Cost effective, Multilingual, Privacy-driven voice-enabled Services

www.compriseh2020.eu

Call: H2020-ICT-2018-2020
Topic: ICT-29-2018
Type of action: RIA
Grant agreement Nº: 825081

WP Nº7: Communication, dissemination and exploitation

Deliverable Nº7.3: First dissemination and communication report

Lead partner: ROOT

Version Nº: 1.0

Date: 28/04/2020

Document information	
Deliverable Nº and title:	D7.3 – First dissemination and communication report
Version Nº:	1.0
Lead beneficiary:	ROOT
Author(s):	Álvaro Moretón, Conrado Castillo, Laura Merlo (ROOT)
Reviewers:	Raivis Skadiņš (TILDE), Aurélien Bellet (INRIA)
Submission date:	28/04/2020
Due date:	30/04/2020
Type[1]:	R
Dissemination level[2]:	PU

Document history			
Date	Version	Author(s)	Comments
08/04/2020	0.1	Álvaro Moretón, Conrado Castillo, Laura Merlo	First draft of the deliverable
21/04/2020	0.2	Álvaro Moretón, Laura Merlo	Deliverable after the first review
28/04/2020	1.0	Zaineb Chelly & Emmanuel Vincent	Final version revised by the project manager and the coordinator

[1] **R**: Report, **DEC:** Websites, patent filling, videos; **DEM:** Demonstrator, pilot, prototype; **ORDP:** Open Research Data Pilot; **ETHICS:** Ethics requirement. **OTHER:** Software Tools

[2] **PU:** Public**; CO:** Confidential, only for members of the consortium (including the Commission Services)

Document summary

This deliverable is related to Task 7.3 “Dissemination and collaboration” and Task 7.4 “Communication activities” of Work Package 7 “Communication, dissemination, and exploitation” of COMPRISE. The main objective of these tasks is to maximise the impact of the project’s results worldwide through the execution of dissemination and communication activities.

The purpose of this document is to provide an update on the dissemination and communication actions, their goals, and the activities carried out to fulfil them. Actions to be taken to reach a wider audience will be indicated as well.

This document will also describe the communication and dissemination activities carried out during the first period of the project (more specifically from month 1 to month 16 – from December 2018 to March 2020), as well as the results obtained as a consequence of these actions. It will also include the communication materials and the list of contents already planned to be generated during the upcoming months.

This document will be further updated during the course of the project, and will result in the submission of Deliverable 7.5 “Second dissemination and communication report” by month 35 (October 2021).

Table of contents

1. Introduction

To guarantee highly visible and fruitful dissemination and exploitation actions, COMPRISE has a specific Work Package (WP), namely WP7 "Communication, dissemination and exploitation", dedicated to communication, dissemination and exploitation activities. The main objective of this WP is to define communication, dissemination and exploitation activities to make COMPRISE a reference crossroad on the broad topic of software engineering for multilingual voice-enabled applications.

In month 3 (February 2019), Deliverable 7.1 "Dissemination and communication action plan"[3], which aimed to define the dissemination and communication activities to be carried out during the project to maximise its impact worldwide, was submitted.

The current deliverable, D7.3 "First dissemination and communication report", is the third deliverable of WP7 and is directly related to Task 7.3 "Dissemination and collaboration" and Task 7.4 "Communication activities".

The purpose of this document is to describe the dissemination and communication activities carried out within COMPRISE from the start of the project (December 2018) until the end of month 16 (March 2020), as well as the results achieved. Additionally, it will include the planning and description of the activities that will be carried out in the upcoming months.

2. Targeted stakeholders

The dissemination and communication activities described in this document aim to maximise the impact of the project's results amongst stakeholders. The main groups of stakeholders targeted by COMPRISE as well as the activities and channels that should be used to reach them were identified in D7.1.

As reflected in Table 1, the stakeholders that have contacted the Consortium partners have shown different grades of interest in COMPRISE. Some of them were interested in obtaining additional information on the project; while others have shown interest in exploiting the project's results or use products based on COMPRISE technologies.

Table 1: *Stakeholders.*

Type of stakeholder	Has shown interest in COMPRISE (e.g., requesting information about the project)	Have expressed interest in exploiting the results of COMPRISE (e.g., develop products based on COMPRISE technologies, carry out additional research based on COMPRISE's results)
SMEs	Llorente y Cuenca, Aflorithmic, Modality.AI, Norsys, Vox.care, ZIroh Labs	i4FS, DRS, Assistentis, ProMed, Mediaskopas

[3] https://www.compriseh2020.eu/files/2019/04/D7.1-1.pdf

Large companies	Google, BBVA, Mahou, Marktel, Intel Corporation, Tet	
Medical service organizations	Ministry of Healthcare of Latvia, Ministry of Healthcare of Estonia	Children's Clinical University Hospital (COMPRISE EAB member)
Researchers	University of Wisconsin-Madison	
EU Projects	SAAM	ZDMP, ELITR, ELG
Media (e.g., magazines, journals, television, radio, etc.)	Diario Economista, Diario EL país, Latvian State TV and Radio, Delfi (EE, LV, LT), Franc e3, L'Usine Nouvelle, Rue89 Strasbourg	EURECOM, LIA, NII, registered participants of the VoicePrivacy challenge (17 academic or industrial teams from China, France, India, Japan, Spain, Switzerland, Turkey, UK, USA), KU Leuven, Karlstad University
Event organizer	VoiceTech Paris	
Policymakers	Garante per la protezione dei dati personali, EDPB	CNIL
Others	SPEAKER (Fraunhofer IIS)	Almond, SPEAKER

3. Dissemination and communication activities

As indicated in D7.1 "Dissemination and communication action plan", from the beginning of the project, several communication and dissemination actions have been carried out to publicise the COMPRISE project's activities and results and to maximize its impact. All communication and dissemination actions as well as their results have been tracked along the project, and the progress achieved through each of them is described in the following sub-sections.

3.1. Hunting

We have initiated a monitoring action to locate forums related to the project activities (voice recognition systems, data privacy, etc.), where partners can participate actively, and promote the project and its results. Hunting actions will also facilitate networking with companies, researchers and professionals that work in similar fields. The hunting actions that we have carried out are detailed as follows:

3.1.1. Online targeting of stakeholders

Different stakeholder groups have been targeted and followed or contacted through social media to draw their attention towards our project. Many of these stakeholders have been identified through the COMPRISE Twitter account. The main profiles targeted include, among others, app developers (especially those that work with voice technologies), SMEs developing voice-enabled solutions, researchers, other European projects, and privacy experts. The list of profiles followed by COMPRISE on Twitter can be found at https://twitter.com/compriseh2020.

COMPRISE's Twitter is updated periodically (new profiles are followed, while those with poor or no interactions are unfollowed after some time) to focus on the most interesting

profiles at each stage of the project, as well as to maintain an active image of the account (e.g., new followers, interactions, posts, etc.).

Additionally, two separate campaigns were initiated on LinkedIn by Vikki Academy (see Section 3.7.1 for more information about Vikki Academy) on February 2020, targeting:

- Policy makers: European and overseas data protection authorities.
- Tech companies: C-level (CTO, CPO & CMO) personnel of medium to large companies developing voice-driven apps or marketing such apps to customers. Specific sectors included e-commerce (online retailers: CTO, CPO & CMO) and e-health (hospitals: CPO).

According to the summary metrics provided by Vikki Academy, 30% of all policy makers who were invited were interested in COMPRISE technologies (i.e., 125 of them accepted the invitation) with 25% of them who responded to the initial e-message that they have received, i.e., an invitation e-message that was prepared and sent on the name of the coordinator from his personal account. Additionally, 36% of all Tech company personnel who were invited were interested in COMPRISE technologies (i.e., 171 of them accepted the invitation) with 37% of them who responded to the initial e-message.

3.1.2. Discussion groups

Rooter has searched for and joined several discussion groups related to the project's activities (e.g., privacy, AI, software developers, voice technologies) on LinkedIn, to disseminate the publications and results achieved so far. Table 2 shows the different discussion groups in which Rooter is disseminating the various COMPRISE materials (e.g., deliverables, videos, blogs, news, etc.).

Table 2: Discussion groups.

Group name	Number of members
Voice AI - Amazon Alexa, Google Assistant, Apple Siri, Samsung Bixby, AI & More	1,417
Voice: We talk voice apps, conversational AI, VUI, VX, Amazon Alexa, Google Assistant and all more	28
Voice in healthcare alliance	67
Chatbots, Virtual Assistants, Intelligent Agents, Conversational AI	3,612
Interactive Voice Response	1,379
Analytics and Artificial Intelligence (AI) in Marketing and Retail	121,662
International Association of Privacy Professionals (IAPP)	27,238
EU Data Protection Regulation	3,305
Data Protection and the EU	14,112
Cyber Privacy Security Commercial	7,465
European Data Protection Forum EDPF	10,006
GDPR Technology	19,035
Privacy and Data Security	12,866
Privacy Professionals	14,532
Freelance Mobile App Developer	11,755

Mobile App Development IOS, Android, Windows Developers	6,225
iOS App Developers Worldwide	22,794

3.1.3. Event monitoring

Hunting activities involve spotting events (e.g., conferences, congresses, fairs, workshops, etc.) where the project results can be disseminated. A Word document has been created and shared with the project partners, where they can contribute by adding events they are aware of, and that can be considered for future dissemination actions. Table 3 lists the events identified in 2019.

Table 3: 2019 events.

