SELMA

Stream Learning for Multilingual Knowledge Transfer

https://selma-project.eu/

D6.2 Impact Plan

Work Package	6
Responsible Partner	DW
Author(s)	Ksenia Skriptchenko
Contributors	Kay Macquarrie, Tim Koch
Reviewer	Afonso Mendes
Version	1.0
Contractual Date	31 December 2021
Delivery Date	31 December 2021
Dissemination Level	Public

Version History

Version	Date	Description
0.1	16/11/2021	Initial Table of Contents (ToC)
0.3	06/12/2021	Draft Deliverable is ready for initial feedback
0.4	10/12/2021	Feedback from reviewer
0.5	16/12/2021	Finalized comments from reviewer
0.6	17/12/2021	Sent to the Project Coordinator
1.0	29/12/2021	Publishable version

Executive Summary

> The SELMA mission statement shows a central impact objective:
>
> **Shaping AI speech and text technologies for media & the newsroom**

This deliverable "D6.2 Impact Plan" presents a strategy of how SELMA is aiming to reach out to a relevant audience in (media and language technology) business, academia and general public and to create the desired impact among its interest community.

This document has three purposes:

- Dissemination Plan for SELMA: Provides an project dissemination plan by highlighting target groups and defining a dissemination strategy based on three phases: inform, involve and convince.

- Communication Plan for SELMA: Provides a set of communication means for a strong visual appearance and a coherent project identity. It provides metrics on reporting and an initial plan for sustainability.

- Exploitation Strategy for SELMA: Investigates initial ways of making use of results and exploiting SELMA NLP components, including the open-source platform. Exploitation activities will start in the second year of the project.

SELMA is a "Research and Innovation Action" project. Therefore, the project has two impact focuses: a) advancing the state-of-the-art in various NLP related tasks and technologies through improvements in research, and b) bringing technology to the market through tools, components and an open-source platform.

This document will be updated within the course of the project's developments after the second project year (M24) and at the end of the project (M36).

Table of Contents

Table of Figures

Table of Tables

1.Introduction

This impact report is divided into three main parts: Dissemination, Communication, and Exploitation.

Section 1 (Dissemination) outlines plans and shows efforts to inform, inspire and involve future SELMA users and the research community. SELMA's dissemination goal is to engage potential users, HLT (Human Language Technology) providers, and early adopters to establish a feedback culture and build a network with other projects, researchers, and technology users.

Section 2 (Communication) establishes the communication strategy and names the main communication channels and activities. It also summarizes the already carried out activities aimed at raising awareness about the project itself and its progress.

Section 3 (Exploitation) shows an overview of exploitation activities planned to start from M13. The initial exploitation path is summarized in this part.

The dissemination and communication plan as well as the exploitation strategy will be constantly revised, updated and refined throughout the course of the project. There will be two following deliverables: D6.4 Interim Impact Report (M24) and D6.6 Final Impact Report (M36).

2. Dissemination

The purpose of this section is to provide a project dissemination strategy by highlighting targeted groups and communities, define internal dissemination/communication guidelines and procedures, outline the foreseen channels, and report on the efforts in the first project year.

This section describes the efforts of defining, identifying, and reaching our target audience, focusing on its two primary target groups, i.e., the scientific communities and the media world. It provides the plan of selecting, setting up and supplying the right dissemination channels and a survey of the dissemination events to promote the results in the related fields of research.

Following the dissemination activities framework outlined in the DoA, SELMA will adopt a multi-channel and multi-target approach. SELMA pursues a clearly defined strategy, which will be outlined and specified further below in the sub-chapters.

2.1 Target Groups

Overall, we divide our target audience for dissemination activities as follows, and will address primarily:

- Broadcasting and media world

- Industries using language processing technologies:

- Translation agencies

- Media monitoring organizations

- Industries in need of monitoring multilingual content across the world media

- Scientific and research community

- Stakeholders and their networks

- Policy makers and interest groups

- Human Language Technology users

- General public (interested users)

2.1.1 Network

All SELMA consortium partners have well-established networks and they do their best to reach their target audiences both individually as well as in concerted project efforts.

The SELMA consortium also commits to support the Big Data Value Association in all events relevant to the activities of the project.

2.1.2 User Group

The User and Advisory Board serves as an advisor to the Innovation Manager and the project's Steering Board. At the first User Group meeting (scheduled for early 2022), an Advisory Board consisting of 3 to 5 people will be formed.

User group activities include identifying the strong and weak points with respect to the objectives of the project (with emphasis on the innovation objectives) and providing recommendations. Furthermore, the members of the User Group help us maximize our industry outreach, serving as links between the consortium and external key industry players. Currently, around 15 representatives from research organizations, European media companies and technology providers are forming the SELMA user group.

