Global Under-Resourced MEedia Translation (GoURMET)

H2020 Research and Innovation Action
Number: 825299
D5.2 – Use Cases and Requirements

Nature	Report	**Work Package**	WP5
Due Date	31/06/2019	**Submission Date**	27/06/2019
Main authors	Andrew Secker (BBC), Julie Wall (BBC), Peggy van der Kreeft (DW), Susie Coleman (BBC)		
Co-authors			
Reviewers	Wilker Aziz (UVA), Mikel Forcada (ALAC, subreviewer)		
Keywords	Use Cases, Requirements, Personas		
Version Control			
v0.1	**Status**	Draft	17/06/2019
v1.0	**Status**	Final	27/06/2019

Contents

1 Introduction

This document describes the use case definitions and user requirements for the technologies and prototypes developed by the media partners.

GoURMET envisages three different use cases: global content creation and media monitoring are the primary use cases. In the third year, a third use case on financial news will be added. This present deliverable deals primarily with the first two primary cases.

The use cases as described in this deliverable will form the basis for the development of prototypes and demonstrators at the two participating media partners, i.e., Deutsche Welle and BBC. These prototypes and demonstrators will be used to test or otherwise demonstrate the efficacy of the developed machine translation (MT) technologies in modern news organisation working in multilingual environments.

The project will develop two generic interfaces. One browser-based user interface (UI) which can be used to quickly demonstrate the underlying translation technologies, and an API which will be used to allow machine access to the translation technologies. Each media partner will then develop their own newsroom prototypes (drawn from the use cases described in section 3), or extend existing newsroom prototypes to exploit the translation technologies developed by research partners.

The list of use cases described in section 3 represents the culmination of a comprehensive requirements gathering exercise and is presented in their entirety in order to give an indication of the variety of problems that can be solved with MT technology in media organisations. A subset of these will be developed and prototyped during the remainder of the project.

This deliverable is the first in a series of three that cover the integration and prototyping process. D5.3, Initial progress report on integration is due M18 and D5.5, Final report on integration is due M36.

This document is in three main parts. Firstly a set of personas are detailed, all from the media partners' perspectives, in order to describe the primary actors that are involved in the two overarching use cases considered in this document. Secondly, use cases are described. These are separated out between the two media partners as each has their own internal priorities. Thirdly, the use cases are used to derive a set of functional and non-functional requirements.

2 Persona Descriptions

This section presents the personas applicable to GoURMET prototypes with regard to the the two primary use cases: *media monitoring* and *global content creation*. For each of these two use cases, this section details the relevant personas, with a professional profile describing the role, the key skills, and key responsibilities, challenges, their interest in the project goals, etc.

All personas listed herein have a connection to the Use Cases described in section 3 and would be expected to be a user of least one of the tools outlined in that section. The two media partners - DW and BBC - have a broadly similar structure for their journalistic teams. As such, and in order to save unnecessary repetition in this section, each persona represents a merging of the equivalent roles between the two organisations. That is to say, the inclusion of this section is aimed at allowing the reader to understand the prototypical users that would have interest in using the MT technologies produced by this project within the context of large news organisations such as DW and BBC.

2.1 Content Creation

2.1.1 Journalist

Profile: Language Department

Organisation: Language Department

Industry history/length of service: 3–10 years

Languages: Native speaker of the language covered

Area covered for production: One particular language area such as English, German, Spanish, Arabic, Portuguese for Africa or for Latin America, and so on.

Key skills: Interest in people, news, current affairs and a good knowledge of the area covered. Have a good eye for a story. Be creative, be able to generate original ideas and find good and reliable source material. Power to convince editors of value of suggested story. Excellent writing production skills. Close cooperation: Fellow journalists and editors working in that language

Reports to: Editor

Key Statement: "I need to be aware of newsworthy events which are relevant **for** my language area but not happening **in** my language area. I need to find original ideas and produce good stories."

Key Responsibilities: Produce high-quality (online text or audio-visual) items in my native language, meeting the set deadline. Find good and reliable source material for journalistic production,

Interest/project goals: 1) Get easy and direct access to external (agency and social media content), as well as items published by my own organisation that may be of interest to my work, in a language that I understand (original or translated into English). 2) Help me translate content.

Current frustrations: Reuse of items produced by my organisation is hindered by lack of direct access to such items, time constraints and language barriers (it needs to be translated even for ingestion).

2.1.2 Language Team Editor

Profile: Language Team Editor

Organisation: Language Department

Industry history/length of service: 10–15 years

Languages: Native speaker of the language covered

Area covered for production: One particular language area such as English, German, Spanish, Arabic, Portuguese for Africa or for Latin America, and so on.

Key skills: Experience in broadcasting editing, including text and audio-visual content. Interest in people, news, current affairs and a good knowledge of the regions in which that language is spoken in terms of politics, economy and society as a whole. Ability to spot and assess good stories related to the language area. Good management, coordination and editing skills.

Reports to: Head of Region

Key statement: "I need to be aware of newsworthy events in my language area and ensure coverage of these events by my language department."

Key responsibilities: Ensure all major events related to my language area are covered and plan, manage and coordinate journalistic production in my team.

Interest/project goals: 1) Get a quick and reliable overview of news events of interest to my language area. 2) Enhance efficiency of journalistic production through improved access to other published material that could potentially be reused (content topics or content items).

Current frustrations: Access to potentially interesting items from other languages is not easy and becomes slow as it requires a request for human translation.

2.1.3 Head of Region

Profile: Head of Region

Organisation: Regional Department

Industry History/Length of Service: 15–20 years

Languages: Basic understanding of all languages covered by the assigned region, for example Europe, Near/Middle East, Africa, East Asia, Latin America, etc.

Key Skills: Experience in broadcasting editing, including text and AV content. Interest in people, news, current affairs and a good knowledge of the region in terms of politics, economy and society as a whole. Ability to spot and assess good stories related to the assigned region. Flexibility to manage balanced coverage of entire region, while ensuring reporting on key events.

Reports to: Chief Editor

Key statement: "I need to be aware of newsworthy events in the assigned region and ensure coverage of these by the languages in my team"

Key responsibilities: Plan, manage and coordinate production across all languages for that region. Ensure all major events are covered and avoid duplication of work.

Interest/project goals: 1) Get a quick and reliable overview of news events in my assigned region from external as well as internal sources 2) Keep track of production from all language departments within my region, as well as reuse of content.

