knowledge and understanding.</td>
</tr>
<tr>
<td>FactCheck.org</td>
<td>It tries to reveal truth, but as PolitiFact it is based on human analysis. Moreover, when users ask for a question they get responses from a group of journalists and not in an automated way.</td>
</tr>
<tr>
<td>TinEye</td>
<td>It can find copies of the image on the Internet, or modified versions of an image, and report the date and time at which they were posted. Thus, it can reveal if an image is not original, but is under modification. TinEye does not recognise outlines of objects or perform facial recognition, but recognises the entire image, and some altered versions of that image.</td>
</tr>
<tr>
<td>FotoForensics</td>
<td>It deploys techniques from the field of multimedia forensics and can be used to validate JPG and PNG pictures. It uses Error Level Analysis (ELA) to identify areas within an image that are at different compression levels, and are likely to have been digitally modified. Some important caveats in the use of this tool as mentioned at their own FAQ (<a href="http://fotoforensics.com/faq.php?faq">http://fotoforensics.com/faq.php?faq</a>) are the following: Will this tell me if the picture is fake? No. This site is like a microscope -- it will show you data, but it does not draw any conclusions. FotoForensics includes tutorials to help you understand what to look for in the analysis results. In some cases, the analysis may not provide the answer you wanted. For example, you may want to know if a picture was edited. However, if the picture is a low quality, then the results may not permit identification of anything beyond &quot;low quality, multiple resaves.&quot; As a concrete example, consider analyzing a picture from Facebook. Facebook strips out all original metadata and replaces it with their own metadata. So the metadata analysis will not identify anything beyond &quot;Facebook&quot;. Facebook also resaves the image at a low quality, so the JPEG quality (JPEG %) will report a low quality image. Error Level Analysis will typically return a dark result with large colored rectangles -- indicating a low quality image and multiple resaves (a solid description of what Facebook provides). Even if the picture is visually altered, the algorithmic results may not detect much more than an image resaved by Facebook.</td>
</tr>
</tbody>
</table>

From the above review of verification and validation capabilities, we have identified three main features that are lacking in other tools, and could give our foreseen REVEAL applications a significant advantage:
- fast, near-real-time automatic verification capability
- consideration of verification modalities from all three aspects: contributors, content and context
- provision of the possibility to see the results of individual modalities (e.g. influence, reputation, etc.).
9 APPENDIX B: Questionnaire for the news use case scenario

The questionnaire used for the news case interviews can be accessed via the link: https://www.soscisurvey.de/socialmediaverification/index.php (password protected)

This screenshot shows the landing page.

Hello and thank you for participating in this questionnaire. This should not take longer than 30 min.

We are participating in an EC co-funded Research & Development project called REVEAL. You can find out more about the project at www.revealproject.eu.

The project tackles a number of issues in the area of multimedia content analysis. A core element is the analysis of content (audio, video, text, images) from various sources (especially Social Media sources such as YouTube, Twitter, Flickr and the like) and the verification process that comes with dealing with all these sources as a journalist.

The above-mentioned content analysis aims to:
* extract content items (e.g., images of a particular event)
* provide information about content items (what is it, in which context does it belong etc)
* provide information about the contributor of individual content items (who contributed it, how trustworthy is that person, what is his/her reputation etc)

A primary aim is to aid in answering the question whether particular content items are what they pretend to be. In other words, we want to answer the question: is a particular piece of information that is provided on the web, in particular in Social Networks, trustworthy?

Ideally, the outcome of the REVEAL project will be a platform that aids in the process of verification of content from Social Networks. (We are NOT focusing on content distribution via Social Networks.) As such a tool/service could be of great benefit to the journalistic profession, we would like to talk to you.

Primarily, we want to find out from you:
* how you work (in particular when it comes to information gathering, research and verification processes)
* what obstacles there may be
* what would make your work easier and what you are longing for in this context

All this information is intended:
* for us to know what your requirements are (focus = verification of Social Media content)
* to be fed to the project's developer partners in order for them to develop tools and solutions that meet clear journalistic needs

So thank you very much for you time and willingness to participate.

Revealproject.eu

This questionnaire was developed and subsequently used by DW for the interviews. It was filled out in the course of the interviewees by the person conducting the respective interview. That way, the answers were available for analysis purposes. The questions are a combination of closed questions that allow comparing the answers statistically, as well as open questions in order for interviewees to talk more freely, include their opinions, address unforeseen issues etc.