2.2 Strategy

The strategy defines the key channels and activities to support the objectives of the project. The dissemination strategy also determines events, the materials created to describe and show project results, and targeted publications.

The project consortium has defined three major stages.

Year 1	Year 2	Year 3
		Share & Convince
	Involve & contribute	
Inform & inspire		

Figure 1 Three dissemination stages

2.2.1 Inform and inspire

The *"inform and inspire"* phase already started in the first year and will be active during the whole lifetime of the project.

The focus in the first phase is on:

- Outlining the project's vision, aims and goals,

- Setting up the dissemination channels and activities to spread the word and to inform target audiences about SELMA and its major objectives,
- Introducing and interacting with relevant communities, which might differ in the scope of different use cases.

2.2.2 Involve and contribute

The *"involve and contribute"* part will start in the second year and will continue until the end of the project.

This phase will be consisting of the following steps:

- Identifying key influencers to involve into the feedback loop, in testing early prototypes, sharing research results,
- Utilizing early adopters as multipliers and to spread further awareness,
- Obtaining user feedback on development and creating solutions for obstructions.

2.2.3 Share and convince

The last part of the dissemination strategy *"share and convince"* will start in the last year and will mainly focus on:

- Demonstrating progress by making available a variety of open-source resources and project outcomes,
- Showing prototypes and innovative features to targeted audiences and third parties,
- Engaging with target groups and individual users to support the SELMA exploitation activities.

2.3 Dissemination Plan

The chapter above outlined the overall SELMA dissemination approach and strategy. The details, on how we plan to reach the pre-set goals, on what has been planned and already achieved in the first twelve months of the project, are outlined in the following section.

2.3.1 Tools (Website, Twitter, LinkedIn, GitHub)

Website and Blog

An important channel of the SELMA communication and dissemination strategy is its website.

The SELMA homepage can be reached at *www.selma-project.eu*. It was launched by DW at M2. It provides information on the project, main goals and project partner descriptions, and contact person information. Currently, it has six sub-pages:

- Info
- Partners
- Output
    - Deliverables
    - Publications
    - Project One-Pager
    - Press Material
- User Group
- Related Projects
- Blog

The blog section of the homepage represents SELMA library of the relevant human language and artificial intelligence technologies contributions created and curated by consortium partners. It is mainly aimed at the general public.
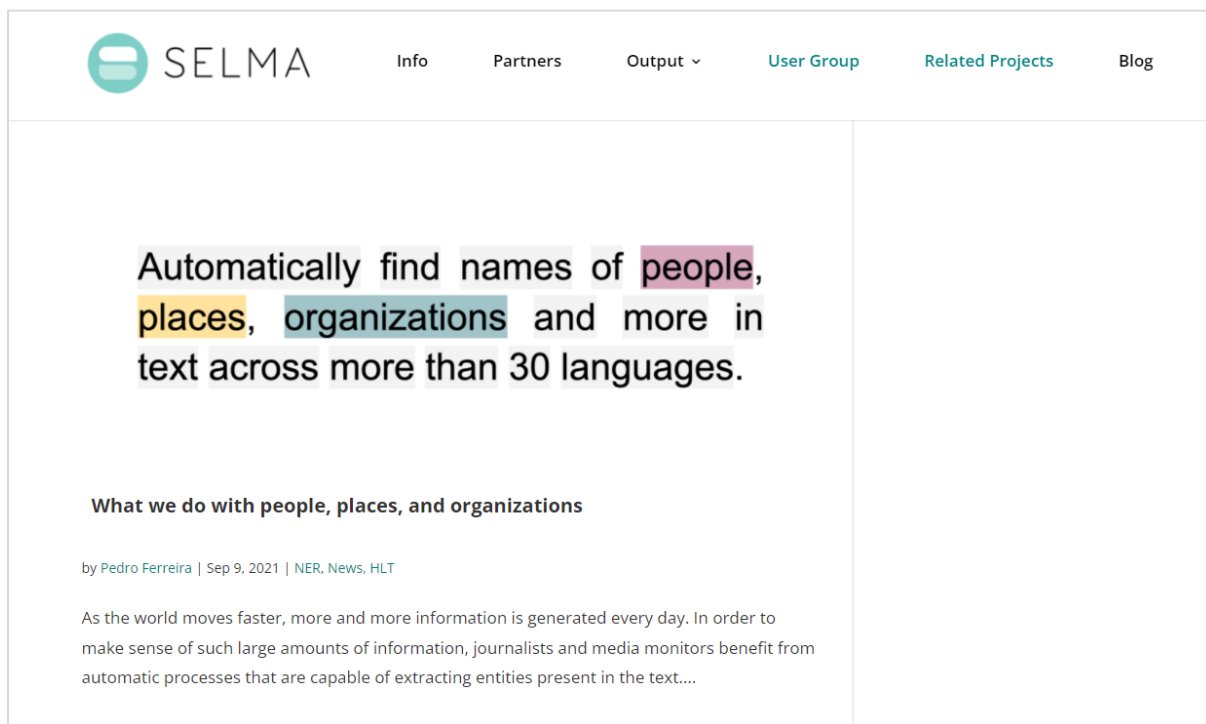