Event name	Link	Date and Location
PRIVACY		
5th IEEE Workshop on Security and Privacy in the Cloud	https://cns2019.ieee-cns.org/workshop/spc-5th-ieee-workshop-security-and-privacy-cloud-2019	10-12 June, Washington DC
10th French Privacy and Data Protection workshop	https://project.inria.fr/apvp2019/	9-11 July, Saint-Valéry-sur-Somme
Data Protection World Forum	https://www.dataprotectionworldforum.com/	11-12 July, Dublin
16th International Conference on Trust, Privacy and Security in Digital Business	http://www.dexa.org/trustbus2019	26-29 August, Linz
Dagstuhl Seminars	https://www.dagstuhl.de/en/about-dagstuhl/	TBD
Data Protection Commissioners 41st International Conference	https://privacycalendar.org/event/data-protection-commissioners-41st-international-conference/	13 October, Tirana
IAPP Europe Data Protection Congress	https://iapp.org/conference/iapp-europe-data-protection-congress/	20-21 November, Brussels
AI /VOICE TECHNOLOGIES		
BDV PPP Meetup	http://www.bdva.eu/node/1217	26-28 June Riga
Multilingualism at the intersection of Knowledge Bases and Machine Translation	https://moment2019.insight-centre.org/	19 August, Dublin
1st International Conference on the European Industry on Language Technology	http://www.pret-a-llod.eu/else-if19/	9 September, Karlsruhe
Digital Excellence Forum @ ICT Proposers' Days	https://ec.europa.eu/digital-single-market/en/news/digital-excellence-forum-ict-proposers-day-2019	19-20 September, Helsinki

Deutscher EDV Gerichtstag	https://www.edvgt.de/es/veranstaltungen/deutscher-edv-gerichtstag/2019-2/	20 September, Saarbrücken
Meta Forum	https://www.european-language-grid.eu/meta-forum-2019/	8-9 October, Brussels
Global forum on AI for Humanity	https://gfaih.org/	28-30 October, Paris
Voice Tech Paris	https://www.voicetechparis.com/2019/	26-27 November, Paris

Table 4 lists the events that have been identified for 2020.

Table 4: 2020 events.

Event name	Link	Date and Location
AI / VOICE TECHNOLOGIES		
CPDP 2020: Data Protection and Artificial Intelligence	https://www.cpdpconferences.org/	22-24 January, Brussels
First European Open Virtual Assistant Workshop	https://oval.hipeac.net/2020/#/	20 March, Palaiseau
Shonan Meeting	https://shonan.nii.ac.jp/seminars/	23-27 March, Kanagawa
Second Workshop on e-Commerce and NLP	https://sites.google.com/view/ecnlp/www-2020	20 April, Taipei
First International Workshop on Language Technology Platforms	https://www.european-language-grid.eu/iwltp-2020/	16 May, Marseille
APPS DEVELOPERS		
Mobile World Congress	https://www.mwcbarcelona.com/	23-28 February, Barcelona
Codemobile	https://www.codemobile.co.uk/	28-29 April, London
App Builders	https://appbuilders.ch/	10-12 May, Lugano
Applause	https://applause.io/	22 May, Barcelona
Reacteurope	https://www.react-europe.org/	14-15 May, Paris
We are developers	https://www.wearedevelopers.com/events/world-congress/	28-29 May, Berlin
Targetsummit	http://targetsummit.com/	Several events
Droidcon	https://www.droidcon.com/	Several events
Mobile Tech Conference	https://mobiletechcon.de/de/	26-28 October, München

App Growth Awards	https://www.businessofapps.com/event/app-growth-awards/	3 December, Berlin
PRIVACY		
CPDP Conference	https://www.cpdpconferences.org/	22-24 January, Brussels
The regulation of new technologies	https://www.tilburguniversity.edu/research/institutes-and-research-groups/tilt/call-papers-regulation-new-technologies	11 June, Tilburg
Second Workshop on Security in Machine Learning and its Applications	http://dsp-lab.eng.uci.edu/simla/	19-22 October, Rome
34th Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy	https://dbsec2020.ur.de/cfp	25-27 June, Regensburg
5th IEEE European Symposium on Security and Privacy	https://www.ieee-security.org/TC/EuroSP2020/cfp.html	7-11 September, Genova
Data for Policy	https://dataforpolicy.org/data-for-policy-2020/	15-16 September, London
IAPP Europe Data Protection Congress	https://iapp.org/conference/iapp-europe-data-protection-congress/	16-17 November, Brussels
Call for papers (several events)	http://www.wikicfp.com/cfp/call?conference=privacy	Several

The events in Table 4 were expected to take place during 2020. However, due to the Covid-19 pandemic, their status is uncertain. Many of them have been cancelled, while others have been delayed, with new dates yet to be confirmed.

3.1.4. Journals and magazines

A Word document has been created and shared with the project partners so they can contribute by adding specialised magazines, journals or other publications they have knowledge of, where papers and articles related to the project's results can be published, with a special focus on privacy and data protection. Table 5 lists the journals currently identified.

Table 5: Privacy and data protection journals and magazines.

Publication	Link
Revista SIC	https://revistasic.es/
Revista de privacidad y derecho digital	https://www.rdu.es/articulos/revista/RPDD
Privacy & Data Protection Journals	https://www.pdpjournals.com/overview-privacy-and-data-protection
Journal Information Privacy & Security	https://www.tandfonline.com/toc/uips20/current
Revista Byte	https://www.revistabyte.es/

Revista CPO	https://www.cpomagazine.com/
Innotech today	https://innotechtoday.com/
International Journal of Law and International Technology	https://academic.oup.com/ijlit
Legal IT Professionals	https://www.https://www.lawtechnologytoday.org/contributors/leghttps://www.tandfonline.com/loi/uips20alitprofessionals.com/
Technology and Society Magazine	https://technologyandsociety.org/technology-and-society-magazine/
Journal of Information Privacy and Society	https://www.tandfonline.com/loi/uips20
Legal Technology Today	https://www.lawtechnologytoday.org/
International Association of Privacy	https://iapp.org/
Brussels Privacy Hub	https://brusselsprivacyhub.eu/publications/
Journal of Data Protection and Privacy	https://www.henrystewartpublications.com/jdpp
International journal for the Data Protection Officer	https://idpp.info/submissions/
Journal of Privacy and Confidentiality	https://journalprivacyconfidentiality.org/index.php/jpc
Ingenta	https://www.ingentaconnect.com/content/hsp/jdpp
International Journal of Information Privacy, Security and Integrity	https://www.inderscience.com/jhome.php?jcode=ijipsi
EDPL	https://edpl.lexxion.eu/

3.1.5. Cooperation with other websites

COMPRISE reached a cooperation agreement with the Association of Understanding of Artificial Intelligence for disseminating COMPRISE contents through their website AIHub[4]. This will help COMPRISE to reach a bigger audience and to bring visits to the project's website.

3.2. Joint actions with other H2020 projects

Regarding COMPRISE cooperation with other H2020 projects, ELITR[5] has expressed interest in COMPRISE privacy-driven speech and text transformation tools to process private meeting data prior to its public release. The COMPRISE partners have shared the first version of the project's tools with the ELITR team in September 2019.

The COMPRISE partners have also setup a collaboration with ELG[6]. ELG will provide a webpage to demonstrate some of the COMPRISE tools which may be run on the ELG platform. This integration will be done after the ELG platform is officially released. It is

[4] https://aihub.org/about/
[5] https://elitr.eu/the-project/
[6] https://www.european-language-grid.eu/

also worth mentioning that COMPRISE submitted a paper to the International Workshop on Language Technology Platforms organised by ELG.

3.3. Events

To ensure successful dissemination of the COMPRISE results, the consortium partners have participated in various events (17 in total) during the first period of the project, both as organisers and speakers, which are listed in Table 6. These events included, among others, conferences and workshops.

Table 6: Events.

Date	Name	Type
December 2018	ICT 2018 networking session	Activities jointly organized with other H2020 projects
March 2019	AI and Language for Citizens	Workshop
April 2019	Panel discussion “Positioning at the European level, an opportunity to be seized”, General Assembly of the Materalia Competitiveness Cluster	Other events
April 2019	Showcase H2020 and INEA	Activities jointly organized with other H2020 projects
June 2019	BDV PPP Summit 2019	Conferences
June 2019	MeetUp Intelligence Artificielle Nancy #4	Other events
July 2019	10th French Privacy and Data Protection workshop	Workshop
September 2019	Interspeech 2019	Conferences
September 2019	ICT Proposers' Days	Conferences
October 2019	META-FORUM 2019	Conferences
October 2019	Global Forum on AI for Humanity	Conferences
October 2019	Workshop for Children's hospital	Workshop
October 2019	Science Festival of University of Lorraine	Other events
November 2019	Voice Tech Paris 2019	Conferences
November 2019	Theory and Practice of Differential Privacy	Workshop
November 2019	Privacy Preserving Machine Learning	Workshop
November 2019	Panel discussion organised by Rue89 Strasbourg	Other events

The estimated number of persons reached through the above events is 1,709. Table 7 categorises the people reached.

Table 7: People reached.

Category	Number of people reached
Scientific Community	732
Industry	355

Civil Society	91
General Public	224
Policy Makers	96
Media	86
Investors	20
Customers	60
Others	45

3.4. Publications

The main purpose of this action is to publish papers and articles concerning the COMPRISE results in international conferences and workshops, scientific journals or specialised magazines. Until month 16, four papers have been published or accepted for publication, as listed in Table 8.

Table 8: Published or accepted publications.

Paper title	Conference or Workshop
"Privacy-Preserving Adversarial Representation Learning in ASR: Reality or Illusion?"	Interspeech 2019
"Evaluating Voice Conversion-based Privacy Protection against Informed Attackers"	2020 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
"The COMPRISE Cloud Platform"	1st International Workshop on Language Technology Platforms
"Private Protocols for U-Statistics in the Local Model and Beyond"	23rd International Conference on Artificial Intelligence and Statistics (AISTATS)

These papers can be found on the COMPRISE website[7]. Several more papers have been submitted to the 37th International Conference on Machine Learning, Interspeech 2020, the IEEE International Conference on Distributed Computing Systems, and the European Data Protection Law Review.

3.5. Website

The COMPRISE website is one of the most important channels for external communication. It acts as a central dissemination and communication portal, and as a repository of publications, articles, software components and other contents related to the project's activities.

[7] https://www.compriseh2020.eu/papers/

The website (a WordPress site) is run by INRIA, with ROOT's support. All partners participate by providing contents to enrich it.

3.5.1. Website sections

The website is divided into six different sections:

- **About the project:** In this section, visitors can find a brief description of the project.