Current frustrations: Exchanging information between the language departments within my region and other regions is slowed down hindered by language barriers. I don't have easy accessible editorial oversight to some of the output under my overall management for the same reason.

2.2 Media Monitoring

2.2.1 Journalist

Profile: Media monitoring language team

Organisation: Monitoring

Industry history/length of service: Approximately 5 years

Languages: Specific to Language Team

Data sources covered: Specific to Language Team

Key skills: Understand the relevant language, write well and clearly in English, have the ability to identify important messages from the media sources being monitored

Reports to: Reports to a team manager (in the future will be called an editorial lead)

Systems/tools used: MS Word, A bespoke video monitoring interface, social media tools such as Tweetdeck.

Key statement: "I need to know when events are breaking or developing."

Key responsibilities: Monitor relevant sources and identify breaking events and developments Interest/Project.

Interest/project goals: 1) Stay across and keep up with more developments and more sources than is possible by existing manual monitoring methods. 2) Get alerts for breaking events.

Current Frustrations: Need to constantly monitor lots of sources, not being able to physically keep across every relevant source. Occasionally need to sub-edit material that refers to sources in a language I don't speak.

2.2.2 Researcher

Profile: Media monitoring research team

Organisation: Member of research team, (will also be mixed into relevant language team)

Industry history/length of service: 2–3 years

Languages: Specialise in the language of a country or countries

Sources covered: As needed and available, will also use output from language teams

Key skills: Research skills (checking facts, sources), able to trawl large amounts of information, IT/database skills - put information into editorial research systems

Reports to: Report into team manager, in the future will report to language team editorial lead

Systems/tools used: MonRes, Monarch (in a more minor capacity)

Key statement: "I need to be able to find relevant facts about significant figures and organisations for my coverage country"

Key responsibilities: Trawl relevant data sources as necessary, support editorial output, populate database (MonRes)

Interest/pproject Goals: A good ongoing search ability, using phrases of interest, alerts on key figures/organisation

Current Frustrations: Not always easy to find information needed (dependent on good metadata on media sources), difficult to extract good information from broadcast sources without constantly watching them.

2.2.3 Editorial Lead

Profile: Leads a media monitoring language team

Organisation: Specific to a language team

Industry history/length of service: 5–10 years

Languages: Specific to Language Team

Data sources covered: Specific to Language Team and other external sources such as Reuters

Key skills: Ability to recognise broader trends and stories emerging from media sources monitored, ability to direct resources across top stories, ability to report to the wider team or department. Leads a language team (20 people on average).

Reports to: Editorial director

Systems/tools used: MS Word, A bespoke video monitoring interface, social media tools such as Tweetdeck. Familiarity and knowledge of more complex social media sentiment tools (able to direct others to use them).

Key statement: "I need to understand the broader context and handle resources"

Key responsibilities: Oversee resources, recognise broader trends and stories emerging, deploy resources against top stories, report to the wider department

Interest/project goals: 1) Obtain a broader summary view of information viewed by their team 2) Get alerts for breaking events

Current frustrations: Unable to stay across all breaking developments, this impacts ability to assign resources and respond to breaking developments quickly flexibly.

2.2.4 Senior Editorial Lead

Profile: Media monitoring

Profile organisation: Sits across language teams

Industry history/length of service: 10–15 years

Languages: None in addition to native language of media partner organisation.

Data sources covered: General news sources and internal

Key skills: Awareness of stories, trends and issues for relevant region, contact with other news agencies, editorial leadership, ability to define best sources and tools for a story, impartiality and balance, oversee quality of news, liaise with wider news teams.

Reports to: Multiple language teams for their region, report into Editorial Director

Systems/Tools used: see Team Manager (editorial lead)

Key Statement: "I need to understand broader regional trends over longer periods of time in a wider context, looking across different locations and regions"

Key Responsibilities: Oversee language teams for the region, maintain an overview of output and quality, liaise with wider organisation

Interest/Project Goals: Need to see a broad overview in English of major trends and storylines

Current Frustrations: Not having a clear view of broader up-to-date events around the world across languages.

2.2.5 Editorial Director

Profile: Leads media monitoring teams including overseas language teams

Organisation: One Editorial Director sits across Monitoring teams

Industry history/length of Service: Approximately 15 years

Languages: None in addition to native language of media partner organisation.

Data sources covered: General news sources and internal multilingual sources

Key skills: Maintain a strategic Overview of the breadth and quality of output Reports to: All senior editorial leads report into the Editorial Director, the Editorial Director reports to the Director of Monitoring[1]

Systems/tools used: see Team Manager (Editorial Lead)

Key statement: "I need to understand the entire breadth and depth of monitoring department's output over the long term"

Key responsibilities: Overseeing the editorial output (news stories published online, on broadcast, etc.),ensuring adherence to BBC editorial standards, liaise with senior editorial staff.

Interest/project goals: Ensure media monitoring team has adequate coverage of current main global stories

Current frustrations: Not having a comprehensive view of all output from media monitoring teams.

2.2.6 Director of Monitoring

Profile: Monitoring Director

Organisation: Director of BBC Monitoring

Industry history/length of Service: N/A

Languages: N/A

Data sources covered: Needs to know what media monitoring teams are doing and also the big stories around the world (big news organisations and internal output). Concerned with what's happened and quality of this media organisation's news coverage

Key skills: Senior Executive skills, strategic, leadership skills, connections within the BBC

Reports to: Head of the organisation

Systems/tools used: Nothing specific

[1] The Director Of Monitoring role is unlikely to use any of the MT tools or prototypes and as such is not described here.

Key statement: "I need to know that monitoring is meeting the strategic requirements of all our users and stakeholders"

Key responsibilities: Ability to position monitoring within the media industry landscape, report into high-level new team.

Interest/project Goals: 1) Overview of news events 2) Increased media monitoring efficiencies and 3) New ways to cover sources

Current frustrations : Lack oversight of what media monitoring department is doing across languages in total on a given day.

3 Use Cases

There are two *levels* of use case described in this section. Firstly the project has three overarching use cases for the translation technologies. These will be described in subsection 3.1. The media partners, BBC and DW, have specific use cases for the use of the translation technologies. These specific use cases are described in sections 3.2.1 and 3.2.2 and will all fit under one of the project use cases from section 3.1. In addition to the media partners' use cases, two generic interfaces will be produced as part of the project. These will be a generic browser-based UI used for basic translation, and an API used for machine access to the translation technologies. These are described in Section 3.3.