Figure 2 Screenshot of SELMA Website (Blog)

The following content was published in the first year of the project:

- Fraunhofer contribution on Machine Learning Methods

https://selma-project.eu/2021/12/16/a-yummy-piece-of-cake/

- Priberam contribution on Named Entity Recognition

https://selma-project.eu/2021/09/09/what-we-do-with-people-places-and-organizations/

- IMCS UL contribution on Infoboxes & Knowledge Graphs

https://selma-project.eu/2021/08/01/diving-into-infoboxes-knowledge-graphs/

- LIA contribution on Curriculum Learning Methods

https://selma-project.eu/2021/07/20/all-about-spoken-languages/

- DW contribution on Biased AI

https://selma-project.eu/2021/05/14/artificially-intelligent-yet-humanly-biased/

All SELMA partners are constantly encouraged to actively contribute content to the site, especially for the blog part.

As the project progresses, we will provide more regular updates including:

- various documentations on project activities and outcome,
- links to prototypes, demos and tutorials,
- code, datasets and resources.

Social Media

A SELMA Twitter account serves as the social communication channel. The account is accessible at *https://twitter.com/selma_project*.

Figure 3 *Twitter Account*

The SELMA Twitter account is used to post primarily scientific, and technology-related content, to inform about project activities and achievements, and to reach out to relevant target audiences.

The consortium also maintains its own LinkedIn page. It is mainly used for notifications regarding SELMA blog content and for announcements for upcoming events. The SELMA page is accessible at *https://www.linkedin.com/company/selma-project-eu/*.
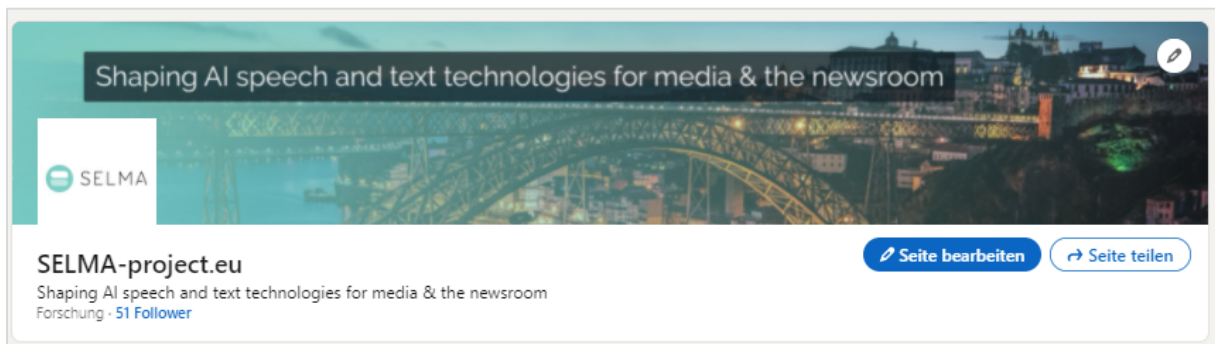
Figure 4 *LinkedIn Page*