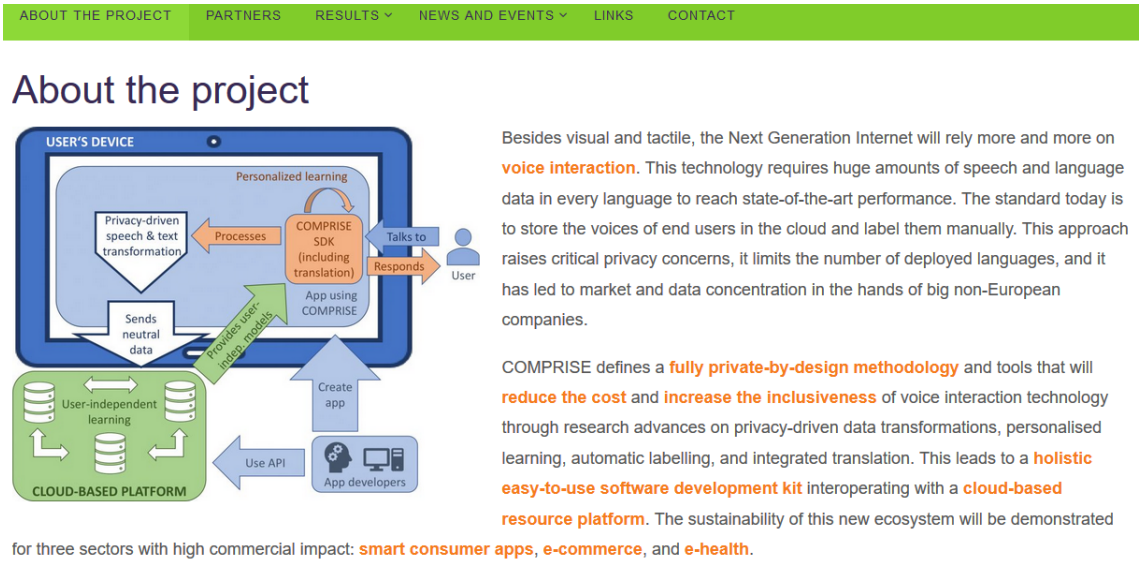

Figure 1: About the project

- **Partners**: In this section, visitors can find information on the organisations participating in the project as partners. Information on the different functions the members of each organisation have in relation to the project can be found in this section as well.

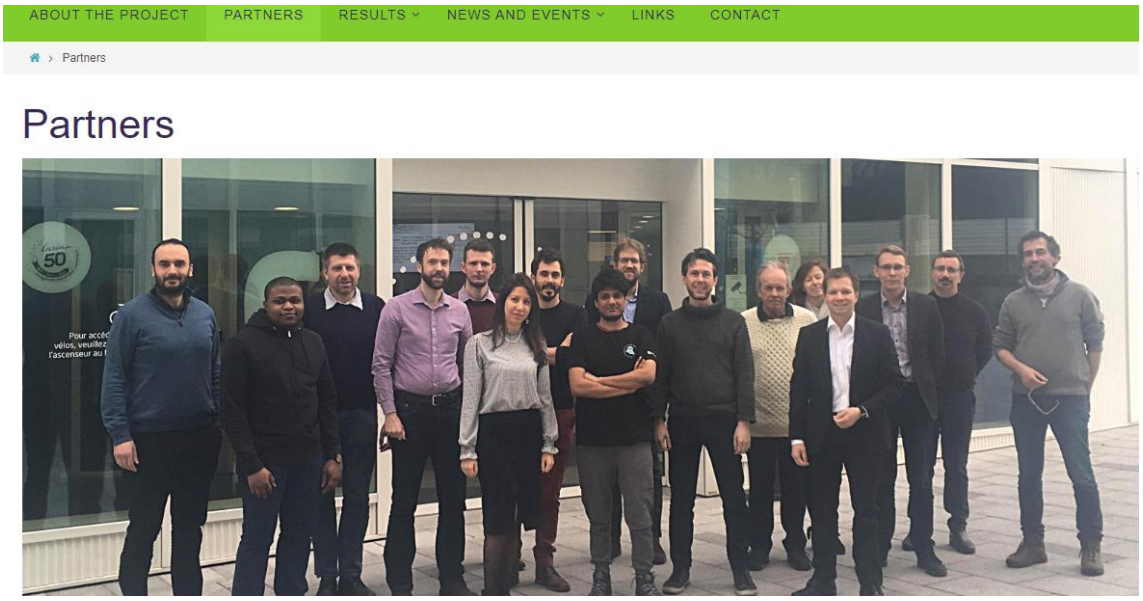

Figure 2: Partners

- **Results.** This section is divided into three subsections:
    - **Deliverables:** Visitors can access and download public deliverables that have already been uploaded into this section. There are 8 deliverables available so far.

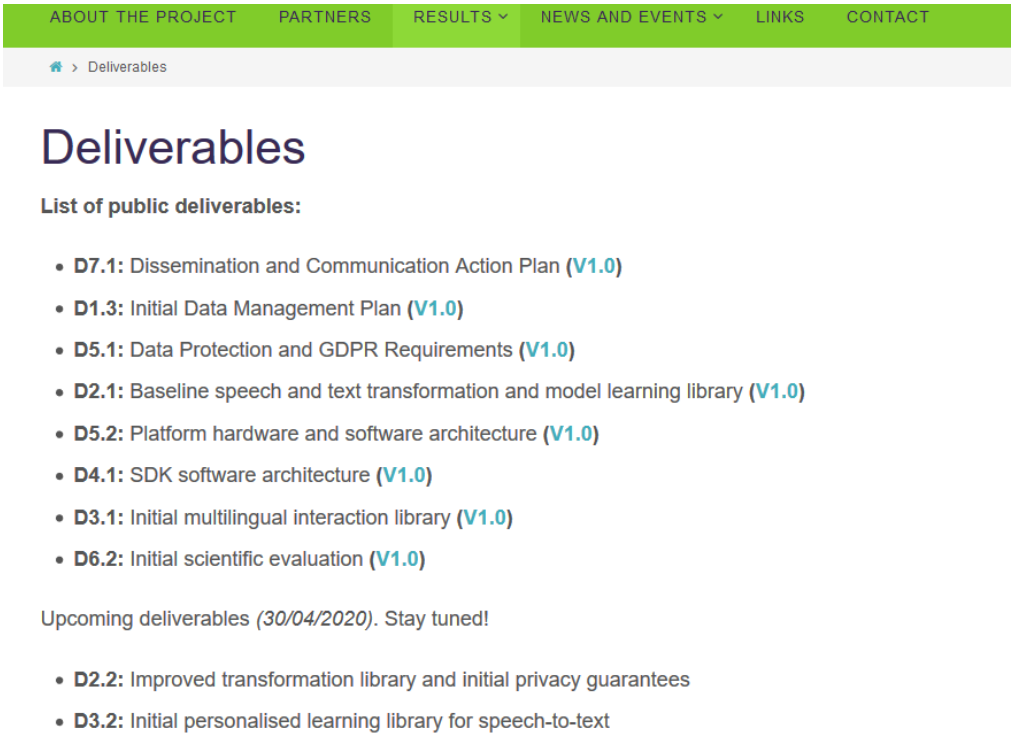

Figure 3: Deliverables

    - **Software:** Visitors can access the Gitlab repository where the public COMPRISE results software libraries can be found.

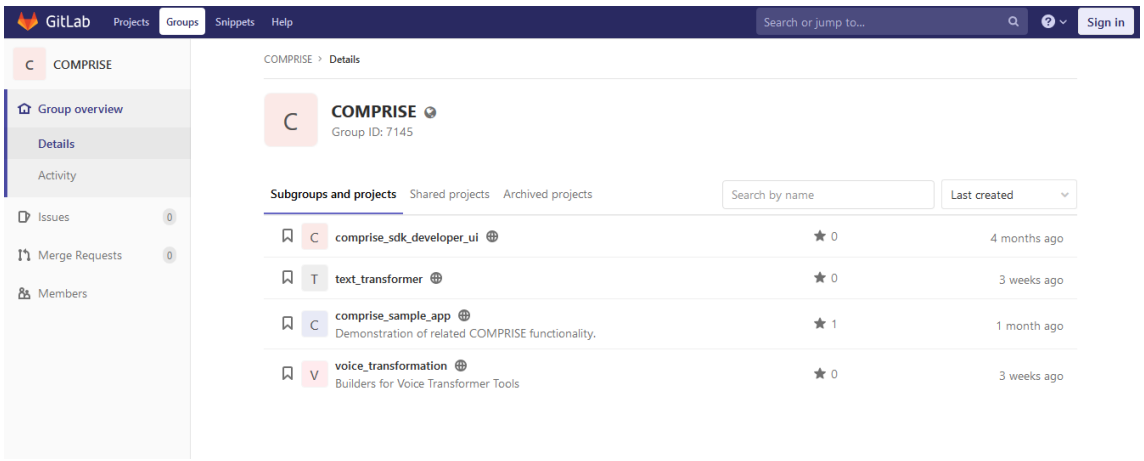

Figure 4: Software

    - **Papers:** In this section, visitors can access project-related papers.