3.1 Project Use Cases

3.1.1 Global Content Creation

Global content creation covers the process of creating media content in multiple languages. Primarily, this use case addresses the question of how multilingual content creation can be made more efficient for the media partners.

This project will focus on the production of news content in a multilingual newsroom. In this context, content will primarily include text news articles, but will also include the reversioning of video and audio media. Reversioning is the editing of the content of a translated news story such that the text will better suit the intended audience (extra information added where assumptions have been made, unnecessary detail removed, etc.) and is a part of post-editing which is a step over and above merely cleaning mis-translations.

The BBC World Service has recently undergone a large expansion, now broadcasting news in 40 languages around the world. BBC's flagship Arabic and Persian services operate 24-hour TV news channels while many others, including Kyrgyz, French, Russian, Ukrainian, Pashto, Burmese, Hausa and Tamil broadcast daily 15 minute TV news bulletins via local partner stations. Deutsche Welle, the German counterpart, currently publishes in 30 languages. In addition to its four live channels, Deutsche Welle focuses on distributing its content online, with text, audio and video (mostly on demand) in 30 languages. Both media partners distribute content via multiple platforms: radio, TV, online text, podcasts, YouTube and various social media, etc.

All foreign language services in both the BBC World Service and Deutsche Welle publish news online. This is extremely important to promoting the reach of the news published by these world services, especially to under-served markets.

The focus of this use case is in four areas:

- How can news stories be made discoverable across languages?
- How can it be assured that content (video, audio, text) created in one language is available to other language departments for re-use?
- How can the re-use of content and the efficiency of the news organisation be monitored?
- How can we develop reporting capabilities for new languages with limited resources?

Each of the four areas above requires a translation step, and with the appropriate user interfaces and machine assisted translation underpinning the translation step, a journalist is able to take a news story or script (in the case of an audio or video report) and quickly translate the original text.

At present, this use case assumes that any machine translation must be manually edited to ensure it is in a state which is of sufficient quality to be presented to an audience. No matter how good the machine translation, this is an important step for media organisations such as the BBC and DW who pride themselves in the quality of the output. It is also important to recognise that when content is translated, it is usually being prepared for publication in a different geographic region from where it originated. Local knowledge, assumed geographical or cultural knowledge and colloquialisms must therefore be expunged or explained in the translated copy. For this use case, and those described in section 3.2 an appropriate, and preferably integrated, user interface is essential to supporting rather than hindering a journalist performing this task.

3.1.2 Media Monitoring

Media Monitoring is the monitoring of broadcast and written media in order to facilitate newsgathering. Monitoring the international news media is of critical importance to news broadcasters such as BBC News and DW and is typically used to spot breaking news stories or spot trends in, for example, the news narrative over time.

News, however, does not only break in English for the BBC or German for DW, instead an effective monitoring operation will monitor the media landscape in a vast variety of languages. Considering the huge growth in the number of streams of data that could potentially be monitored, the current, predominantly manual, processes for media monitoring fail to scale adequately. It is becoming imperative for technology to be used in this process to automate tasks, such as translation, in order to free monitors and journalists to perform more journalistic tasks that cannot be achieved with technology.

BBC Monitoring undertakes one of the most advanced, comprehensive, and large scale media monitoring operations world-wide, providing news and information from media sources around the world. Around 200 journalists and analysts monitor TV, Radio, internet, and Social Media sources in order to detect trends and changing media behaviour, and to flag breaking news events. The BBC Monitoring operation covers over 150 countries and 100 languages worldwide. Media monitoring journalists are on the lookout for emerging themes—political, societal, and economic—and aim to anticipate future developments and events. The experience and expertise of monitoring analysts and journalists are also required to understand changes in the behaviour of media sources, especially those that may be associated with particular points of view and underlying political powers. This use case will focus on the use of MT to help specialist media monitors and journalists efficiently track stories across languages, including in particular low-resourced languages.

This use case is subdivided into two types of use for media monitoring. Namely:

External media monitoring which is monitoring of media across the world, external to that broadcast or otherwise distributed by the media partner.

Internal media monitoring which is the monitoring of a media partner's own multilingual output to promote efficiency across multilingual editorial teams.

The internal media monitoring use case is therefore intrinsically linked with the global content creation use case as both support editorial teams producing output.

Both media partners have interests in both external and internal media monitoring and will generate specific use cases accordingly (section 3.2).

3.1.3 International Business News Analysis

We will demonstrate the commercial potential of our research by delivering an international business analytics platform. Currently there are no solutions for reliable machine translation where the domain is highly specialised and there is little or no in-domain training data. The financial domain is highly specialised with terminology relating to products, processes and organisations, all of which have to be translated correctly for the meaning of the content to be preserved. We will deploy solutions developed for low resource language pairs to handling the problem of low-resource domains. This will demonstrate the wide applicability of our research, and broaden the commercial potential of the project. This final use case will be developed by focusing on international news topics covered by our partners, and on languages of high commercial interest.

3.2 Media Partner Use Cases

During months 1-6, each media partner undertook an internal study of potential uses for translation technologies within their respective organisations. Here the results of that research are described as a series of specific use cases.

The media partner specific use cases described in the following subsections will be used during the remainder of the project to shape thinking about how the translation technologies may be evaluated in a real-life business setting. That is to say, they are used as enablers for the relevant human evaluation methods as described by deliverable D5.1.

Additionally, these use cases and examples thereof have been written in such a way to give the reader an *illustration* of the range of possibilities for application of MT technologies for an large, multilingual, news organisation.

It should be stressed here that too many potential use cases exist to prototype exhaustively under GOuRMET. As such, a subset of these use cases (or variations thereof depending on business need and specific needs regarding evaluation) will be chosen for prototyping during the course of the project.

3.2.1 BBC Use Cases

This section describes the findings of the BBC's research into possible uses of MT. In order not to limit the use cases generated, all avenues for the use of MT have been explored, not only those reliant on low resourced languages.

Overview

As a multilingual organisation the BBC operates in more than 40 languages across the globe. With this comes many benefits, but also challenges. Internally, language can be a barrier to giving team's visibility of, and sharing, their editorial content.

Journalists can often be frustrated that they cannot use or review content published in languages beyond those they speak while those working in media monitoring want wider accessibility to local content in the vernacular in areas of the world important to them.

Reversioning of stories with wide international appeal into multiple languages is an efficient and economical approach to helping grow content and audiences. Reversioned stories are reported by journalists at some services to account for up to 70 per cent of output across online, radio and TV. Multi-lingual staff do the vast majority of this reversioning manually.