We also set up a GitHub Page to be able to collaborate and share our output. It is accessible at https://github.com/SELMA-project. It is in a private mode for now.
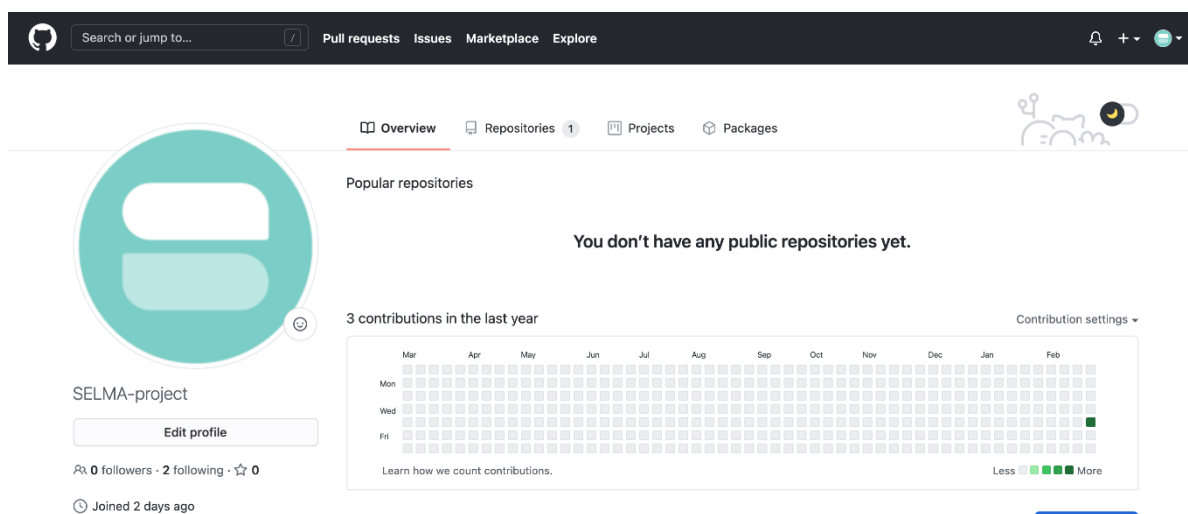
Figure 5 *GitHub Page*

2.3.2 Materials (Fact Sheet, Poster, Flyer, Roll-up)

In order to reach relevant influencers, the following physical dissemination materials were produced and promoted through the website and social media channels.

Project Fact Sheet

Very early-on in the project, we produced a one-page summary of SELMA to outline the primary project aims, objectives and main use cases.
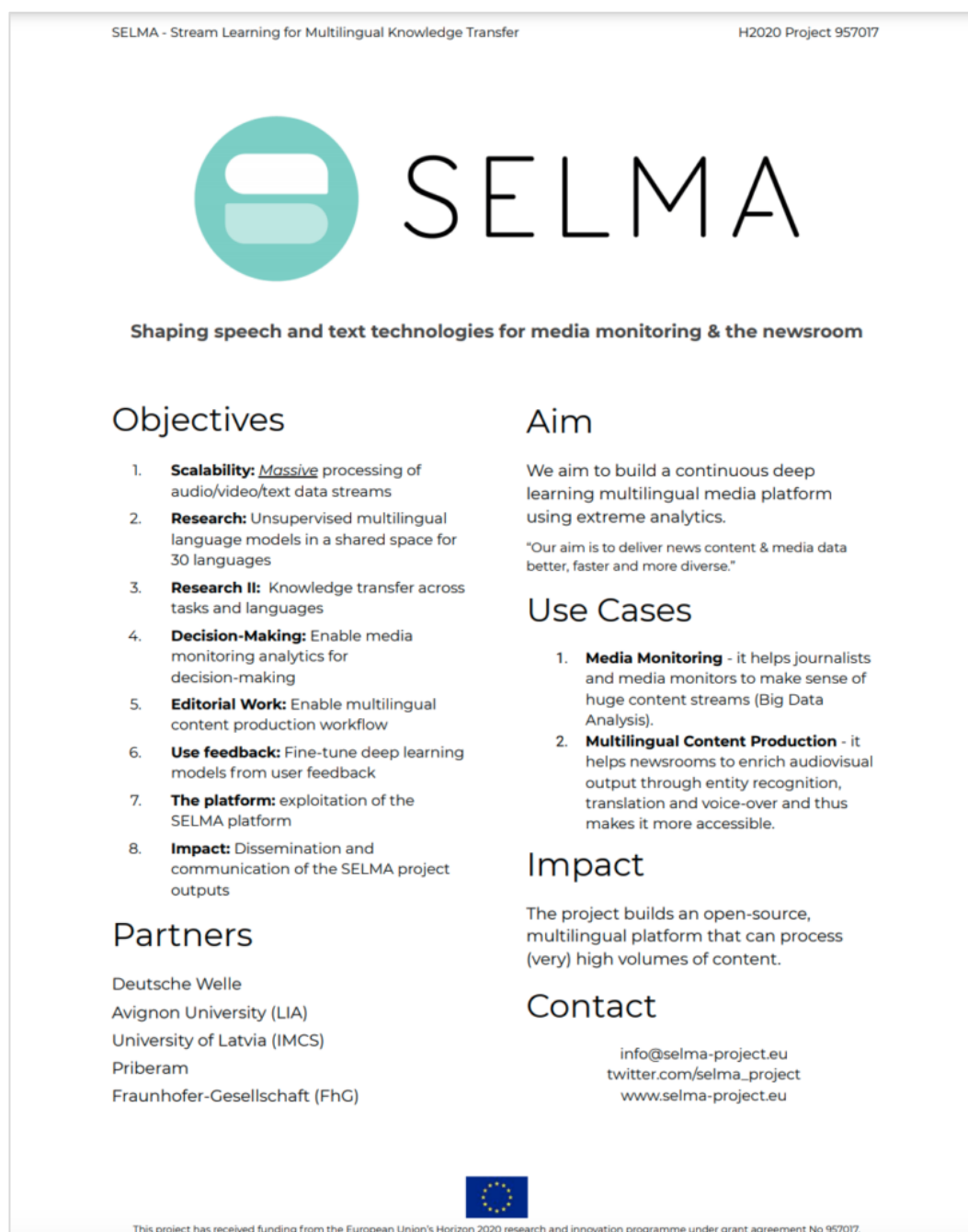
Figure 6 *Project Fact Sheet*