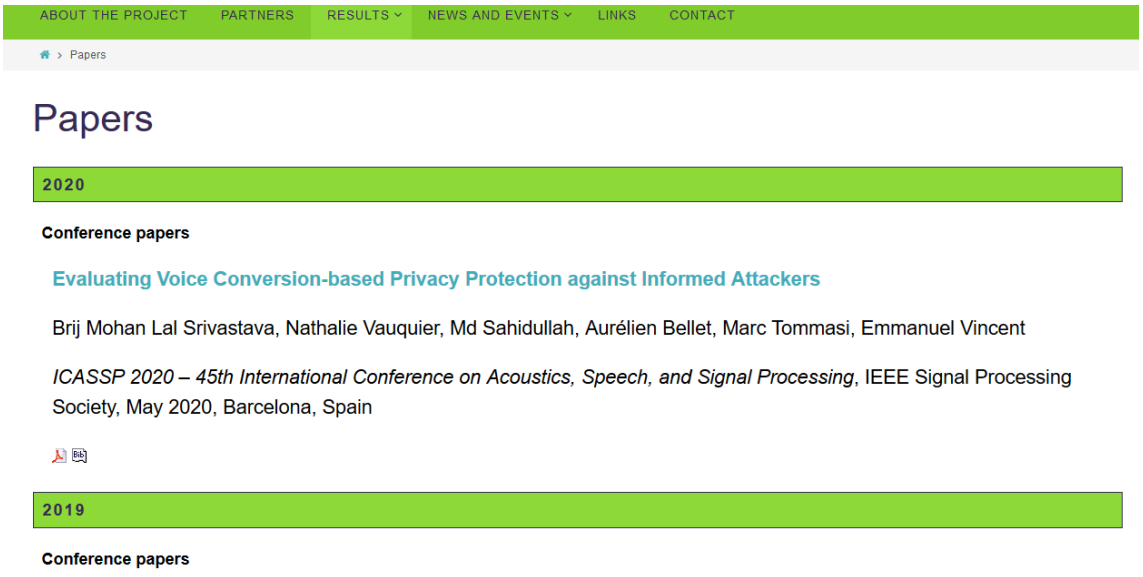

Figure 5: Papers

- **News and events.** This section is divided into four subsections:
    - **Project news:** In this section are published all news related to the project's progress.

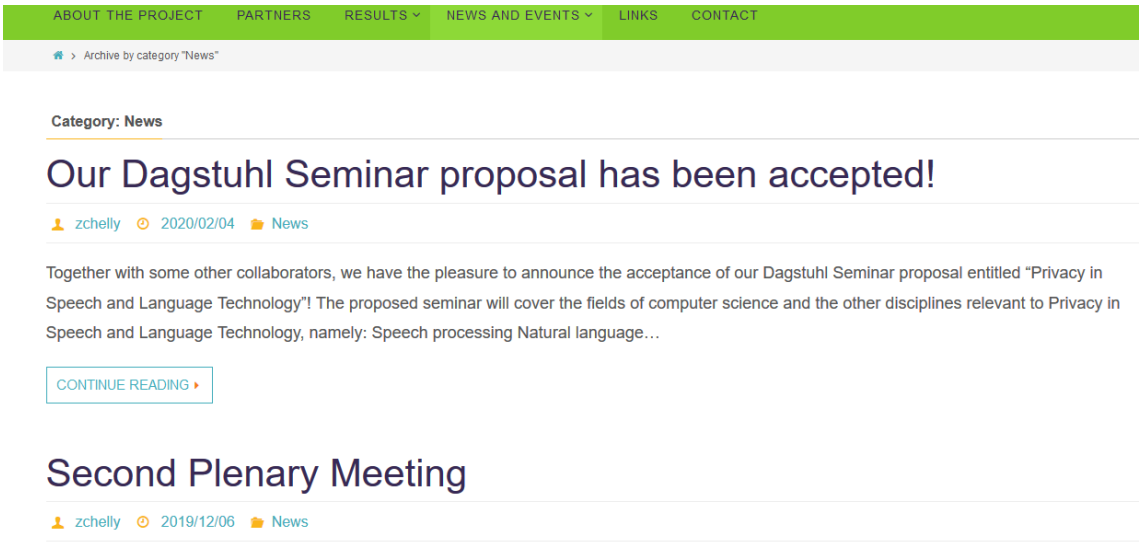

Figure 6: Project news

    - **Events:** In this section are published all news related to events in which partners have participated or will participate in disseminating the project.

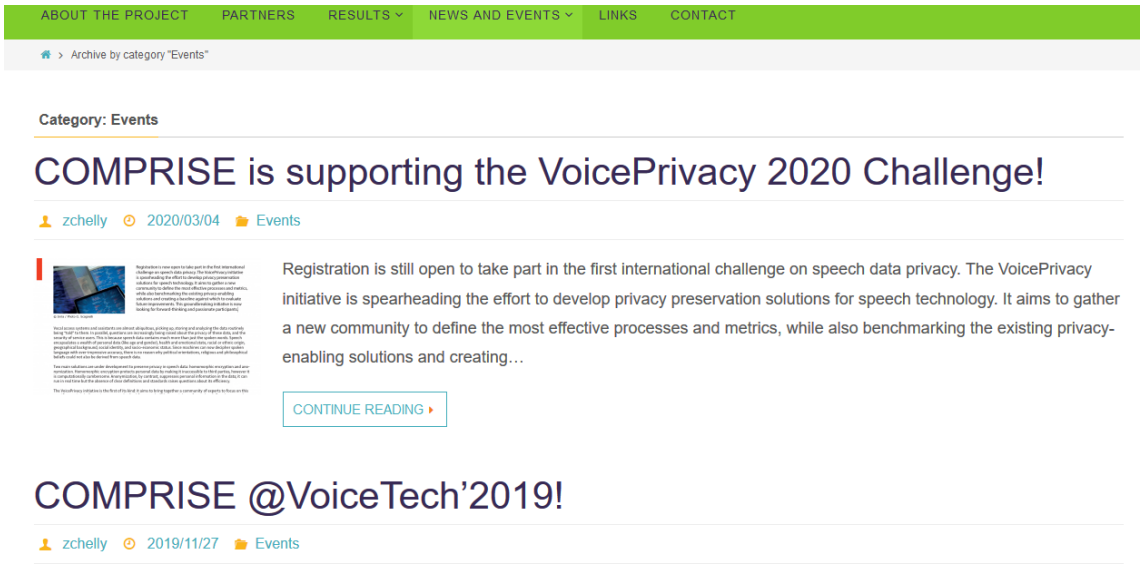

Figure 7: Events

- **Blog:** In this section are published all blog entries and infographics related to COMPRISE activities. All project partners are involved in the generation of blog entries. There are currently 15 blog entries published so far.

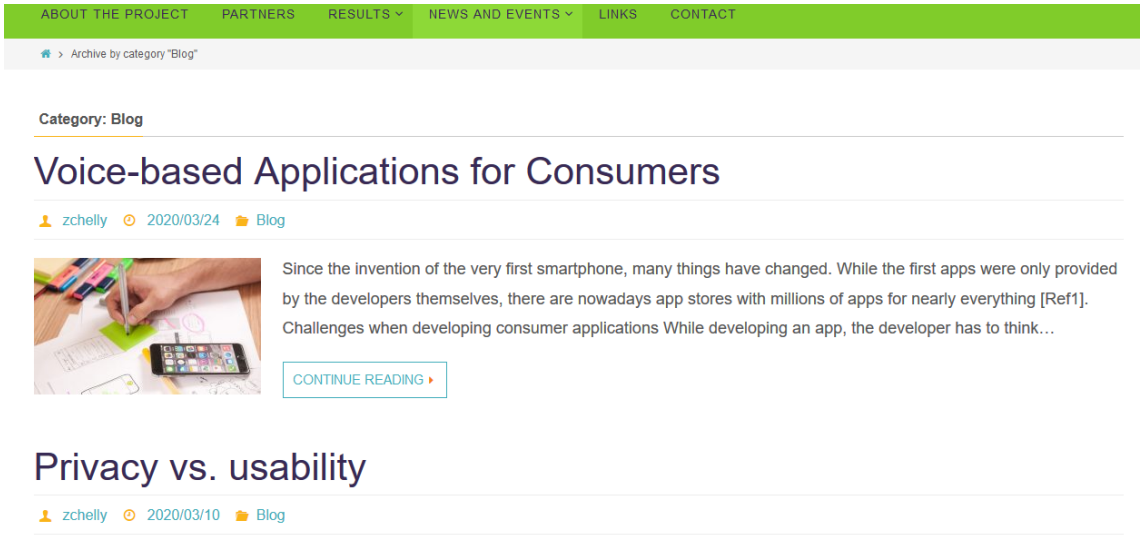

Figure 8: Blog

- **Videos:** In this section, visitors can access all project-related videos. There are currently 7 videos (5 videos produced by COMPRISE, and 2 videos from different media channels where COMPRISE partners have participated).

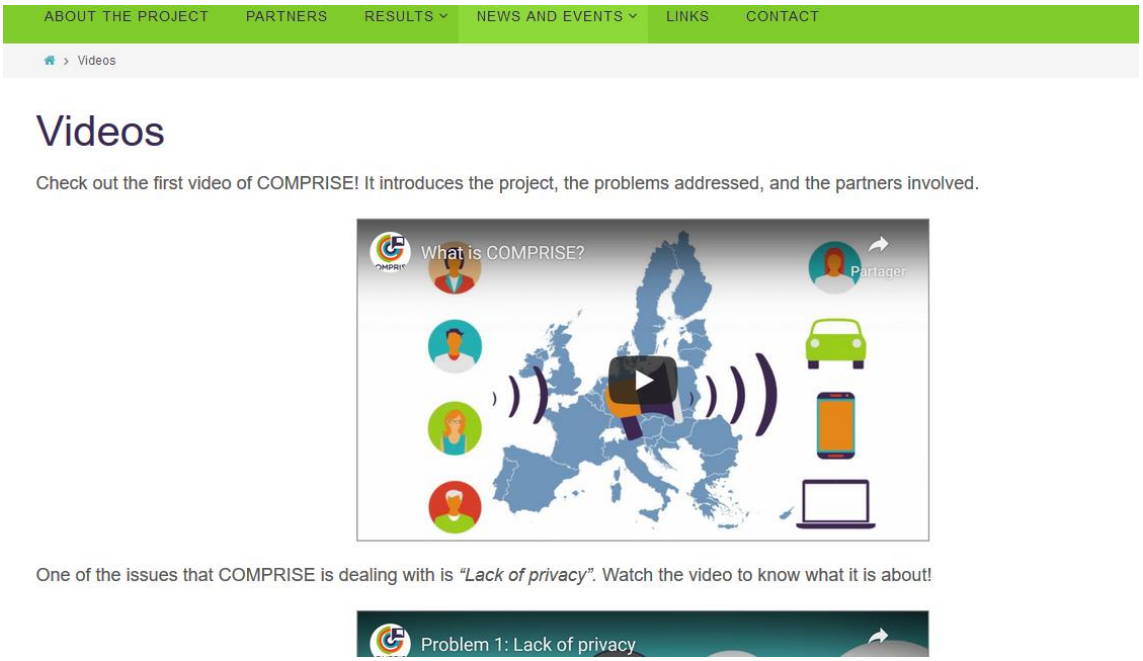

Figure 9: Videos

- **Links.** In this section are included the links to the websites of H2020 projects related to COMPRISE and the websites of the External Advisory Board members.