Below we provide interview questions as well as the options for answering:

1. Do you generally use Social Media platforms, such as Twitter, Facebook or YouTube for your professional use?
   a. Yes

Project Title: REVEAL
Project Coordinator: INTRASOFT International S.A.
Contract No. FP7-610928
www.revealproject.eu
2. How often are you using the following Social Media platforms for professional use? (Never to daily)
   a. Facebook
   b. Twitter
   c. YouTube
   d. Pinterest
   e. Instagram
   f. Flickr
   g. LinkedIn
   h. Tumblr
   i. Path
   j. Foursquare
   k. Blogs
   l. Google+
   m. Other

3. If you use other social media platforms, please list the platforms you are using for professional use (open question).

4. How do you use these Social Media platforms?
   a. For professional use only
   b. For private purposes, as well

5. If you use Social Media platforms both professionally and privately, do you
   a. Mostly use the same account/identity for professional and private use in individual Social Media platforms
   b. Have separate accounts/identities for professional and private use in individual Social Media platforms
   c. Use some Social Networks primarily, and other primarily privately, the way separating private/professional identities
   d. I also use company accounts when using Social Media platforms professionally
   e. Other scenarios, please specify

6. Please tell me in your own words how you use Social Media professionally
   a. As a distribution platform for already existing content
   b. As a means to get in contact with experts/eyewitnesses
   c. As a way to speak to/get in touch with audience
   d. As a source to obtain information in newsgathering process
   e. Please specify

7. When do you professionally turn to Social Media platforms for the following specific purposes? (Never to daily)
   a. Looking for experts in a specific field of interest
   b. Looking for eye-witnesses
   c. Cross-checking information
   d. Looking for pictures
   e. Looking for videos
   f. Looking for audios
   g. Looking for statements
   h. Looking for comments
   i. Asking for comments

(See example screenshot)
6. Please tell me in your own words how you use social media professionally.

After the participant is finished answering the open question, please go through the below checkbox-

- Please Specify

- As a distribution platform for already existing content
- As a means to get in touch with experts/witneses
- As a way to speak/keep in touch with audience
- As a source to obtain information in newsgathering process

7. When do you professionally turn to social media platforms for the following specific purposes?

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Never</th>
<th>Almost never</th>
<th>Once a month</th>
<th>Once a week</th>
<th>Daily</th>
</tr>
</thead>
<tbody>
<tr>
<td>Looking for experts in a specific field of interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looking for eyewitnesses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-checking information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looking for pictures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looking for videos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looking for audios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looking for statements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asking for comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. If you use social media platforms for further purposes, please specify here.

---

8. If you use social media platforms for further purposes, please specify here (open question)

9. How often are you on Social Media platforms for professional use? (Never to daily)
   a. Staying in contact with others via chatting/messaging
   b. Posting/sharing textual information
   c. Finding/reading useful information from others
   d. Posting/sharing pictures
   e. Posting/sharing videos
   f. Distributing content for promotional purposes
   g. Newsgathering/Researching
   h. Other

10. How do you use Social Media once you found information you would like to use?
    a. Integrate own body of work
    b. Double check with other information from other sources
    c. Contact author/contributor
    d. Contact other witnesses
    e. Please specify

11. Are you following specific people/organisations/offers?
    a. No
    b. Yes, other journalists on Social Media networks
    c. Yes, other media organisation on social media networks
12. Please tell me in your own words how you use Social Media as a source to obtain information in the news gathering process
   a. Researching for stories
   b. Researching for events
   c. Researching for experts
   d. Pls specify

13. Please tell me in your own words how you use Social Media as a way to interact with the audience/collaborate with audience
   a. Asking friends/followers/subscribers for stories
   b. Asking for contributions for specific topics/stories
   c. Asking for user generated content
   d. Pls specify

14. Do you specifically verify content that comes from Social Media platforms if it is part the story?
   a. Yes, I always verify any content coming from Social Media platforms
   b. No, I don’t verify content coming from Social Media platforms
   c. Pls specify

15. When it comes to verification of content residing on social media platforms, how important is verification for you and your work in this context?
   a. 5-point scale from little importance to very important, plus option
   b. I don’t verify content

16. Please tell me in your words how you verify content in Social Media.
   a. I use verification tools, such as Storyful or Newswhip
   b. I primarily use content from people I trust (e.g. verified Twitter accounts, official news accounts, famous journalists)
   c. I contact the author of the post/tweet/video/photo
   d. I check if content was used by other (retweets, shares etc.)
   e. Pls specify

17. What kind of Social Media verification tools so you use? (open answers)
   a. For the verification of photos, I use ...
   b. For the verification of videos, I use ...
   c. For the verification of authors/contributors, I use ...
   d. For the verification of textual information, I use ...
   e. Most useful for verification, I find ...