Journalist effort is required for each stage of the reversioning process. It is the availability of this effort that is the constraining factor on the amount of content that can be made visible to a wide internal audience and, as a knock on, to the external audiences.

The time burden, for smaller language teams in particular (who may spend much of their time translating English World Service reports for their own TV and radio broadcasts), may mean that making their own original journalism accessible is not an achievable priority.

Offering greater choice to audiences is an undoubted aim and the inevitable benefit of more efficient workflows.

There is massive scope within the BBC for high quality machine translation and computer-assisted translation.

It is an area in which the BBC has long seen potential and has previously experimented and prototyped. However this work has so far had little impact on business as usual workflows. Quality of outcome is perhaps the overriding factor for this.

Quality of translation and the ability to improve with training are likely to positively influence the outcome.

Primary areas of benefit for MT within BBC News, World Service and Monitoring can be summarised as follows. Direct references to personas in Section 2 are provided where appropriate.

- World Service journalists (2.1.1) want to have greater visibility of the best and most relevant content across the BBC regardless of language so that their service can make autonomous decisions about what to reversion for their audiences.
- World Service journalists (2.1.1), particularly at smaller services, want greater opportunities for their best original content to have a platform for sharing across the BBC and not be isolated by language.
- Regional and Central Services news editors (2.1.2 and 2.1.3) who may not speak all relevant languages, want greater independent editorial oversight of the output for which they are responsible.
- Monitoring journalists (2.2.1 and 2.2.2) want greater independent access to source material in local vernacular in areas important to them.
- Monitoring editors (2.2.3, 2.2.4 and 2.2.5) want to widen access to local vernacular content to facilitate flexible coverage to fit fluctuating demands of the news agenda and avoid bottlenecking at, or overburdening of, individual journalists of little spoken languages.

Use case details

USE CASE A: Improving internal visibility

In an ideal newsroom all journalists would be able to freely access the work of their peers, regardless of language. This would mean that quality of research and journalism would be free to speak for itself and all output within the organisation would be on an equal footing.

However, despite efforts to make output as visible and accessible as possible to all staff throughout the organisation —through a number of internal content systems— language barriers remain a blocker in realising this.

This lack of visibility provides a number of issues for staff around the BBC, perhaps most noticeably for World Service staff working in regions where service target audiences are neighbouring but output is in mutually unintelligible languages.

For smaller services in particular this would also help provide a platform for showcasing the best of their original journalism, something that too often currently, remains confined to a single service.

Specific examples include:

Example A1: Azeri service access to BBC Persian website

As a northern neighbour to Iran - and with a substantial Azerbaijani community resident there – Iranian stories are often of interest to Azerbaijanis both sides of the border. This makes BBC Persian a potential source of stories for BBC Azeri.

However Persian and Azeri are from different language families and many within Azerbaijan itself do not speak Persian or read the Persian script.

If a parallel version of the BBC Persian website was internally available in Azeri for example, service staff could easily browse content and make their own selections from the full offering.

With the bulk of the translation already complete the Azeri team could continue to work with Persian staff to assist in providing the best translations for quotes, confirming facts and ensuring nuances are correctly captured.

Example A2: Everything accessible to everyone in a single bureau

In the BBC's large international hubs a number of language services work side by side.

The geographical proximity of the target audiences for these services means that there is a potential for a large cross over in content of interest.

An example of such a large regional bureau is that in Delhi, which has teams publishing content in Hindi, Gujarati, Bengali, Marathi, Punjabi, Tamil, Telugu and Urdu.

Example of the issue in action: screens showing the most clicked-on news stories are prominently positioned in regional hubs, in order that journalists are informed of stories which are of the most interest to audiences at that time. However, as content can only be shown as published, not all teams may understand the headlines published on this live information display.

Example A3: Optimising visibility in internal content systems

BBC Internal systems allow journalists, regardless of division or service, to see and search continually updating news output from across the BBC alongside agency wire intake. Though devised to facilitate full transparency of material at a journalist's disposal these tools only carry text from world service language sites in the language/s as published.

Catch lines (equivalent to headlines within a feed) within these systems are in English and are able to attract a journalist's interest for reversioning or research purposes e.g. `IRAN-EARTHQUAKE`. However journalists in a multi-lingual environment may then find the story headline and content itself are in a language inaccessible to them.

If there was an option for a user to translate non-English content the full range of material could be made searchable and accessible. Perhaps as a first step the headline and summary - as presented in the default interface - and the full item on demand.

Example A4: Allowing experienced journalists to feed back

Guided by the BBC's editorial standards, more senior colleagues review and sub-edit ("sub") the work of junior colleagues. When there is a language barrier between the subbing journalist and source material it becomes an arduous and time-consuming process. For example Delhi-based staff of BBC Monitoring are asked to sub work from Afghan contributors. This has left subbing journalists reliant on commercially available MT tools to check the veracity of the claims and validity of the translations from source material.

If source links were submitted as part of this process, MT could streamline the subbing journalists quest to understand the material via easy access high quality translations.

Example A5: Facilitating greater ability to share

Though the author of a piece of original journalism may flag it up for reversioning, they may also conclude they are not in a position to offer the extra effort to translate just in case someone wants to use it. Time pressures on journalists, particularly in smaller services, make the latter a likely position.

If journalists were able to submit a piece for machine translation, it may facilitate a low-production-effort 'shop window' for high quality original journalism.

Universal visibility (of an acceptable standard of translation) of their piece and responses to it would allow language teams to make informed decisions on whether producing a completely re-versioned text is a productive use of time.

Example A6: A system of alerting journalists to stories doing well across languages

Specialist teams within BBC News and World Service identify potentially well-performing wide-appeal stories and make them internally visible to all journalists by commissioning translations into English, but some great pieces may slip under the radar.

This is not a new idea and has been previously prototyped via the "Sandcastle" prototype[2] as part of the SUMMA[3] project. Sandcastle was a multi-lingual alerting system prototyped by the BBC using language technologies resultant from the SUMMA project. The Sandcastle prototype aimed to identify news stories published across the BBC's non-English news websites and alert when an abnormally high clickthrough rate is identified for a particular story. This allows translation and reversioning of the popular story to be prioritised in the hope of publishing the translated stories across many languages at the height of that story's interest.

There exists a desire to revisit the Sandcastle prototype, to improve existing translations and widen the number of languages alerts were available in.