The one-pager also identifies five project partners and provides links to our digital channels – website and Twitter.

Flyer

The flyer consists of four different pages, as shown below.

Figure 7 *Flyer*

In addition to the general information regarding SELMA project, it also includes the list of languages covered by the project activities.

Poster

For presentations at events, conferences, workshops and other dissemination occasions, we have produced a SELMA poster and flyer.

The poster includes the simplified state of art workflow of the platform and the achievements of SELMA. It summarizes the aims and goals of the project, mentions and clarifies the SELMA target groups.

Figure 8 *Poster*

All SELMA digital channels and consortium partners including the EU commission are named at the bottom of the poster.

If needed, specific technical posters will be produced for poster presentation sessions focusing on certain SELMA technologies.

Roll-Up

We have also produced a SELMA banner. It can be rolled up for easier handling, hence the term "roll-up".

Similar to other dissemination materials, the roll-up was produced using the same design and color scheme, to provide a coherent "look & feel".

Figure 9 *Roll-Up and Poster Live and in Action at Festival IA 2021*

The following table gives a brief report on the status of planned dissemination tools and materials, explained and described above.

Tools/Materials	Status	Year
Website	set-up/ongoing updates	Y1/Y1-3
Blog	set-up/ongoing updates	Y1/Y1-3
Twitter	set-up/ongoing updates	Y1/Y1-3
LinkedIn	set-up/ongoing updates	Y1/Y1-3

Github	set-up/ongoing updates	Y1/Y1-3
Project Fact Sheet	completed	Y1
Poster	completed (can be adjusted for specific poster presentations, if required)	Y1
Flyer	completed	Y1
Roll-Up	completed	Y1

Table 1 Dissemination Tools & Materials

New versions of the aforementioned dissemination materials will be produced in the course of the project if required.

Publications

Publications in the scientific and academic spheres are an essential component of SELMA dissemination activities.

List of publications and research papers published in the first year is provided below.

Consortium Partner/ Names	Title	Published in	Open Access & Golden Standard (+URL)
Priberam: *Pedro Ferreira, Ruben Cardoso, Afonso Mendes*	***Priberam Labs at the 3rd Shared Task on SlavNER***	*Proceedings of the 8th Workshop on Balto-Slavic Natural Language Processing*	*Yes*
LIA: *Solène Evain, Ha Nguyen, Hang Le, Marcely Zanon Boito, Salima Mdhaffar,*	***Task Agnostic and Task Specific Self-Supervised***	Proceedings of the Neural Information Processing	Yes

Sina Alisamir, Ziyi Tong, Natalia Tomashenko, Marco Dinarelli, Titouan Parcollet, Alexandre Allauzen, Yannick Estève, Benjamin Lecouteux, François Portet, Solange Rossato, Fabien Ringeval,Didier Schwab, Laurent Besacier	***Learning from Speech with LeBenchmark***	Systems Track on Datasets and Benchmarks	
LIA: *Solène Evain, Ha Nguyen, Hang Le, Marcely Zanon Boito, Salima Mdhaffar, Sina Alisamir, Ziyi Tong, Natalia Tomashenko, Marco Dinarelli, Titouan Parcollet, Alexandre Allauzen, Yannick Estève, Benjamin Lecouteux, François Portet, Solange Rossato, Fabien Ringeval, Didier Schwab, Laurent Besacier*	***LeBenchmark: A Reproducible Framework for Assessing Self-Supervised Representation Learning from Speech***	Proc. Interspeech 2021, 1439-1443, doi: 10.21437/Interspeech.2021-556	Yes
LIA: *Sahar Ghannay, Antoine Caubrière, Salima Mdhaffar, Gaëlle Laperrière, Bassam Jabaian, Yannick Estève*	***Where are we in semantic concept extraction for Spoken Language Understanding?***	Speech and Computer: 23rd International Conference, SPECOM 2021, St. Petersburg, Russia, September 27–30, 2021, Proceedings	Not yet

Table 2 *List of Publications Year 1*

2.3.3 Activities

A very important part of the dissemination strategy is being present at relevant events for SELMA project work. It helps us to stay informed and up to date in the scientific areas, present project achievements and results, meet relevant stakeholders for future collaboration and cooperation, and prepare for exploitation activities. Given the pandemic restrictions, we have been attending industrial and academic events, the majority of which was held digitally.