18. If you had one wish concerning the use of Social Media tools for verification processes and the way it would make your life as a journalist easier, what would this be? (open question)

19. Please tell me in your own words which are the major obstacles you encounter at present when you verify content from Social Media. (open question)

20. What major obstacles are you encountering at present when you verify content? (strongly agree to strongly disagree)
   a. Too time-consuming
   b. Too much information
   c. Too little verified information
   d. Difficult to separate marketing from real information
   e. No one platform that suits all needs
   f. Many platforms not geared towards journalistic needs
g. External verification tools not allowed in my organisation

21. Is there anything else, you would like to mention when it comes to verification of content from Social Media platforms? (open question)

22. Are you aware of what other journalists in other organisations use
   a. No
   b. Yes (pls specify)

23. If yes: What do they so that you would also like to do (e.g. workflows, tools etc) (open question)

24. Please tell us which organisation you work for [type and name, if possible and willing to disclose.]
   a. Media Organisation
   b. News Agency
   c. Social Media Platform
   d. Freelancer
   e. Other

25. What is your role within your organisation?
   a. Journalist
   b. Editor
   c. Researcher
   d. Developer
   e. Manager/Co-ordinator
   f. Head of Departments
   g. Editor-in-chief
   h. Graphics
   i. Other

26. What projects do you usually work on?
   a. (Breaking) News
   b. Feature Stories
   c. Technical Work
   d. Project Management
   e. Other

27. If you want, you can tell us your name and email address.

10 APPENDIX C: Questionnaire for the enterprise use case scenario

The following screenshots provide the guideline for the interviews conducted by the SAG research team.
Part 1: Your General User Behaviour

1. Please tell me how you use social media for your daily work.

2. Please describe your activities and the frequency (from “never” to “daily”) in the communities.

3. Which specific people / organizations/ offers do you follow in the media?

4. Do you specifically verify content that comes from social media platforms, and, if so, how?

Part 2: Your Estimation of the Significance

1. In your view, what role do Social Media play for the development and the distribution of our products.

2. Did you make your own experiences (real-life examples)?
Part 3: Your Requirements

1. If you could wish a method/tool that support users in verification processes of information what would be interesting for you

<table>
<thead>
<tr>
<th>Content</th>
<th>is interesting for me, because</th>
<th>is uninteresting for me, because</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detecting of thought leaders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detecting of trends in the BPM area</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detecting of competitors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting verification of users (see also Question 2 in Part 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature Requests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifying of bugs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception of product news (SAG)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception of topics addressed by SAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception of marketing campaigns (SAG)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement of user groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifying of negative discussions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifying of industry-specific content</td>
<td></td>
<td></td>
</tr>
<tr>
<td>others</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Part 4: Handling

Imagine you’re starting up your computer and a mashup with brandnew data of the described modalities is appearing on the screen (f.e. divided by industries).

a. Would you welcome it?

b. How do you think such a mashup should look?

c. Could you imagine other kinds of reporting forms, like dashboards, reports, alerts…?
Part 5: Real-life Scenarios

Imagine a new product version is releasing and there is an overwhelming number of negative contributions.

a. What do you think of a tool revealing backgrounds and data of contributors?

b. Which kind of informations would be helpful?

Part 6: Crowdsourcing through Gamification

Definitions

**Crowdsourcing** was coined by Jeff Howe referring to a business model through which undefined groups of connected individuals contribute innovative impressions in order to present a solution.