USE CASE B: Increased workflow efficiency for reversioning output

[2] http://summa-project.eu/2019/03/21/using-sandcastle-to-cope-with-a-tidal-wave-of-stories/
[3] http://summa-project.eu/

Removing the need for time-hungry routine manual translation processes by providing machine assistance to reduce reversioning effort.

Specific examples include:

Example B1: Reducing time to Broadcast

The Languages TV Unit (TVU) supports World Service languages by curating video content for reversioning. Videos and transcripts of the BBC's best original TV journalism from UK network News and other News broadcasting (such as Newsnight) are supplied to World Service teams for translation into local languages and pictures re voiced.

An internal English language speech-to-text tool assists transcription effort of the original English-language soundtrack. TVU journalists tidy the textual output before it is sent to language services for translation.

The limitations of this process are apparent when in the context of the speed needed in the live news environment. Sometimes the latest news items cannot be processed sufficiently quickly for inclusion in an upcoming broadcast.

Compression of the timeline from acquisition of the original report to ready-to-air reversion would give greater options to editorial decision makers and improved output for audiences.

A tool producing a high-quality translation and reducing journalist effort could see time for this process dramatically cut.

Example B2: Efficient subtitling of video content

Many great online and TV broadcast videos are produced each week around the BBC and language services often reversion these – translating captions and subtitles – for their own audience.

Reversioning systems have enabled users to efficiently update closed-caption text by filling in a simple online form, where they can read the English original and type their translation into an adjacent box, before applying to the video.

'Stitch'[4] is the BBC's experimental reversioning system for converting English-language digital videos for social media platforms to other languages. This process can be time-consuming since videos are assembled on an element-by-element basis using internal BBC systems. Stitch allows journalists to use a simple web form which automates the reversioning process, reducing the need for specially-trained editors. Users do not need knowledge of video editing -– they enter a line-by-line translation of any text appearing in a video or graphic, while the software preserves the time coding of the original media asset.

Stitch has saved both journalist time and the need for each reversion to be carried out on expensive editing software. However as these videos can be several minutes long there can be tens of blocks of text to translate and type. Videos can take up to an hour to reversion.

The next step for 'Stitch' could be to offer a pre-translated transcript of the original video for editing.

With the need for minimal corrections, high quality machine translations could free up journalists to produce more original content.

Example B3: Subtitle and dubbing script computer-assisted translation for TV

When the BBC sells its content abroad the localisation team ensure that it is fully prepared for audiences in the destination country.

[4] https://bbcnewslabs.co.uk/projects/stitch/

Script translation is part of the localisation process resulting in the creation of a dubbed or subtitled programme.

Most of the content handled by the team is long-form factual and drama programming and currently external vendors are commissioned to do this work.

Machine translation could potentially enable some of this work to be brought back in-house

Example B4: Taking the pain out of localisation of metadata

Localisation of TV programmes includes production of metadata in the local language. Such metadata includes, for example, the name and brief synopsis of the programme for on-screen guides.

The process of translating programme metadata is currently either outsourced or completed by a localisation team largely reliant on commercially available MT tools.

The use of a bespoke metadata reversioning tool incorporating MT has the potential to improve the team's translation process.

USE CASE C: Editorial oversight

Removing language barriers to allow senior editorial staff - responsible for quality and compliance across services - to monitor, understand and guide output equally across all languages.

Specific examples include:

Example C1: Editorial oversight (online news)

Each language service has an editor responsible for its output. Above this level news editors are often responsible for multiple language services clustered by region.

This leaves even multi-lingual editors reliant on others to keep them updated to output in at least some of these services.

Where there are few language speakers among staff at certain times of the day this overview of output can be particularly challenging.

News editors are required to oversee the editorial direction of the stories published in their region and are responsible for ensuring consistency and quality across all languages for which they are responsible. These include consistent format across the region, ensuring all relevant information is present, and ensuring the story is balanced and within editorial policy,

Reliable high quality MT may allow more detailed and timely oversight for language editors across a region's languages.

Example C2: Editorial oversight of compliance (TV)

Any output by the BBC must meet standards for legal, editorial and safety risk compliance within the context of the target nation(s) for broadcast. Individual service heads hold the editorial responsibility within their service for all material broadcast – whether this is original or reversioned.

A tool that allows those responsible to follow what has gone on air could improve standards and consistency between services.

To achieve this a two-part process would be required. Firstly transcribing reports in the vernacular followed by translation of the text into English

Potential for inaccuracies would be exacerbated by pre-translation transcription imperfections. However the realisation of an acceptably performing tool making multi-lingual broadcast content

accessible and searchable in a single language (English) could ease the burden of internal output tracking of difficult to reach content.

Example C3: Editorial oversight (Social media)

Senior journalists from the central Digital Service oversee social media accounts across all BBC output languages. Their role is to monitor editorial compliance and engagement statistics as well as provide support and representation, such as speaking to stakeholders to develop policy and sharing best practices to improve quality.

Monitoring of all platforms is carried out by a small team of 4 or 5 who may have up to 10 languages under their jurisdiction. Each has languages amongst their share that they themselves are unable to speak.

The platforms monitored are primarily Facebook, Twitter, Instagram and YouTube but 10 – 12 other regionally more popular platforms, such as VKontakte in Russia, are used by the appropriate services.

There are some clear avenues for exploitation of MT for the social media team such as allowing all social media updates from relevant accounts can display in English for easily accessible editorial oversight.

USE CASE D: Media insight

BBC Monitoring observes and reports on media developments in over 100 languages from around the world.

This coverage not only includes well staffed teams working in widely spoken languages of large global players such as Russian, Arabic and Persian, but also a breadth of lesser spoken languages from smaller nations and regions of political significance, such as Kurdistan and North Korea.

As a result of editorial priorities and economic practicalities staffing levels for some languages are smaller than others. This means news agenda fluctuations, commercial commissions and staff leave can mean a relative shortage of speakers of a given language at a given time.

Machine translation could allow a broader pool of journalists with regional knowledge to meaningfully support language speakers work.

Example D1: Kurdish

A small BBC Monitoring team covers Kurdish-language media from London and in the region.

There are 20–30 million native Kurdish speakers based largely in a region that covers parts of Turkey, Iraq, Iran and Syria. The language is divided into three main groups: Northern (known as Kurmanji), Central (Sorani) and Southern (Pehlewani). They are not mutually intelligible without acquired bilingualism.

Kurdistan is a politically important region and on-going coverage in BBC Monitoring output is crucial. However, there is not always a Kurdish speaker available to search websites and Social Media to find the latest content.