List of attended or organized conferences (academic & industry), workshops, and other dissemination events in the first year of the SELMA project is provided in the table below.

Date	Title	Partner
10.1.2021	EBU presentation: DW Benchmarking Efforts	DW
5.2.2021	79th International Conference of the University of Latvia	IMCS UL
18-19.3.2021	BDVA PPP "project meet-up"	DW
23.3.2021	Seminar "A very brief history of Spoken Language Understanding"	Priberam, LIA
20.4.2021	3rd Shared Task on SlavNER (part of the 8th Workshop on Balto-Slavic Natural Language Processing)	Priberam
26.5.2021	Open "NLP" Workshop of GoURMET H2020	Fraunhofer
21.6.2021	EBU Workshop: MT for low-resourced languages applied at DW	DW
28.6.2021	17th Baltic Conference on Intellectual Cooperation (BCIC)	IMCS UL
5-6.8.2021	SIGSLT - Special Interest Group on Spoken Language Translation (SIGSLT)	LIA
27.9.2021	EBU workshop on AI applications, with BBC, RAI, DW and other EBU members	DW
30.8-3.9.2021	Interspeech 2021	LIA
6.10.2021	Interview AI in language technology	DW
25.10.2021	EBU Workshop: AI at DW	DW
15.11.2021	Presentation at Annual Meeting Goethe - plain X Application	DW

17-19.11.2021	Festival IA 2021 – Avignon Université	LIA
18.11.2021	GoURMET User Event	DW
24.11.2021	Presentation at ARD: plain X Application Outlook	DW
6-14.12.2021	NeurIPS conference 2021 (Thirty-fifth Conference on Neural Information Processing Systems)	LIA

Table 3 *List of Dissemination Events Year 1*

3.Communication

This section outlines SELMA's communication activities. It consists of three parts.

The communication strategy provides the project identity and guidelines for communicating internally and externally.

The "Evaluation" part of this section measures and evaluates the performance in the first twelve month of the project. It also suggests if and where it is necessary to re-adjust strategy and activities according to the *communication success metrics table*.

And finally, the chapter "Sustainability" provides the information regarding the accessibility and availability of the outcome and results provided on publicly accessible channels after the active period of the project.

3.1   Strategy

The SELMA's communication strategy describes the channels and methods used to efficiently implement communication:

- within the project (internally)

- with peer researchers, related projects

- with end users and the public at large.

It is important to ensure that the entire consortium is informed of the development status and achievements, also of the challenges to overcome, and the efforts invested. Equally important is communication towards the relevant research community, related projects, and our main target group consisting of media and broadcasting professionals to inform and ensure that the SELMA results can be used in professional communities outside of the consortium and relevant feedback can be provided during the project's lifespan.

The SELMA output can be used by a wide range of users interested in consuming content in other languages, using subtitles, full-text transcripts, or voice-over applications. Hence, and to ensure the continued use of the outcomes and results, it is success-critical to inform the possible end-users and general public about advantages, limitations, and overall progress of the SELMA's outcome.

3.1.1   Project Identity

The vital part of the communication (and hence the dissemination) strategy is having a clearly defined project identity to guarantee the perseverative and immediate recognition and to stay distinguishable from other projects. The project identity consists of a coherent concept which defines the logo, mission statement and all media outlets such as the website, social media

activities, and marketing material. The stronger and more coherent a project identity is, the more awareness can be raised.

In addition, every communication channel and document have a common style and appearance, to increase project visibility, recognition, and consistency.

3.1.2   Logo

The project identity is expressed by consistent use of the project logo in all its communication, creating a direct visual recognition of the SELMA project. Hence, the logo is simple, appealing, and easily recognizable. It represents a simplified portrayal of two speech bubbles in a dialog.

Figure 10 Logo

The logo appears on the project website, as well as on all external social media channels and communication documents, i.e., deliverables, flyers, and posters.

3.1.3   Mission Statement

For successful communication, it is crucial to have a mission statement that provides a clear and concise definition of the unique selling point. The SELMA mission statement consists of following slogan:

"Shaping AI speech and text technologies for media & the newsroom"

The slogan is meant to attract people's attention and make them curious to learn more about the SELMA project and its products. It is used in combination with a corresponding image of a bridge, symbolizing the connection between the different user groups through the SELMA project.