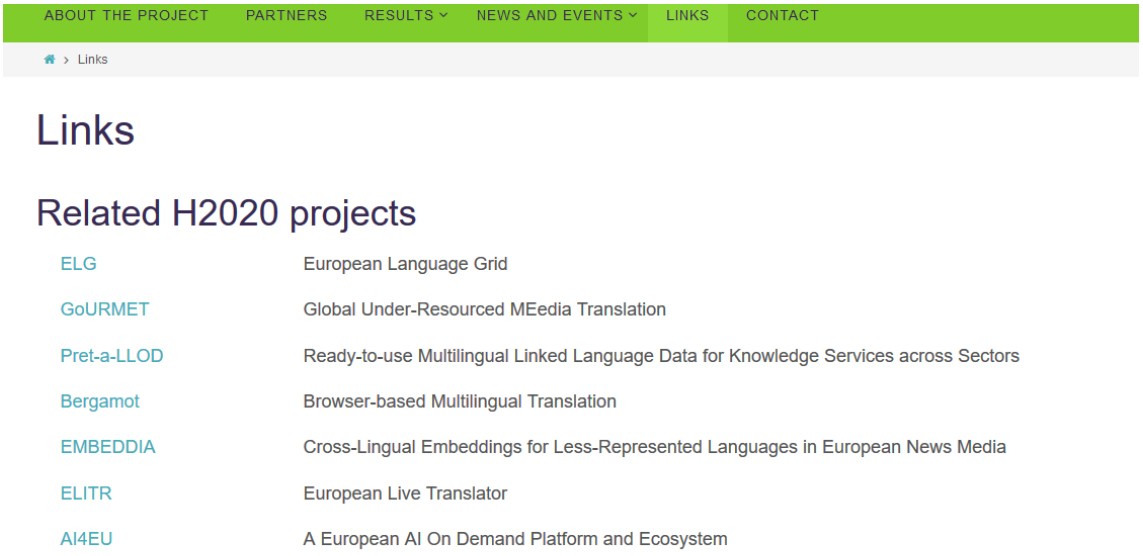

Figure 10: Links

- **Contact:** This section covers the contact details of the COMPRISE project manager as well as links to the COMPRISE social networks.

Figure 11: Contact

3.5.2. Website audience

To analyse the trails left by website visitors we are using Matomo, an open-source web analytics tool, which protects privacy by anonymizing IP addresses. Table 9 lists the visit statistics until March 31, 2020.

Table 9: Overview of website visits.

Overview of visits	
Visits	3,527
Unique visitors	1,708
Average length of visit	3 min 3 s
Bounced visits (left the site after viewing a single page)	50%
Visitor actions (page views, downloads, outgoing links and internal searches within the site)	3.3
Page views	9,843
Unique page views	7,025
Downloads	652
Unique downloads	590
Outbound links	979
Unique outbound links	854
Maximum number of shares in one visit	78

The number of unique visitors is 1,708. The average time spent on the website is 3 min.

The bounce rate, which is the percentage of visitors who enter the page directly from www.compriseh2020.eu, and who do not interact with it, is 50%. Although this is not an exaggerated percentage, we aim to reduce it in the future.

Another fact worth noting is that each visitor performs 3.3 actions on average before leaving the website (e.g., download of documents, exploring tabs within the page, etc.).

Table 10 contains the metrics for some example pages within the COMPRISE website (e.g., content, categories, where visitors click, etc.).

Table 10: Webpage metrics.

URL of the page	Page views	Unique page views	Bounce rate	Average time on page	Output percentage
Comprise	2,818	1,891	51%	59 s	44%
deliverables	570	469	32%	1 min 28 s	67%
partners	476	387	75%	1 min 12 s	51%
videos	356	265	57%	1 min 53 s	67%
papers	293	109	11%	46 s	48%
results	189	128	13%	20 s	11%
spoken-dialog-systems-are-they-fully-reliable	134	87	58%	2 min 50 s	56%
voice-based-applications-for-e-health	128	108	89%	41 s	87%
links	116	103	65%	55 s	49%
news-and-events	110	87	20%	17 s	3%
personal-data-concept-and-categorisation	107	86	74%	1 min 33 s	72%

According to Table 10, there have been a total of 1,891 unique visits to the COMPRISE homepage. The blog entry “Spoken dialogue systems – are they fully reliable?” has the longest visitor stay on the page, with an average of almost 3 min, and a bounce rate of 58%, which means that, although visitors probably read the whole article, once they have finished, they leave the website.

One of the improvements proposed is to place a search box on the website that allows us to identify what content visitors are looking for and, based on these, make additional improvements if the searches are considered significant.

Matomo also analyses the number of unique downloads on the website. As of March 31, there have been a total of 596 unique downloads. These downloads have been mostly deliverables of the project, but we can also highlight, with more than 10 downloads, two of the infographics uploaded to the website. The documents with the highest number of downloads are:

- D5.1: 149
- D7.1: 100
- D2.1: 75
- D1.3: 44
- Infographic - 10 things you should know about voice technologies: 24
- D5.2: 19

- Infographic - 10 Privacy risks associated with voice technologies: 19
- D4.1: 15

The profiles associated with the most downloaded documents are as follows:

- D5.1 “Data Protection and GDPR Requirements”: Mostly Spanish, American and French visitors. There are also several profiles from Panama and Germany. The number of shares oscillates around 3-5 actions on average for each visitor.
- D7.1 “Dissemination and Communication Action Plan”: Mostly Spanish, French and American visitors. There are also some profiles from Bulgaria and the UK. The number of shares oscillates around 3-5 actions on average for each visitor.
- D2.1 “Baseline speech and text transformation and model learning library”: Mostly Spanish, French and German visitors. The number of shares oscillates around 5-6 shares on average for each visitor.
- D1.3 “Initial Data Management Plan”: Mostly Spanish and French. There are also some profiles from Canada. The number of shares oscillates around 3-5 actions on average for each visitor.
- 10 things you should know about voice technologies: Mostly Spanish and French visitors, although there are some profiles from Ukraine, Austria, South Korea and Australia. The number of shares is quite high, with peaks of up to 42 shares.
- 10 Privacy risks associated with voice technologies: Mostly Spanish and French visitors. The number of shares oscillates around 9-20 on average for each visitor with peaks of up to 35 shares.

The number of actions is higher for infographics because, being inside the blog, visitors must browse to find them. This could be an indicator that there are profiles interested in the extra content provided on the website. In addition, these visitors continue browsing the website checking other contents after downloading or viewing the infographic. Visitors who only look for the project’s deliverables generate fewer actions because they go directly to where that content is, without spending time navigating on the website.

Another range of data concerns the hours and days of the week at which visits, and actions take place within the website.

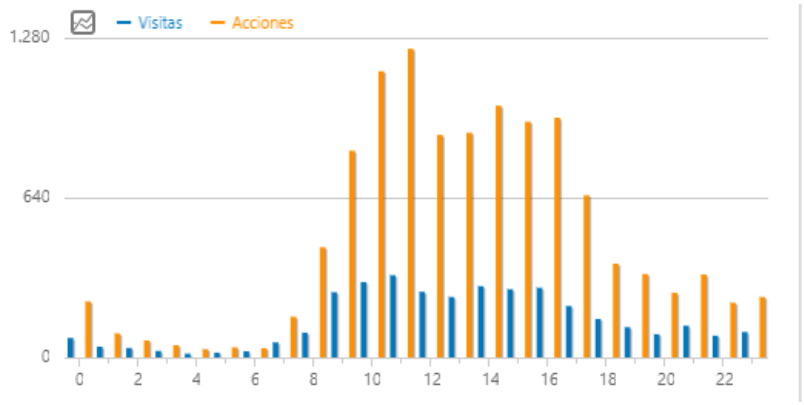

Figure 12: Visits by local time

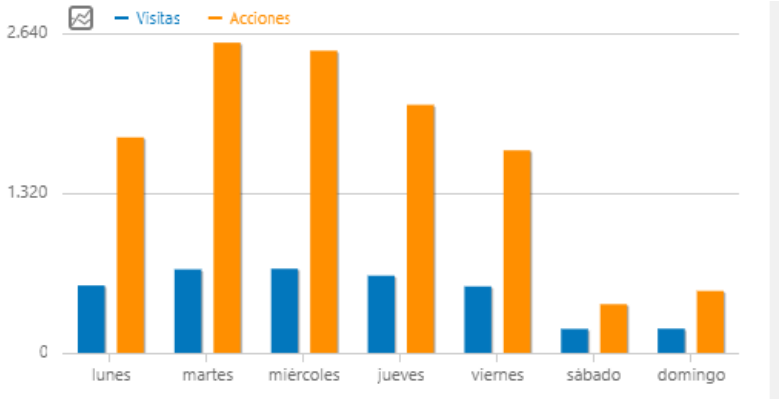

Figure 13: Visits by day of the week

As shown in Figures 12 and 13, the best time to upload content on the website is on Tuesdays and Wednesdays between 10:00 and 12:00 since, on average, more interactions take place at that time. The time between 14:00 and 16:00 is also a good time to upload content, as this is when the highest number of visits is registered.

We have also analysed in Table 11 the statistics of returning visitors compared to new visitors.

Table 11: Returning vs. new visitors.

Statistics	Returning visitors	New visitors
Average visit duration	4 min 41 s	1 min 32 s
Number of actions	4.2	2.4
Bounce rate (left the website after one page)	38%	62%

The average time spent on the page is higher among returning visitors, and the same pattern can be observed in the number of actions, which is almost double. Lastly, the percentage of visitors who leave the website after consulting a single page is considerably lower for those who access the page more frequently.