Sorani Kurdish to English Machine translation could assist the team to:

- Stay informed as to what Kurdish media are reporting
- Get a sense of tone in the Kurdish media
- Spot relevant content that can be referred to freelancers

- Note interesting-looking stories that may be lost in more recent content by the time Kurdish-speaking staff are back on shift

Example D2: North Korean

Korean specialists are supported by a wider group of editorial staff who have good knowledge of the North Korean political situation but are not language speakers.

There are a variety of propaganda and state-linked websites of interest. Effective machine translations of Korean online sources would mean that a wider range of staff could use their knowledge to keep coverage capacity consistent.

A North Korean specific translation model would be valuable as commercially available alternatives are likely to be specialised for the South Korean vocabulary. This has evolved differently to that in North Korea, which retains a more traditional use of language and more words borrowed from Russian, as opposed to Americanisms in the South.

USE CASE E: Research and experimentation with semi-automated content production

Reversioning of the best wide-appeal stories is an efficient and economical approach to helping grow content and audiences across languages.

However, this still demands substantial amounts of journalist time. The result is that the treatment is generally restricted to prominent, less time-sensitive stories and inevitably some desirable content can't be worked with.

There is a desire to investigate future possibilities that could economise on production effort of suitable website pages for a greater breadth and immediacy to content that can be offered.

Journalists have noted that MT often results in specialist terminology and proper nouns translating poorly. It would be interesting to experiment in this area but with these previous criticisms in mind. A new model, trainable by journalists able to create custom dictionaries, could perhaps rapidly improve on previous results and provide valuable learning to feed into future approaches.

Experimenting with delivery of content to users after minimal editorial intervention could also incorporate new avenues of user research and provide insight into openness to automated content and perceptions of quality and acceptability.

3.2.2 DW Use Cases

Overview

This section describes the potential use of MT in general, but also in particular that offered by GoURMET at Deutsche Welle. Deutsche Welle is already using automated translation in its workflow to some degree and aims to optimise this, working on internal integration in its existing infrastructure and expansion to more and better engines.

Obviously, as BBC World Service and Deutsche Welle operations and objectives are very similar, much of the overall description of the use cases in section 3.2.1 applies to Deutsche Welle as well —it's mostly a difference of departmental organisation and workflow. To avoid repetition, this subsection focuses on the more specific approaches and objectives at DW, as they differ from those of the BBC (section 3.2.1). That section also rightly states the potential use cases are too numerous to completely cover, so an informed selection shall be made.

Below a selection of use cases are presented for two of the broad use cases in section 3, i.e. global content creation and media monitoring.

USE CASE F: Translation and Adaptation for Content Creation

Deutsche Welle has recently introduced the news.bridge tool[5] for pilot use by its editorial departments for computer-assisted translation. It is a customised tool developed in a Google Digital News Innovation fund project coordinated by DW, and is built upon the SUMMA architecture. It includes automated ingestion of content, transcription, translation, voice-over —with post-editing after every phase— and publishing in 100+ languages, incorporating a large set of engines.

In addition, many editorial departments use online tools at their discretion for translation of text, including Google, DeepL, and some smaller ones. This is very language dependent. For Hindi DW has used Go Translate[6]. The Portuguese team uses DeepL, with a fairly good translation into Portuguese.

As this is meant for production, DW requires the best possible automated translation output – of course this depends on the language pair – which is then post-edited to publication quality.

Why then, is another engine needed if for many languages there is already one? For most smaller —or low-resourced languages— Google is the only tool that covers these. Besides, Google does not cover all languages, for instance GoURMET is working on Tigrinya and Afaan Omoroo, which are currently not in Google Translate's offer. Two major problems exist in this respect. First of all, if it exists for that language pair, Google is often not good enough, in particular for reuse of content (translation for content creation), but even for comprehension and gisting (translation for monitoring). DW hopes to train and customise the tools on internally produced content to achieve a better results. Secondly, Google may not be allowed based on company policy, for data protection reasons.

Example F1: Translate English or German content for smaller language departments

Deutsche Welle's biggest language departments are English and German, and thus content from those two languages are most widely used and distributed. It is the widest content repository for reuse within our organisation and is a major source for the smaller language departments. Exactly these smaller departments face a shortage of staff and have to be very selective in what they choose to produce, due to time and staffing restraints. DW's objective is to speed up the process with Computer Assisted Translation (CAT) and change the workflow so that language departments can increase their production. DW has started using CAT for Hindi and Portuguese, but for many of the low-resourced languages, including Swahili and Turkish, the quality is below expectations. The staff want to use it, but post-processing takes up too much time. DW would like to improve this with the GoURMET MT engines so that these smaller languages get extra support and can serve their communities better.

This includes computer-assisted translation at various levels:

- full translation of text articles, of videos, of audios (for podcasting, for instance) for publishing parallel items
- adaptation, localisation of (parts) of text, video or audio
- translation of manuscripts for extraction and summarisation for inserts in social media videos

[5] http://newsbridge.eu/
[6] https://www.cdac.in/index.aspx?id=ev_corp_gist_go_translate

Example F2: Translate content from low-resourced languages into English for reuse

A specific DW focus is human rights and the situation of people in developing countries, for example, interviews with people in Zimbabwe, talking about their living conditions. DW would like to capture more of those situations and that means using content in smaller, e.g. African, Asian or East-European languages. Tools would therefore be needed that can easily translate from those languages into English to facilitate reuse and publication from those languages. It will help raise awareness about situations in those areas.

Example F3: Support our new Turkish YouTube Channel

DW manages the project called +90, offering a new YouTube channel in Turkish, launched by four major foreign-service broadcasters collaboratively, i.e. Deutsche Welle, BBC, France24 and Voice of America. Currently, it is a YouTube Channel —adding a linear (i.e. broadcast, not on-demand) TV channel in Turkish is also under consideration. +90 provides a comprehensive offering of news, to users in Turkey, as well as Turks living abroad, partly with the aim of strengthening freedom of speech and the press. The channel features a broad range of sociopolitical topics covered in reports, background stories and interviews and targets especially the longer narrative YouTube format, allowing us to dive deeper into the topics compared to regular news channels. It is still in an early stage, but eventually, the four participating broadcasters will contribute an equal amount of video content to the channel, which can be accessed by users at any time. Thus, the challenge is to provide sufficient interesting and high-quality content in Turkish on a daily basis. The Turkish MT engine to be developed in GoURMET would help the producers of this YouTube channel to find and use interesting content to be published in Turkish. In the case of DW, the idea is to adapt relevant DW source material (primarily from English or German) as efficiently as possible for this new channel into publishable Turkish, on the one hand – and to monitor and translate external content for potential input as clips or rewriting by Turkish staff, on the other. In addition, video items in this channel are subtitled and this makes the use of a combined translation/subtitling tool for Turkish ideal.