Figure 11 Slogan & Picture

3.2 Evaluation & Reporting

The table below evaluates the progress on the communication means and outreach:

Activity	Planned (in total)	Description	Achieved (M12)
Website Visits	24.000	Y1: 7.000 Y2: 7.000 Y3: 10.000	7721
Twitter Followers	180	Y1: 60 Y2: 60 Y3: 60	61
Tweets	150	Y1: 50 Y2: 50 Y3: 50	50
LinkedIn Followers	90	Y1: 30 Y2: 30 Y3: 30	51
LinkedIn Posts	50	Y1: 15 Y2: 15 Y3: 20	12

Videos	10	Y1: 0 Y2: 5 Y3: 5	0
Poster	1	(updated version if required)	1
Flyer	1	(updated version if required)	1
Roll-Up	1	(updated version if required)	1
Events (participated and organized)	55	Y1: 15 Y2: 20 Y3: 20	14
Publications	11	Y1: 3 Y2: 4 Y3: 4	4
User Days	4	Y1: 0 Y2: 2 Y3: 2	0
GitHub Watches	30	Y1: 0 Y2: 0 Y3: 30	GitHub account will be made public as of Y3

***Table 4** Overview Communications Means & Progress*

For 2021 all plans were successfully met.

3.3   Sustainability

To ensure the availability of the work performed by the SELMA project, the consortium aims at keeping dissemination channels available for a period of at least three years after the end of the project. This will include the website, and the social media channels.

All dissemination channels will be kept available after the project end, with very limited editorial input though.

4. Exploitation

This section of the deliverable focuses on capturing the information required and explores options to ensure the outputs of the SELMA can be exploited by consortium partners and others. As the exploitation roadmap, this document sets out the activities required for the successful exploitation of SELMA results. The exploitation activities will start in the second year of the SELMA project, required IPR strategy will be set up at the beginning of the task.

4.1 Exploitation Ways

There are four main ways in which SELMA can be exploited:

- **Orchestration platform**: The focus is to provide an open-source big-data platform able to ingest and orchestrate the pipeline graph of NLP modules and apply stream learning and user feedback. During the project the partners will investigate the options further and will aim to give further consideration to the revenue potential of the best options.

- **Component-based/individual system modules:** The SELMA components/ modules have a high potential value as improvements for existing services, or as the basis for other new services.

- **Integration into other projects/products:** Results will be also integrated in MONITIO project (an H2020 FTI project for AI powered Media Monitoring lead by Priberam) and will be a part of plain X service, a novel human language technology (HLT) platform developed by DW Innovation in cooperation with Priberam.

- **Knowledge**: Through continuous research, each consortium partner is learning and developing new techniques that can be applied to succeeding projects. It also can be utilized for designing new approaches and new services.

4.2 Exploitation Plan

The initial exploitation of the consortium is shown in the following table.

Partner	Exploitation Plan
LIA	Avignon Université (LIA) will use the technologies developed in the project to improve both the university's expertise in enriched transcription and translation and will expand this to its research network. LIA has many collaborations with industrial partners that could be interested to the results of the SELMA project: for instance Vecsys/Bertin, a French company that develops vocal technologies on speech processing commercialized all over the world; Orange has continuously shown interest in semantic information extraction from speech, Airbus is also a regular LIA partner interested in speech and language processing, or the INA that needs automatic speech and language processing to manage millions of video documents. LIA researchers are largely involved in a Master's degree specialized on machine learning and language processing proposed by the Avignon Université. Research in the SELMA project will feed courses and (lab) seminars. Lastly, LIA aims to publish high-level scientific publications in major peer-reviewed international journals and conferences. LIA has also very strong relations to the Language and Speech Technology team of LIUM, a public research laboratory hosted in Avignon Université. Some future collaborations between LIA and LIUM researchers will have a positive impact on the SELMA project.
DW	Deutsche Welle will be directly involved with the system as user partner. Its data is ingested into and processed through the platform. A high level of customisation is envisaged, for instance, voice synthesis is trained on and customised for some DW news readers. Punctuation as well as summarization is trained on and tested with DW content. Deutsche Welle journalistic staff and the development team are involved from an early stage, to enable and facilitate early implementation and beta testing in-house. The SELMA system is compatible with an HLT platform currently under implementation at DW, resulting from the SUMMA and news.bridge projects. SELMA will ensure continuous improvement and sustainability of HLT applications in a large number of languages. Thus, Deutsche Welle intends to use the platform to process and distribute its content in more languages, serving also smaller language departments, and introduce automated subtitling and voice-over, with post-editing, as standard procedures. It will work with other broadcasters in its network for wider use and feedback.
Fraunhofer	Fraunhofer will use the technologies developed in the project in industry projects for the German media industry, especially for German public broadcasters. In its long-term collaboration with the public broadcaster WDR as the lead buyer of AI technologies for the ARD group, Fraunhofer will introduce the developed technologies to fully search the vast media archives of the public broadcasters. The technologies and advances described in the proposal are closely correlated to demands in the media industry, especially related to improved speech recognition, punctuation prediction, live subtitling and accurate speaker recognition. The results obtained in the project are consequently directly supporting public broadcasters in Germany to improve their programm, e.g. by more efficient search of their archives or faster production of media assets and subtitles, thus reducing costs and delivering news faster and more reliably. Besides the media industry, Fraunhofer is supporting German state parliaments to automatically subtitle their parliamentary debates to ensure barrier-free access for the Deaf and hard-of-hearing. Improvements in speech recognition quality and punctuation will directly benefit this and enable us to scale these subtitling efforts to more regional or even the federal parliament. Fraunhofer will exploit the transfer learning techniques developed in this project to scale to more languages as well as different branches of industry (e.g. banking, call center, health industry), a scaling which was previously not possible due to lack of sufficient industry specific training data.

