



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 N°: 825081

WP N°7: Communication, dissemination and exploitation

Deliverable N°7.5: Final dissemination and communication report

Lead partner: ROOT

Version N°: 1.0

Date: 31/10/2021



Document information	
Deliverable N° and title:	D7.5 – Final dissemination and communication report
Version N°:	1.0
Lead beneficiary:	ROOT
Author(s):	Álvaro Moretón, Conrado Castillo, Laura Merlo (ROOT)
Reviewers:	Raivis Skadiņš (TILDE), Carolin Wübbe (ASCO)
Submission date:	31/10/2021
Due date:	31/10/2021
Type ¹ :	R
Dissemination level ² :	PU

Document history			
Date	Version	Author(s)	Comments
04/10/2021	0.1	Álvaro Moretón, Conrado Castillo, Laura Merlo	First draft of the deliverable
14/10/2021	0.2	Carolin Wübbe	Proofreading and review
23/10/2021	0.3	Álvaro Moretón, Laura Merlo	Deliverable after the first review
26/10/2021	0.4	Raivis Skadiņš	Proofreading and review
31/10/2021	1.0	Akira Campbell and Emmanuel Vincent	Final version reviewed 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 T7.3 “Dissemination and Collaboration” and Task T7.4 “Communication Activities” of Work Package 7 “Dissemination Communication and Exploitation” of the COMPRISE project. The main objective of these tasks is to maximise the impact of COMPRISE results worldwide through the execution of dissemination and communication activities.

This is the second and final dissemination and communication report which aims to provide a final overview on the dissemination and communication actions carried out from Month 1 (December 2018) to Month 34 (September 2021) of the project.

This document highlights and describes in more detail the communication and dissemination activities carried out within the project from April 2020 to September 2021 (the final reporting period) as well as the results achieved due to these actions.

Table of contents

1. Introduction	6
2. Targeted stakeholders	6
3. Dissemination and communication activities	8
3.1. Targeted dissemination	8
3.1.1. Stakeholders online targeting	8
3.1.2. Discussion groups	8
3.1.3. Events monitoring	10
3.1.4. Journals and magazines	11
3.1.5. Cooperation with other websites	11
3.2. Joint actions with other H2020 projects	12
3.3. Events	12
3.4. Publications	15
3.4.1. Scientific and specialised publications	15
3.4.2. Popularised publications	17
3.5. Website	18
3.5.1. Website sections	18
3.5.2. COMPRISE website impact	25
3.6. Social media	27
3.6.1. Twitter	28
3.6.2. LinkedIn	28
3.6.3. YouTube	28
3.7. Generation of content	28
3.7.1. Videos	28
3.7.2. Leaflet	31
3.7.3. Roll-up	32
3.7.4. Poster	33
3.7.5. Infographics	33
3.7.6. Blog articles	39
3.8. Press releases and media activities	41
3.8.1. Press releases	41
3.8.2. Other media activities	44
4. KPIs	45
5. Specific actions to attract software developers and final users	46
5.1. Website	46

5.2. Re-organisation of Gitlab	46
5.3. Proper documentation	46
5.4. How-to videos	47
5.5. Developer survey	47
5.6. Websites for COMPRISE assets	47
5.7. Materials	48
5.8. Other actions to attract developers	48
6. Conclusions	48

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 exploitation, dissemination and communication activities. The main objective of this WP is to carry dissemination, communication and exploitation activities to make COMPRISE a reference crossroad on the broad topic of software engineering for multilingual voice-enabled applications and maximise its impact worldwide.

In Month 3 (February 2019), Deliverable D7.1 “Dissemination and Communication Action Plan”, which aimed to define the dissemination and communication activities to be carried out during the project, was submitted.

Deliverable D7.3 “First Communication and Dissemination Report”, the second deliverable of WP7, was submitted on Month 17 (April 2020). This deliverable offers an overview of the activities directly related with Tasks T7.3 “Dissemination and Collaboration” and T7.4 “Communication activities” carried out during the first reporting period of the project as well as the results of these activities.

The purpose of this document, Deliverable D7.5 “Final Communication and Dissemination Report”, is to provide a summary of the dissemination and communication activities carried out within COMPRISE from Month 1 (December 2018) to Month 34 (September 2021) of the project, paying special attention to the activities carried out from Month 17 to Month 34 (the final reporting period). Readers will have access to a summary of the information provided in the previous deliverable (D7.3) plus an update on all the activities carried out from its submission until September 2021 (the final reporting period).

2. Targeted stakeholders

The dissemination and communication activities described in this document aim to get the maximum impact of the project results amongst its 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 Deliverable D7.1.

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

The table below offers an update of the stakeholders (organisations) that have shown interest in COMPRISE and that have already contacted at least one of the partners of the Consortium (the ones in bold have contacted COMPRISE partners during the last reporting period).

Table 1: Stakeholders.