Example F4: DW distribution to partner broadcasters

Featured items, long documentaries produced by DW are shared with a large number of partner broadcasters, e.g. in Africa, Asia and Latin America, for distribution in those regions. Preparing the items for those target audiences is still very much a manual process which would be responsive to automation to a certain degree. The transcripts are provided with the video material and are thus available in text format, and subsequently translated into other languages, such as Swahili. DW and some of the partner companies are currently collaborating to improve and automate this workflow using the news.bridge tool. Currently, Google Translate is the only available engine for many of these languages, if it even exists, but considered barely adequate for publication quality, even with post-editing. DW would like to compare Google Translate output with GoURMET output, and train this module on that specific type of content to get better results. If so, the customised engine can be incorporated in the DW online subtitling/translation platform used for this by the partner broadcasters.

Example F5: Distribution to other regions with translated subtitling

Deutsche Welle Distribution also distributes its content for broadcasting and dissemination in other countries and regions and thus in languages outside its DW "standard" list of 30 languages. Use of computer-assisted translation of material for such third-party countries, e.g. Vietnam and Korea, would greatly benefit the workflow, as it now requires subcontracted manual translation in those countries, resulting in huge time delay and high costs.

USE CASE G: Translation for Cross-Lingual Media Monitoring

Deutsche Welle as a international broadcaster covers 30 languages and therefore multilingual availability of the content is vital to its operations. This takes different forms: making its own content available within the organisation (internal monitoring), accessing content from external sources from a variety of languages (external monitoring), and enabling monitoring and raising awareness of DW content by making its content available to partner organisations across languages (enabling monitoring), leading to wider dissemination and collaboration. GoURMET supports this by providing engines that can translate from smaller or low-resourced languages into a lingua franca, i.e. English, thus making the content from the original languages available in a generic language. The purpose of monitoring is understanding, gisting, awareness and does not require post-editing. DW journalists require as high a translation quality as possible, but realistically can work with less than excellent output for this purpose. For monitoring, low cost is a major factor due to the volume to be processed. A lot of content – possibly all content from specific feeds – are translated to allow monitoring of all the content for specific languages. Processing and translation happens automatically, regardless whether the items are actually read/used. The fact that GoURMET engines are customised, produced within the consortium and made available for further use to the consortium and in general as open-source, allows cost-effective usage.

Example G1: Internal monitoring

Here DW would make use of a fully automated translation process to make DW content available to all language departments in-house. The purpose is to distribute our content to our own staff, regardless of the original language and language knowledge of the staff. Currently, content produced in English and German is widely spread in the organisation, but all other language material is largely lost to other language departments. This is, of course, a major problem. Thus, DW has developed a monitoring platform in the SUMMA project and DW is currently in the process of integrating it and rolling it out within the organisation. The platform will allow DW to be disseminated internally. The benefits are clear:

- Give more credit to the work of our journalists from smaller departments.
- Provide more source material to journalists in other language departments.
- Provide a cross-lingual overview, regardless of source language, clustering items by topic and theme and including those from smaller language departments.
- Make our material searchable in a lingua franca and enrich our archives. This can be English, as is currently, the case foreseen in GoURMET, but it can be made available in German by adding another translation level using internal tools developed in SUMMA, compatible with the GoURMET dockerised module. English is obviously a language that most journalistic staff in DW master, but many are in the end more comfortable with German.
- Speed up the exchange of information on items that are planned or in production and thus improve editorial planning cross-lingually .

The SUMMA platform only covers 8 of the 30 DW languages. In order to make the tool really useful for making the staff aware of all DW content - and give them access to that content - all DW languages should be covered. Therefore, GoURMET will develop a special module that makes dockerised engines (compatible with the SUMMA platform), enabling dissemination of content in Swahili, Bulgarian, etc. through this platform.

Example G2: Enabling monitoring of DW content externally

Similarly, this monitoring platform will allow us to expand this monitoring of DW content to other collaborative partner organisations and exchange content in a specific framework, such as the ARD or EBU, or our collaboration partners in Africa or Asia. The result will be increased awareness and reuse of our content and wider dissemination of DW content into other language regions.

Example G3: External monitoring

Of course, Deutsche Welle does not only want to be aware of what it has published in-house, but is equally in need of knowing what others are publishing. Using the monitoring platform for a large number of languages and expanding it in particular to low-resourced languages will lead to a better awareness of what is out there. Some of the benefits are:

- Increased monitoring and access to content available elsewhere (outside of DW) .
- Increased ability for monitoring trending news and increasing awareness of what the news is.
- Expanding monitoring to regions outside of the current coverage.
- Provide access to content the journalistic users do not master themselves.
- Increasing the variety of source material for journalists.

TOOLS AND PLATFORMS

A few final words on how Deutsche Welle intends to use the GoURMET engines and make them available to the user.

In this section, two platforms have been under consideration, in particular in the use case descriptions above: SUMMA for media monitoring and news.bridge for content creation, i.e., translation and subtitling with post-editing options. The GoURMET engines will be among the drop-down list and can be selected by the user as MT tool for a specific language pair. These will be integrated into the platform via API access.

In addition, accessing and assessing the output will be through the UI to be developed for direct access by the user. See section 3.3.1 for more details.

It is important that the engines are made available in different ways (API for integration) and UI for direct user access. This allows ample opportunity for testing, evaluation and pilot use and ensures potential for continued use.

3.3 Machine Translation Access Modalities

The GoURMET project plan requires the creation of two methods of accessing the research Machine Translation technologies. While not use cases per se, they have requirements that are important to define, and as such these are covered in this document.

3.3.1 Generic Browser-Based UI

The project requires a generic browser-based user interface which will expose all MT language pairs and versions of the MT system as they become available. This interface will serve a number of purposes, such as supporting basic human evaluation (see deliverable D5.1) and allowing quick and easy demonstration of the MT technologies at outreach events (Workpackage 6). This basic UI should accept a text input in the source language and provide a text output in the target language.