IMCS	IMCS as the integrator of the SELMA project will continue to develop the technologies conceived in the project and will disseminate by making them commercially available for the broader user group. IMCS will also exploit project results by implementing them in the LETA news agency multilingual (Latvian, English, Russian media) video news production workflows and for joint knowledge graph and large language model technology development with the PiniTree.com startup. IMCS will also use the SELMA voice-over system to extend the media monitoring and storyline summarization system built in SUMMA to automate the generation of video from the news items and synthetic narration of the summary.
Priberam	Priberam is currently productizing the results of SUMMA in MONITIO and news.bridge in Plain X and as such plans to incorporate the results of SELMA into those new offerings. Priberam will also incorporate results from SELMA in its line of NLP SaaS products, which are sold directly and through partners in Portugal, Spain and Brazil to clients which include the biggest media producers. Together with IMCS, Priberam intends to support the platform and its derived products giving commercial support to the open-source parts and reaching agreements with the other parties on the commercial use of specific components. This renewed product line and its case studies will be presented in international roadshows as well as conferences in areas covered by the project. Priberam will disseminate the results across media groups and media monitoring target user companies in European countries, reaching also organizations from Latin America and Africa.

Table 5 Initial Exploitation Plans

4.3   Expected Technology Impacts

SELMA's output is expected to be an open-source platform including improved research and tools for the various NLP technologies.

The following table sums up the expected technology /components achievements as listed in the initial DoA.

Technology / Component	TRL (2020)	Expected TRL (2022)	Status M12
Punctuation Recovery	5	7	as planned
Speaker Diarization	6	8	as planned
Speaker Recognition	6	8	as planned
Rich Automatic Speech Recognition (including named entities)	6	8*	as planned

Text Machine Translation	7	9*	as planned
Expressive and Personalized Voice Synthesis	5	7	as planned
Speech Machine Translation	5	7	as planned
Topic Labeling (from crosslingual transfer)	4	6	as planned
Named Entity Recognition and Linking	6	8	as planned
Abstractive Summarization	3	6	as planned
Integration Platform (NLP Components and UX)	6	8	as planned
Integration Platform (Learning/Training of NLP and Automatic Redeployment)	3	7	as planned
*depending on the target languages and language pairs			

Table 6 Expected Technology / Components Improvements Overview

4.4 IPR Management

For a clear IPR Management each component will be broken down into a set of descriptors:

- Component name
- Inputs from
- Outputs to

- Component lead partner

- Component contributors

- Brief description

- What it does (more detail)

- How it works (more detail)

- Key innovative aspects

- Potential applications

- Software & IPR status

- Terms & conditions of use

- Performance requirements

- Further documentation

- Alternatives

- Key contact(s)

- Potential applications of any components aside from SELMA

- The names and details of any libraries used within components and their IPR status

- Answering key questions developers wishing to exploit the component are likely to have such as: How many users can be supported? What specific machines are required to run the code?

- The name of and links to alternative open source components if a component will not be part of an open source release

This work will start in the first half of 2022.

5.Conclusion & Outlook

Creating impact through a coherent communication, dissemination and exploitation strategy is a central building block throughout the project's life cycle.

There are a few factors that define a successful strategy. Besides drawing attention and inspire through good communication means and by being visible on (virtual) events, workshops and fairs, it is very important to convince with a good product: with modules and components which have an added value and can be integrated smoothly into a given NLP landscape.

Collecting feedback from an early stage of the research and development process will help to improve the SELMA project and its outcome. That is where the user and advisory board will play a vital role.

This is the first of three iterations; the interim update is due after the second year (M24) and the final version is due at the end of the project (M36).