**Gamification** Gamification is defined as the "use of game elements in non-gaming context" (Deterding, Dixon, Khaled, & Nacke, 2011), in order to motivate people in a given domain or for a defined problem space.
Part 6: Crowdsourcing and Gamification

1. Are you participating in crowdsourcing and gamification mechanisms in general? If so, please tell us some examples.

2. Do you think it’s useful to reward contributors with reductions or license renewals of SAG products?

2. Could gamification elements provide additional information of higher-level modalities like trustworthiness, completeness, quality, or controversy?
## 11 APPENDIX D: Matrix of Feature Groups of Enterprise Use Case

### Classification of Requirements and Feature Groups

<table>
<thead>
<tr>
<th>Contributor</th>
<th>Content</th>
<th>Technical requirement/Handling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thought leadership</td>
<td>Identifying and following thought leaders (competitors, analysts, leaders (Bill Gates), journalists, colleagues)</td>
<td>Identifying early trends</td>
</tr>
<tr>
<td></td>
<td>Identifying Experts within a company's own ranks: Can TECHCommunity afford to miss such a tremendous important expert and not to link/refer to his contributions? Why he is not yet identified?</td>
<td>Does a posting have an impact for the community? Sustainability is important.</td>
</tr>
<tr>
<td></td>
<td>I know best the few leaders in the world and I'm able to judge them.</td>
<td></td>
</tr>
<tr>
<td>Thought Leadership: a thought leader is very active in social media. So, sometimes you have to ask by yourself: Has he or she nothing better to do?</td>
<td>Which sites succeed?</td>
<td></td>
</tr>
<tr>
<td>Thought leaders and trends are important. But with a factor of 1 to 20 I'm interested in people who are active in a complete different area: Outside-Box-Thinking, lateral thinking. Which method does a sculptor f. e. use to come upon an idea for example? Reputation and influence are most important. But not the popularity</td>
<td>For the development, Social Media are not important. But for the distribution, social media are a cheap and effective way to contact your customers. And a vendor tweet is not annoying.</td>
<td></td>
</tr>
<tr>
<td>Influence is only meaningful if I you regard the overall context, What do 50 likes matter in relation to 100,000 users? Maybe the influence is only significant in a small area.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Influence: A manager has influence, can carry on the message.</td>
<td>Real-Life-Example: We used Social Media to inform about a limited license key for free. It was a great success.</td>
<td>Open interfaces to other tools for linking them together effectively.</td>
</tr>
<tr>
<td>Who contributes in the BPM area and where else is he active?</td>
<td>Search for cause: Had there already been similar discussions relating with this problem?</td>
<td>Should be simple: all essentials at a glance, little text and many illustrations/photos.</td>
</tr>
<tr>
<td>Who is the initiator of a discussion?</td>
<td>A tool should combine both, the push and the pull channels.</td>
<td>There should be an interconnection between the REVEAL-Tool and our internal tool Yammer, because we combine both tools to share and get information and to follow specific people or groups.</td>
</tr>
<tr>
<td>Feature Requests/Bug messages/Comments on the product</td>
<td>If requests reach us, we always have to know where they come from to evaluate, and finally, implement them. We always are interested in the history of a user. So statistics would be helpful here.</td>
<td>The time factor is very important for me. I have to value something quickly.</td>
</tr>
<tr>
<td>Did the requester already get in touch with our support?</td>
<td>For me it's relevant when the content is liked by many people. But it has to be set in relation to the total number of our users. In relation to 100.00 ARIS-User, 30 Likes are not interesting.</td>
<td>My suggestion: a system similar to Klout: I first define which criteria I need to set up a ranking. A score with a certain number of points, which is visible close to the user data, would be significant for me. You could define a threshold value. Is the user positioned above this level than the product manager gets an alert: Follow this user.</td>
</tr>
<tr>
<td>The tool is required to rank contributors according to the number of postings</td>
<td>Negative discussions: You have to distinguish: Is somebody interested in a solution for his problem or does he intend to initiate a discussion?</td>
<td>Suggestion: A negative feedback could lead to an automated reaction: Did you mean...? Do you already know...? Proposed solutions, personal contact will be initiated...</td>
</tr>
<tr>
<td>Is it a real criticism or is</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Project Title: REVEAL  
Contract No. FP7-610928  
Project Coordinator: INTRASOF International S.A.  
www.revealproject.eu
<table>
<thead>
<tr>
<th>Requirements analysis and specifications</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information research</strong></td>
<td></td>
</tr>
<tr>
<td>it a competitor-driven shitstorm?</td>
<td></td>
</tr>
<tr>
<td>Top Posters are required to be sorted in region, social channel and newsgroups</td>
<td>Hashtags: What is being said about specific hashtags, by whom and on what sites?</td>
</tr>
<tr>
<td>Who has been recently joining the community?</td>
<td>A tool should find out the relevant hashtags and relations.</td>
</tr>
<tr>
<td>Trustworthiness stands above everything else. That's why I'm so interested in a history of a contributor. How active does he or she maintain a relationship with us or other users in the ARIS Community?</td>
<td>I don't need to bind deeply with contributors like in the real life. Instead of this I'd like to come together with people who are interested in the same topic like a special interest group in an adult education center.</td>
</tr>
<tr>
<td>Influence is only meaningful if you regard the overall context: What do 30 likes matter in relation to 100.000 users?</td>
<td>The kind of making information available, the information preparation, is very important. Whoever is able to express himself in two sentences makes it significantly easier to convince me.</td>
</tr>
<tr>
<td>Negative discussions: You have to distinguish: Is somebody interested in a solution for his problem or does he intend to initiate a discussion?</td>
<td>It will be a challenge to figure out what is real and what is interpreted? Does the interpretation hit the right tone?</td>
</tr>
<tr>
<td>Negative discussions: Who is the contributor? Is he well connected to cause damage (spread and reach)?</td>
<td></td>
</tr>
<tr>
<td>If I see a new “face” in the network, I try to figure out as much as possible: Where is he working? Is he a consultant? Is he a real customer? So it could lead to a real rapport. For that reason I use</td>
<td>I wouldn't welcome such a tool because I would feel restricted: What you see is what you see.</td>
</tr>
<tr>
<td>Question</td>
<td>Answer</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>How can I identify competitors in our and other communities?</td>
<td>“I’m a creature of habit”. User experience is very important for me, so I’m not interested in a new tool!</td>
</tr>
<tr>
<td>Which products is the user interested in?</td>
<td>I verify check for plausibility more or less intuitively: Does the information meet my expectations or does it match my experiences?</td>
</tr>
<tr>
<td>Is he a technician or a decision maker?</td>
<td>People often tend to post more negative than positive comments. To find out the background could help to assess the posts more objectively.</td>
</tr>
<tr>
<td>There is no need to summarize CVs</td>
<td>Useful parameters for the content/contributor: the balance between negative and positive comments, average assessments in the past, where and which comments, qualitative elements of comments</td>
</tr>
<tr>
<td>Who initiate the most contributions?</td>
<td></td>
</tr>
<tr>
<td>We have to distinguish from industry, country, position/function</td>
<td></td>
</tr>
<tr>
<td>In which other communities spend competitors their time?</td>
<td></td>
</tr>
<tr>
<td>How do our important or potential customers act? Where can I reach them?</td>
<td></td>
</tr>
</tbody>
</table>
Special conditions