Users of this UI will include members of the GoURMET project, both media partners and research partners, users connected to the project such as editorial staff working for the media partners, and 3rd parties granted access for demonstration purposes.

The UI should be as simple as possible, but at the same time as functional as possible and should be able to demonstrate all capabilities of the underlying MT systems.

3.3.2 RESTful API

Machine access to the MT technologies is essential for project partners to be able to create prototypes and extend existing tools which use MT - these are required by the project to understand how tools such as these work in a realistic news environment and to enable human user evaluation (see deliverable D5.1).

The Application Programming Interface (API) should be built using the RESTful paradigm. This is preferred to other solutions as this paradigm is an extremely common industry standard for APIs. As a result there is a large number of frameworks to support the development of a RESTful API. This should result in faster development times and a well-tested and robust base to the API. In addition, an API built on a RESTful paradigm should provide a simple and familiar interface for developers using the service due to the large number of available libraries that support interacting with a REST API and the widespread use of REST across the industry. Finally, the REST paradigm is inherently stateless which fits the translation operation where it is unnecessary to store state between requests to the API (i.e. the state or result of one translation operation has no bearing on subsequent operations).

In the context of Use Case definitions, users of this interface (i.e. the API) will be machines, although for the sake of clarity it should be noted that these shall be machines that are under the ownership/control of consortium partners or selected, authorised 3rd parties.

4 Requirements

This section describes the media partners' requirements on the translation technologies. This section also describes the requirements for the shared interfaces (with scope as defined by task T5.2 - Creation of shared interfaces and use cases in 3).

This section does not describe requirements for the individual prototypes or integrations used to evaluate the quality of the translation technologies themselves.

4.1 MT Technologies

4.1.1 Functional Requirements

- MT technologies are produced in language pairs as agreed following a request by media partners
- Language pairs will be bidirectional
- MT technologies shall accept text input only
- MT technologies shall perform appropriate pre-processing of the raw text input (e.g. tokenisation, truecasing, subword segmentation, etc.)
- MT technologies shall perform appropriate post-processing of the raw translation output (reversing the pre-processing).
- MT technologies shall be sufficiently robust with respect to the spelling or diacritisation variations for languages which are not sufficiently standardised.
- MT technologies shall provide word-by-word confidence scoring for an agreed subset of language pairs.
- MT technologies shall provide a selection of alternative sentences for each translated sentence, which may be ordered by predicted likelihood.

4.1.2 Non-Functional Requirements

- MT Technologies shall accept multiple sentences
- MT technologies shall be robust to zero-length
- MT technologies shall be robust to the inclusion of unexpected characters
- MT technologies shall ignore white-space characters other than as a word separator
- MT technologies shall translate a sentence in not more than 500ms (on average).
- MT technologies shall accept sentences of not more than 80 words.
- MT technologies shall accept UTF-8 character encoding only
- Output of MT technologies will include punctuation

4.2 Generic browser-based UI

4.2.1 Functional Requirements

- A user will be able to select a language pair for translation.
- The interface shall expose all available language pairs.
- A user will be able to input text for translation and receive a response.
- The UI will be able to authenticate a user and establish what actions they are allowed to undertake.
- The UI will surface any errors that arise during translation.
- The UI will surface any errors that arise during authentication. For example if the user is not recognised or their credentials are invalid.
- The user input will be sanitised to protect against attacks such as cross-site scripting.[7]
- The UI will allow the user to provide feedback on the UI and the quality of the translation.
- It shall be possible to add content by copy-paste

4.2.2 Non-Functional Requirements

- The UI will impose a sensible limit on the amount of text that can be translated in a single request in order to keep the performance speed of the tool high.
- The interface will be sufficiently simple that a user will be able to use the service without training, providing they have a reasonable level of IT knowledge.
- The UI will not allow unauthorised users to perform translation actions.
- The UI will allow a user to use and assess the quality the GoURMET translation models without needing to interact directly with the API.
- The UI will correctly function in recent versions of both the Mozilla Firefox and Google Chrome browser.
- The UI will be maintained for the duration of the GoURMET project.
- A process will be defined to allow potential users to request access to the UI.

[7] https://en.wikipedia.org/wiki/Cross-site_scripting

4.3 RESTful API

4.3.1 Functional Requirements

- The API shall expose all available language pairs.
- The API shall allow an incoming request to define which language pair to use for translation.
- The API shall be able to indicate which version of a language pair is in use.
- The API shall accept only text characters which are valid for translation
- The API shall accept UTF-8 character encoding
- The API shall accept only encrypted input for data security.
- The API shall supply only encrypted output for data security.
- The API will be able to authenticate a user/service and establish what actions they are allowed to undertake.
- The API shall respond with an HTTP error code if the input is invalid.
- The API shall respond with an HTTP error code if the service is unavailable.
- The API shall respond with an HTTP error code if the user cannot be authenticated.
- It will be possible for the API owner (in this case the BBC) to invalidate a user's authorisation if necessary e.g. a leaked API key.
- The API request and response shape will remain consistent across endpoints and for different language pairs.

4.3.2 Non-Functional Requirements

- The API will prevent the user from providing input that is too large to be handled by the API within a reasonable time frame.
- The API will not allow unauthorised users to perform translation operations.
- The API will be built to be able to handle spikes in traffic.
- The API endpoints will be documented.
- The API interface design will be such that an individual with a reasonable knowledge of using APIs will be able to use the interface with support from the API documentation.
- The API will allow a user to use and assess the quality of the GoURMET translation models without having to use the models directly.
- The API will be maintained for the duration of the GoURMET project.
- Major changes to the API user facing interface will be done via a new version of the API being released and the old version being supported for a reasonable time period to allow users to migrate to the new service.

5 Conclusion

This document presents an overview of the use case definitions and user requirements for the technologies and prototypes developed by the media partners.

The different personas describe the actors involved in the two primary use cases for the GoURMET project at the two world broadcasters involved, Deutsche Welle and BBC.

A comprehensive set of example use cases was presented, describing a large number of scenarios in which a multilingual news broadcaster can exploit the MT technologies under development in this project.

A list of user requirements is the result of the preparatory work as described in this report.

This deliverable will be followed by two other reports in this series: deliverable D5.3 (Initial progress report) in M18 and deliverable D5.5 (Final report on integration) in M36. For further information regarding how the use case examples described here will be used to enable human evaluation of the research MT technologies, the reader is directed to the companion document deliverable D5.1.

ENDPAGE

GoURMET

H2020-ICT-2018-2 825299

D5.2 Use Cases and Requirements