Finally, Table 12 reports the channels through which visitors enter the COMPRISE website.

Table 12: Channel type.

Channel type	
Direct entries	1,964 (56% of all visits)
From internet sites	406 (12% of all visits)
From search engines	585 (17% of all visits)
From social networks	572 (16% of all visits)
Different search engines	7
Different social networks	6
Different internet pages	63
Different keywords	9

The highest number of visits corresponds to direct entries to the website (56%). 17% of visits have been made via search engines, using keywords such as “COMPRISE

H2020", "videos 2020", etc. Another 16% of visits have been made via social networks. Lastly, 12% of visitors have accessed the COMPRISE website from other websites, including, among others, ROOT's, INRIA's, and ELG's websites.

The statistics for visits made from social networks are provided in Table 13.

Table 23: Social networks.

Social network	Visits	Actions	Actions per visit	Average time on site	Bounce rate
LinkedIn	194	391	2	2 min 15 s	69%
Twitter	188	611	3.3	2 min 42 s	47%
Facebook	162	269	1.7	52 s	83%
YouTube	14	39	2.8	2 min 13 s	43%
Hacker News	11	11	1	-	100%
GitHub	3	3	1	-	100%

Visitors coming from social networks that are not part of the project leave the website sooner, while those coming from the project's own social networks spend more time on the website.

3.6. Social media

Activities carried out on social media aim to attract visitors to the COMPRISE website, disseminate the project's results, contact stakeholders, and inform about the project's progress and events in which the partners participate.

Every week different articles on topics related to the project's activities (e.g., privacy, voice technologies market and functionalities, research advancements on AI and voice technologies) are selected and published in COMPRISE social media accounts to feed them with quality content that attracts new followers. Additionally, the content published by other accounts is monitored and retweeted or shared when considered interesting.

The actions carried out and the results obtained in each of the COMPRISE social networks from month 1 to month 16 are as follows:

3.6.1. Twitter

A total of 174 tweets with 124,376 impressions (i.e., the number of times that COMPRISE tweets have shown up in someone's timeline) have been published through the COMPRISE Twitter account, @COMPRISEH2020 (https://twitter.com/compriseh2020). Additionally, the COMPRISE profile has 160 followers and has been visited 1,707 times.

3.6.2. LinkedIn

A total of 95 posts with 6,254 impressions have been published through the COMPRISE LinkedIn account, COMPRISE H2020 (https://twitter.com/compriseh2020). Additionally, the COMPRISE profile has 132 followers and has been visited 652 times.

3.6.3. YouTube

A YouTube channel has been created to upload the videos produced during the project. The link to the COMPRISE YouTube channel is:

https://www.youtube.com/channel/UCe3sVWWRqYceRgn_sxxT24w.

Currently, there are five videos produced by COMPRISE and that have been uploaded to the COMPRISE YouTube channel (with a total of 749 views). Additionally, two videos from "France 3" and "Rue89 Strasbourg" (with a total of 358 views) can be found in the "Reproduction list" channel in which COMPRISE coordinator Emmanuel Vincent has participated to disseminate the project.

3.7. Generation of content

Several materials have been developed in order to publicise the project. In addition to the materials described in D7.1. "Dissemination and communication action plan" (i.e., the logo, Word and PowerPoint templates), the following materials have been produced:

3.7.1. Videos

The production of at least 15 videos has been planned to disseminate COMPRISE results:

- 1 introduction video
- 3 videos showing the problems that COMPRISE is addressing
- At least 7 small videos showing the new solutions and tools provided. Additional videos such as interviews with the partners and How-To videos for the tools will probably be produced.
- 3 videos for the demonstrators (1 video per demonstrator)
- 1 conclusion video.

Five videos have been produced between month 1 and month 16. Even though the partner responsible for the production of videos is ROOT, it has been decided with the agreement of all partners, that 3 out of 11 videos will be produced by Vikki Academy (http://www.vikkiacademy.com/about/), a media company specialised in the production of animated videos related to academic papers and research activities. Table 14 lists the videos produced to date and their corresponding links.

Table 34: Videos.

Video title	Produced by	Link
What is Comprise?	ROOT	https://www.youtube.com/watch?v=asHD8hBcPPA&feature=youtu.be
Exposed by your own voice	Vikki Academy	https://www.youtube.com/watch?v=gm2cC8za8Us&feature=youtu.be
When voice assistants don't understand	Vikki Academy	https://www.youtube.com/watch?v=-HvADcfEOuE&feature=youtu.be
Why is voice assistants' integration so expensive?	Vikki Academy	https://www.youtube.com/watch?v=5LQb9X3RtUs&feature=youtu.be

How will COMPRISE ensure privacy	ROOT/ INRIA	https://www.youtube.com/watch?v=_LE7XbNPjSs&t=375s

3.7.2. Leaflet

A leaflet has been designed to present the project attractively and understandably. Three hundred units have been printed and handed to the project partners (depending on the number requested by each of them) so they can distribute them at events.

Figure 14. Leaflet outer part

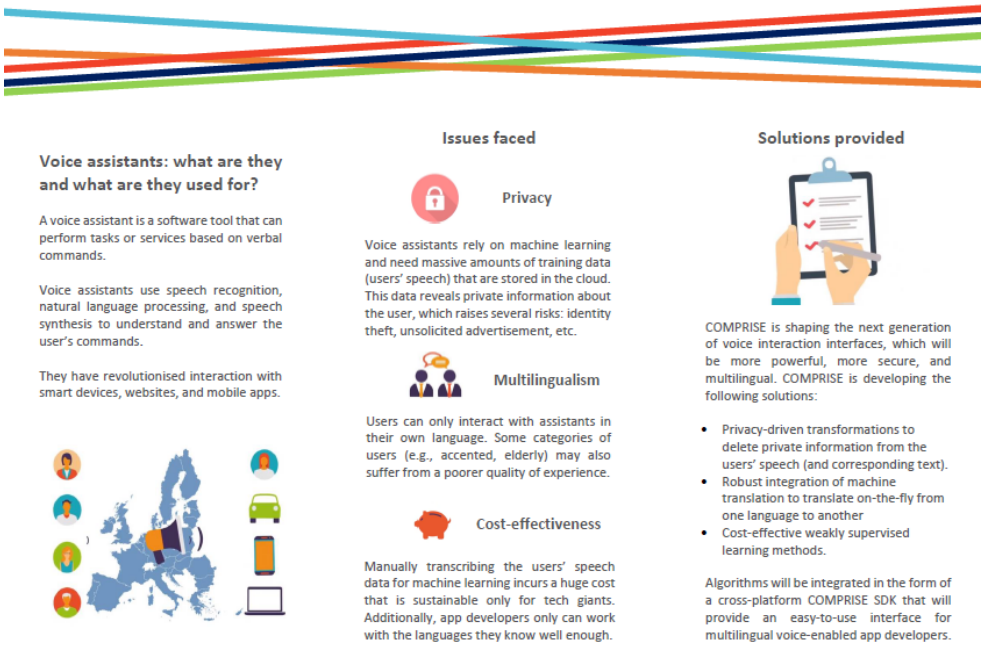

Figure 15: Leaflet inside part

3.7.3. Roll-up

A roll-up design has been created to be used in different events.

Figure 16: Roll-up

3.7.4. Poster

A poster has been designed to disseminate COMPRISE in various events.

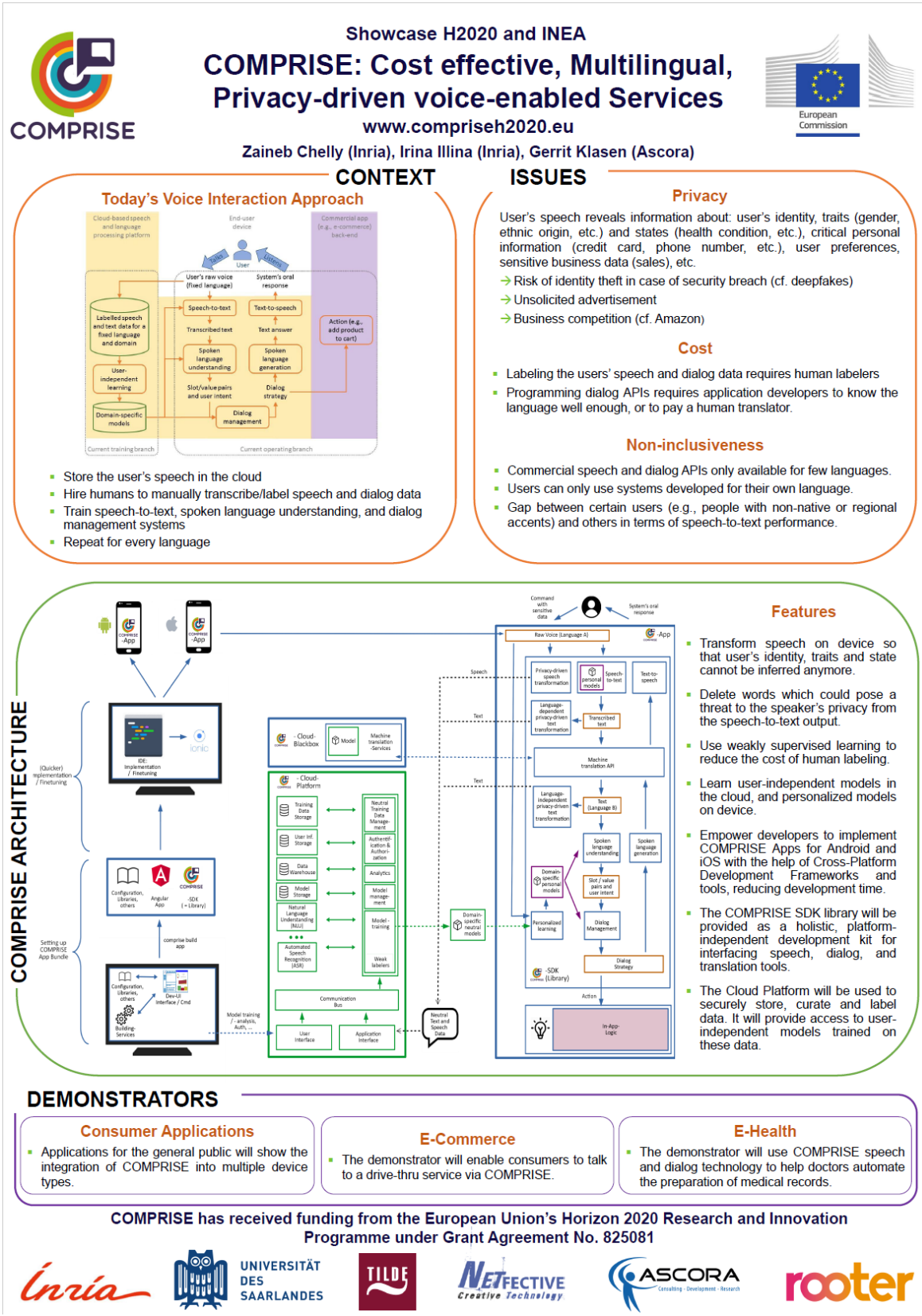

Figure 17: Poster

3.7.5. Factsheets / infographics

So far, three infographics have been produced and uploaded to the COMPRISE website:

- 10 things you should know about voice-based technologies (https://www.compriseh2020.eu/10-things-you-should-know-about-voice-based-technologies/)

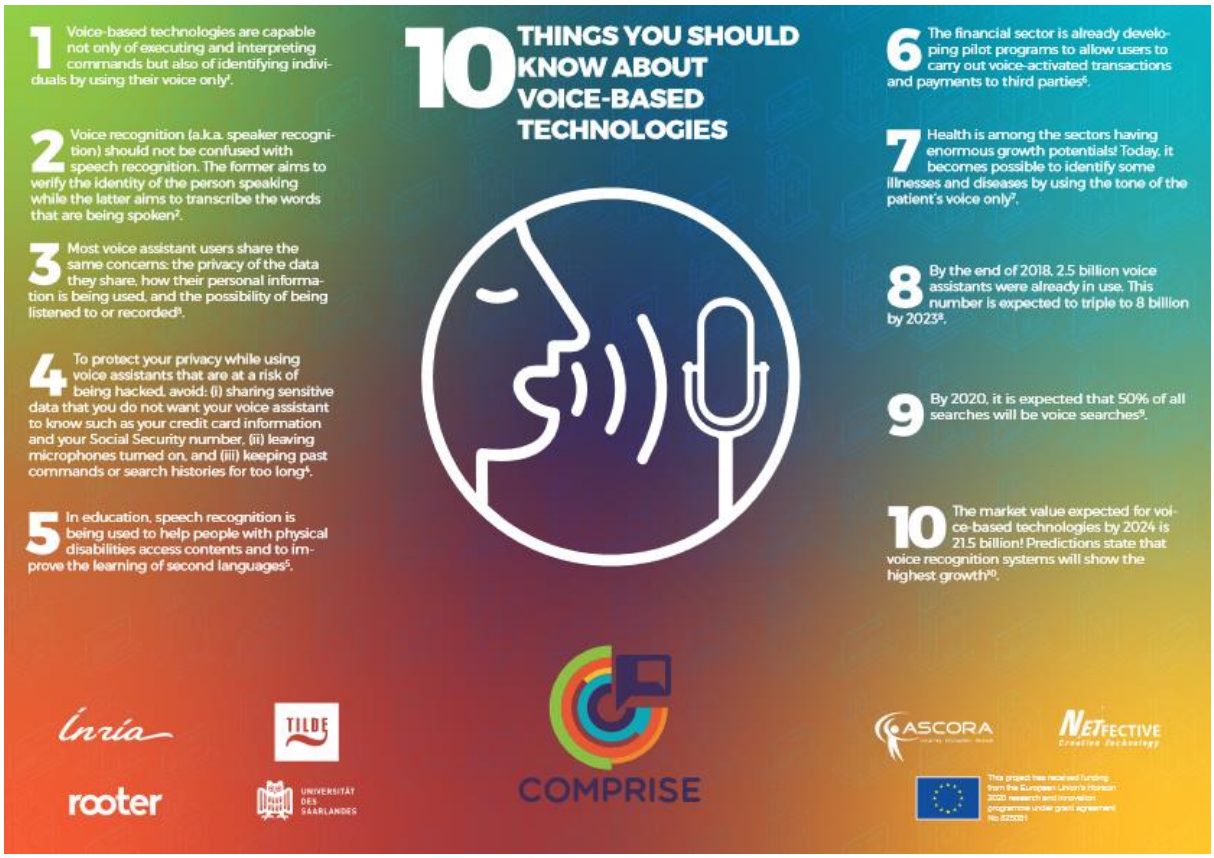

Figure 18: Infographic "10 things you should know about voice-based technologies"

- 10 privacy risks associated with voice-enabled technologies (https://www.compriseh2020.eu/10-privacy-risks-associated-with-voice-enabled-technologies/)

Figure 19: Infographic "10 privacy risks associated with voice-enabled technologies"

- Business opportunities around voice-enabled technologies (https://www.compriseh2020.eu/business-opportunities-around-voice-enabled-technologies/)

Figure 20: Infographic "Business opportunities around voice enabled technologies"

3.7.6. Blog articles

Besides the infographics, twelve articles have been published in the "Blog" tab of the COMPRISE website. Each blog entry has been written taking into consideration the author's (partner) expertise and role in the project.

Below are listed the entries available in the project's blog:

- Spoken Dialog Systems – are they fully reliable? (https://www.compriseh2020.eu/spoken-dialog-systems-are-they-fully-reliable/)
- Voice assistants: is someone listening to us? (https://www.compriseh2020.eu/with-voice-based-assistants-is-someone-listening-to-us/)
- Personal Data: Concept and categorisation (https://www.compriseh2020.eu/personal-data-concept-and-categorisation/)
- Voice-based applications for E-commerce (https://www.compriseh2020.eu/voice-based-applications-for-e-commerce/)

- Architecting a privacy-preserving dialogue system SDK (https://www.compriseh2020.eu/architecting-a-privacy-preserving-dialogue-system-sdk/)
- Voice-based applications for E-Health (https://www.compriseh2020.eu/voice-based-applications-for-e-health/)
- Privacy-driven speech transformation with adversarial learning (https://www.compriseh2020.eu/552-2/)
- Implementation of GDPR principles in machine learning (https://www.compriseh2020.eu/implementation-of-gdpr-principles-in-machine-learning/)
- Is there anything else I should know about you? (https://www.compriseh2020.eu/is-there-anything-else-i-should-know-about-you/)
- Handling private information in text (https://www.compriseh2020.eu/handling-private-information-in-text/)
- Privacy vs. usability (https://www.compriseh2020.eu/privacy-vs-usability/)
- Voice-based applications for consumers (https://www.compriseh2020.eu/voice-based-applications-for-consumers/)

3.8. Press release and media activities

Currently, two press releases related to the project activities have been released. The press releases have been disseminated in all the project partner's websites, social media accounts and in some digital journals. The press releases published are listed below:

- Towards the next generation of voice interaction technology

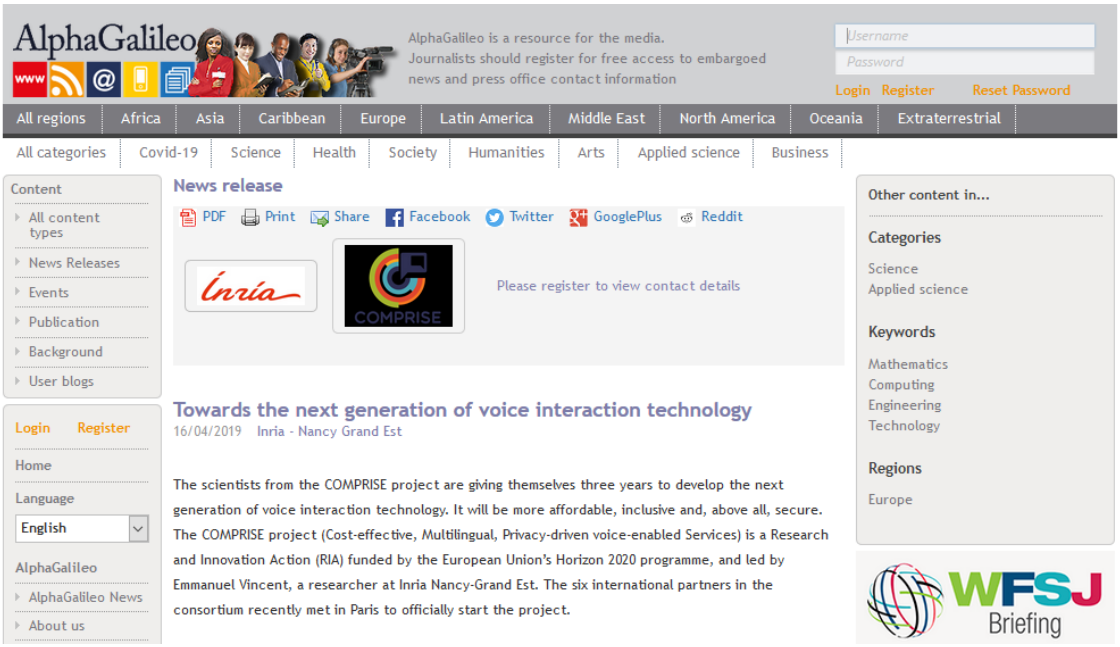

Figure 21: First press release

- Speech data privacy: a real challenge!

Nancy, 12 March 2020

Media Release

COMPRISE

Speech data privacy: a real challenge!

© Inria / Photo G. Scagnelli

Registration is now open to take part in the first international challenge on speech data privacy. The VoicePrivacy initiative is spearheading the effort to develop privacy preservation solutions for speech technology. It aims to gather a new community to define the most effective processes and metrics, while also benchmarking the existing privacy-enabling solutions and creating a baseline against which to evaluate future improvements. This groundbreaking initiative is now looking for forward-thinking and passionate participants.