- Complex Sales: Social media don’t cause a final purchasing decision, because our products are too complex. (Complex Sales)
- Our customers are still a little bit shy in the context of social media. This regards the decision-makers not the concrete product users.
- Keep in mind: We are moving in the B2B not in the B2C area. We don’t expect masses of negative reactions. In contrast to other companies social media play a minor role for us as B2B enterprise.
- Global player: Revealing backgrounds and data of contributors? I have worked for over eight years for SAG and have observed a „German bias“. If somebody posts a tweet, I assume that he is an expert in his area. Otherwise he wouldn’t post a tweet. I think there is a German suspicion. (American point of view)

Just a site note

Who actually is teaching social competence in social media? Are we all able to move in social media? Or do we need a kind of contributor-license?

Gamification / Crowdsourcing

<table>
<thead>
<tr>
<th>Contributor</th>
<th>Technical requirements/Handling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamification is a good way to get information about users you don’t know, but we do not have manpower for manual work</td>
<td>Can gamification elements be automated? Or can more information be extracted out of the existing content? E-Learnings also can be a gamification element.</td>
</tr>
<tr>
<td>Sometimes used for getting reference statements of customers</td>
<td>Restricted by ethical issues, code of conduct, compliance rules</td>
</tr>
<tr>
<td>Access to preview releases could be a way to motivate people to bring up more information</td>
<td>Needs to be included in the forums and monitoring tools Analyzing Q/A behavior</td>
</tr>
<tr>
<td>Decision makers do not have enough time to use it, more applicable for end users</td>
<td>Need to know the profile</td>
</tr>
<tr>
<td>Forum writers could like to have some kind of award / rankings Helping others is a kind of gamification</td>
<td></td>
</tr>
<tr>
<td>Crowdsourcing is not used for professional purpose</td>
<td></td>
</tr>
<tr>
<td>ARIS Connect is used to find out trends in the groups and which new features are interesting so könnt the get proposals for improvements</td>
<td></td>
</tr>
</tbody>
</table>