Vocal access systems and assistants are almost ubiquitous, picking up, storing and analyzing the data routinely being "told" to them. In parallel, questions are increasingly being raised about the privacy of these data, and the security of service users. This is because speech data contains much more than just the spoken words. Speech encapsulates a wealth of personal data (like age and gender), health and emotional state, racial or ethnic origin, geographical background, social identity, and socio-economic status. Since machines can now decipher spoken language with ever-impressive accuracy, there is no reason why political orientations, religious and philosophical beliefs could not also be derived from speech data.

Two main solutions are under development to preserve privacy in speech data: homomorphic encryption and anonymization. Homomorphic encryption protects personal data by making it inaccessible to third parties, however it is computationally cumbersome. Anonymization, by contrast, suppresses personal information in the data; it can run in real time but the absence of clear definitions and standards raises questions about its efficiency.

The VoicePrivacy initiative is the first of its kind: it aims to bring together a community of experts to focus on this specific security challenge. It sets a standardized frame and will be a crucial step forward to define the currently missing metrics and processes to evaluate speech privacy, and the efficiency of the technologies developed to protect it. The participants will be asked to process the same dataset of speech signals to anonymize them, while protecting the linguistic content and speech naturalness.

The co-organizers of the VoicePrivacy initiative are:
- University of Avignon - LIA, France - *http://univ-avignon.fr/international-english-version-3508.kjsp*
- Inria, France - *inria.fr*
- National Institute of Informatics, Japan - *nii.ac.jp/en/*
- EURECOM, France - *eurecom.fr/en*
- University of Edinburgh, UK - *ed.ac.uk/*

The VoicePrivacy initiative is supported by:
- Agence Nationale de la Recherche (projects Deep-privacy, Harpocrates and VoicePersonae) - *anr.fr*
- European Union's Horizon 2020 Research and Innovation Program (project COMPRISE) - *compriseh2020.eu*
- Japan Science and Technology Agency (project VoicePersonae) - *jst.go.jp/EN/*

Figure 22: Second press release

4. KPIs

The current status of the KPIs set for communication and dissemination activities is detailed in Tables 15 and 16.

Table 45: Dissemination KPIs.

Dissemination KPI's	Status
KPI-DISS-1: At least 3 journal papers per year on average	Currently, 0 journals published. See actions to be implemented in Section 5.3.

PI-DISS-2: At least 8 conference or workshop papers per year on average	So far, COMPRISE has four published or accepted conference/workshop papers (see Section 3.4). Several more have been submitted. See actions to be implemented in Section 5.3.
KPI-DISS-3: At least one collaboration with another H2020 project under this call.	ELITR and ELG have shown interest in collaborating with COMPRISE (see Section 3.2).
KPI-DISS-4: At least 500 downloads for public deliverables and prototypes	So far, we have registered 402 downloads for public deliverables and prototypes.
KPI-DISS-5: At least 1 public communication in a networking event organized as part of relevant associations, for example, META-NET, BDVA.	KPI-DISS-5 has been fulfilled as COMPRISE have been presented within two networking events organized as part of META Forum and the BDV Summit.
KPI-DISS-6: At least 1 public communication in an EC event on data protection and privacy	COMPRISE has not participated in an EC event on data protection and privacy. See actions to be implemented in Section 5.2.
KPI-DISS-7: 2 workshops organized	A Dagstuhl workshop on Privacy in Speech and Language is being organized in Jan. 2021.

Table 56: Communication KPIs.

Communication KPI's	Status
PI-COMM-1: Web presence: at least 5000 new visitors per year to project website	So far, 1,708 unique visitors have been registered. See actions to be implemented in Sections 2.4 and 2.5.
PI-COMM-2: Project story factsheets: at least 10 story factsheets	3 factsheets have been published until now. See Section 3.7.5.
KPI-COMM-3: Press releases: at least 10 press releases	2 press releases have been published. See actions to be implemented in Section 5.7.
KPI-COMM-4: Media materials: at least 5 project videos	5 project videos have been already produced. See Section 3.7.1.
KPI-COMM-5: Twitter activity: at least 4 tweets per month	An average of 11 tweets per month are published

5. Next steps

To maximise the impact of COMPRISE and fulfil the KPIs defined for the project and listed in Deliverable D7.1, the dissemination communication actions previously described will continue to be carried out during the whole project; including new actions that have been considered necessary and that are described below.

5.1. Specific actions focused on attracting software developers

One of the main objectives of COMPRISE is to attract developers (as the main group of targeted stakeholders) and help them design voice-enabled software. In this sense, current actions will be reinforced to reach the goal of creating a community of software developers to exploit COMPRISE results. New actions will be implemented for this purpose.

5.1.1. Restructuring of Gitlab

Our Gitlab repository has been restructured to highlight COMPRISE's Assets as listed in Appendix A of D7.2 "Initial Exploitation Plan" and offer a better experience to developers. It now provides a separate top-level public entry for every Asset :

- COMPRISE SDK
- Voice Transformer
- Text Transformer
- Speech-to-text Machine Translation
- COMPRISE Cloud-based Platform
- Speech-to-Text Weakly Supervised Learning
- Spoken Language Understanding Weakly Supervised Learning

5.1.2. Software documentation and How-To videos

We will pay particular attention to software documentation. The documentation shall describe the functionalities of COMPRISE components and how to install and use them.

How-To videos will also be recorded showing how to install and use them.

5.1.3. Other actions to attract developers

Other actions have already been carried out to attract software developers and will continue during the next months. These include:

- Identifying events for software developers where COMPRISE results can be disseminated.
- Paying special attention to developers' profiles when targeting potential followers on social media.
- Joining software developer discussion groups on LinkedIn to disseminate COMPRISE results.

5.2. Events

All COMPRISE partners will continue keeping track of and participating in different events (e.g., conferences, workshops, etc.) to disseminate the project's results. As mentioned above, special attention will be placed on identifying software developer events in which COMPRISE partners can participate in the future.

Regarding the "KPI-DISS-6: At least 1 public communication in an EC event on data protection and privacy", it has been noticed that speakers at EC events on data protection and privacy are typically upon invitation only and cover a disjoint scope from

COMPRISE. Therefore, it has been decided to target events which may not be organised by the EC but cover a broad range of disciplines and stakeholders. These include the Annual Privacy Forum (APF) and the International Conference on Computers Privacy and Data Protection (CPVP), among others, to which we will aim to submit our GDPR-related work.

Finally, regarding workshops, COMPRISE is currently organizing a Dagstuhl workshop on Privacy in Speech and Language Technology which will take place in January 2021. The expected audience will include academic researchers, industry representatives, and policy makers from the fields of speech processing, natural language processing, computer security and privacy, machine learning, and law and ethics.

At a later stage of the project, a second workshop will be organised. This workshop will mainly target application developers.

5.3. Publications

The COMPRISE partners will continue publishing papers and articles in conferences, journals and specialised magazines.

We have listed the articles we have recently submitted or we intend to submit to journals in 2020. Here are the tentative titles:

- How can contextualization be used to identify private information recorded by voice-enabled systems?
- Speech anonymization strategies and attack models
- The VoicePrivacy 2020 Challenge: an analysis of the outcomes
- Text transformation techniques that preserve privacy in text corpora
- Optimal differentially private mechanisms

We have also listed the papers we have recently submitted or we intend to submit to conferences or workshops in 2020. Here are the tentative titles:

- Who started this rumor? Quantifying the natural differential privacy guarantees of gossip protocols
- Robust differentially private training of deep neural networks
- An introduction to the VoicePrivacy initiative
- Baseline speaker anonymization for the first VoicePrivacy challenge
- A comparative study of speech anonymization measures
- On semi-supervised training of LF-MMI acoustic models in limited data scenarios
- Speaker adaptation of neutral ASR acoustic models
- Differentially private adversarial learning for speech anonymization
- Learning embeddings for named entities to improve cross domain adaptation of NER
- Prevention of memorization in word representations for privacy-preserving NLP
- Controlling the content of private embeddings through information theory
- Private binary networks
- Robust text transformation for privacy-preserving SLU training
- Anonymization of age and gender in speech transcripts
- Weakly supervised learning for resource-effective SLU tasks

5.4. Website

The following actions will be carried out to improve the website visibility and gain visitors:

- **Creating a more attractive homepage**. Currently, the home page of COMPRISE is the "About the Project" section. ROOT carried out a survey with professionals who agreed that, although the homepage provides enough information about the project, it is not visually attractive (e.g., the amount of text needs to be reduced). A new homepage is currently being designed.
- **Adding links from other websites**. We will add more links to relevant resources (AI-related blogs, other projects, etc.) into the "Links" section, and ask them to include a link to the COMPRISE website on their own website. This will help better position the COMPRISE website.
- **Adding a search box.** Adding a search box will improve user experience. This will also help to determine what contents visitors are more interested in.

5.5. Contacting other websites

To increase website visits and disseminate the project's materials, we will contact other website administrators (specialised in AI, technology, privacy) and propose them to publish the project's blogs and materials, as we already did with AIHUB.

5.6. Generation of materials

COMPRISE will continue producing videos according to the plan (see Section 3.7.1). Besides, blogs and infographics will continue to be produced by partners on a monthly basis. Other communications materials will be produced if needed.

5.7. Press releases

As most press releases will coincide with the release of COMPRISE components in the coming months, a press release plan has been prepared, containing the estimated dates of publication and the topics that will be addressed.

6. Conclusions

This deliverable describes the main communication and dissemination activities carried out from month 1 to month 16 and the results achieved so far.

Although most of the KPIs defined in the DOA regarding communication and dissemination activities are on track, a few others (i.e., KPIs for publications and website visitors) need a closer follow up. We proposed a range of actions to improve our performance regarding these KPIs and reach the expected objectives.

Based on the comments provided during the Review Meeting in January 2020, particular emphasis is being put on creating a software developer community around COMPRISE to facilitate the exploitation of the project's Assets. This is reflected both in the actions carried out recently and in those to be carried out by all partners in the coming months.

Agreements with other projects, blog publishers or other digital media will also be sought to disseminate and communicate COMPRISE results and thus reach the maximum possible number of stakeholders.