Type of stakeholder	Shown interest in the COMPRISE project ³	Expressed interest in exploiting the results of COMPRISE ⁴	Use products based on COMPRISE technology
SMEs	Llorente y Cuenca, Aflorithmic, Modality.AI, Norsys, Vox.care, Zlroh Labs, IP Soft. Meditec	i4FS, DRS, ProMed, Mediaskopas, Talkamatic, Dir IPS, hel-lomybot.io, SilWorld, MyC, /e/ Foundation	Assistentis, Vi-voka
Large companies	Google, BBVA, Mahou, Marktel, Intel Corporation, Tet, Eesti Energia, Intermarché, Pierres & Vacances, AWS	Latvenergo	
Medical service organisations	Ministry of Healthcare Of Latvia, Ministry of Healthcare of Estonia, Stradiņš University Hospital	Riga Children's Clinical University Hospital, Estonian Institute of criminology, Tallinn Eastern Central hospital, ARS (largest private clinic in Latvia), Lille University Hospital	Vidzeme regional hospital
Researchers	University of Wisconsin-Madison, University of Pau, Imperial College of London	KU Leuven, Karlstad University, DFKI, Norwegian Computing Centre, University of Potsdam, University of California	EURECOM, LIA, NII, registered participants of the VoicePrivacy Challenge (40 academic or industrial teams from China, France, India, Japan, UK, USA, etc.), Queen Mary University of London, Vector Institute
EU Projects	SAAM, StairwAI	ZDMP, ELITR, SPRING, LT-Bridge, FVLLMONTI	ELG
Media (magazines, journals, television, radio...)	Diario Economista, Diario EL país, Latvian State TV and Radio, Delfi (EE, LV, LT), France 3, L'Usine Nouvelle, Rue89 Strasbourg, We Demain, VNet, LETA, Latvian State TV, SIA 4. vara, SIA Helio Media		
Event organiser	VoiceTech Paris, Voice & Text Bots Cologne, Lampa		

³ E.g., Requested information about the project⁴ E.g., develop products based on COMPRISE technology, carry out additional research based on COMPRISE results

Type of stakeholder	Shown interest in the COMPRISE project ³	Expressed interest in exploiting the results of COMPRISE ⁴	Use products based on COMPRISE technology
Policymakers	Garante per la protezione dei dati personali, Ministry of Justice of Latvia , Ministry of Justice of Estonia	CNIL, EDPB , Ministry of Interior of Latvia , State Chancellery of Lithuania , Ministry of Foreign Affairs of Lithuania , Latvian Courts Administration	
Others	ELDA, Sadzirdi (NGO)	Almond, SPEAKER (Fraunhofer IIS)	

3. Dissemination and communication activities

From the beginning of the Project, several communication and dissemination actions have been carried out to publicise the activities and results of the COMPRISE project to maximise its impact. All communication and dissemination actions, as well as their results, have been tracked along the project.

3.1. Targeted dissemination

We have continued our monitoring actions to locate forums related to the project's activities (voice recognition systems, data privacy, developers, etc.), where partners can participate actively, and promote the project and its results. During the final reporting period hunting actions have been more focused on reaching developers and potential final user groups. The targeted dissemination actions that were carried out are detailed below.

3.1.1. Stakeholders online targeting

During the final reporting period new potential stakeholders have been targeted and followed through the COMPRISE Twitter account to draw their attention towards the project. 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 the following link:

<https://twitter.com/compriseh2020>.

The COMPRISE Twitter account 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 for each stage of the project, as well as to portray an active image of the account (e.g., new followers, interactions, posts, etc.).

3.1.2. Discussion groups

We have kept searching, joining, and participating in discussion groups related to the project's activities on LinkedIn, to disseminate the publications and results achieved. Specifically, during the reporting period we have joined new developer discussion

groups, because developers have been identified as one of the most important categories of stakeholders of COMPRISE. Table 2 shows all the LinkedIn discussion groups in which we have disseminated COMPRISE materials (deliverables, videos, blogs, news, etc.). The groups in the green cells are the ones which we have joined during the final reporting period.

Table 2: Discussion groups (LinkedIn).

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
Mobile App Professionals	24,945
Flutter Developer Community	29,363
NLP, AI ML professionals	2,946
Developers - Android, iOS developer, Blockchain, Ethereum, Java, Ruby, .net, php, django, etc	295,435
Freelance Mobile Apps Developers	661

During the last reporting period we have also joined discussion groups on Facebook to disseminate COMPRISE results and materials, as listed in Table 3.

Table 3: Discussion groups (Facebook).

Group name	Number of members
Artificial Intelligence and Machine Learning	175,000
Cisco Voice Xperts	2,600
Artificial Intelligence Machine Learning Data Science Deep Learning	27,900
ChatBots	23,900
Data Science - Machine Learning, Artificial intelligence, Deep Learning	56,000
Artificial intelligence, Machine learning, Deep learning	190,000

3.1.3. Events monitoring

The targeted dissemination activities involve spotting events (conferences, congresses, fairs, workshops, etc.) where the project results can be disseminated. A document has been created and shared with the project partners, where everyone can contribute by adding events they are aware of and that can be considered for future dissemination actions. Deliverables D7.1 and D7.2 listed the events identified for 2019 and 2020. Table 4 lists the events identified for 2021, during the final reporting period.

Table 4: Events 2021.

Event name	Link	Date and Location
PRIVACY		
CPCD	https://www.cpdconferences.org/	March 17 th
IAPP Privacy summit	https://iapp.org/conference/summit-online/	April 17 th – 30 th
PEPR	https://www.cpdconferences.org/	June 10 th – 11 th
IEEE European Symp. on Security and Privacy	https://www.ieee-security.org/TC/EuroSP2021/	September 6 th – 10 th
Bitkom	https://www.bitkom-live.de/de/privacy-conference	September 27 th – 28 th
ISCA SPSC	https://spsc-symposium2021.de/	November 10 th – 12 th
AI / VOICE TECHNOLOGIES		
CES 2021	https://www.ces.tech/	January 11 th – 14 th
Voice Tech	https://voice-tech-conference.com/	March 21 st
Project Voice Worldwide	https://www.projectvoice.ai/project-voice	April 15 th – 16 th
ACL-IJCNLP 2021	https://2021.aclweb.org/	August 2 nd – 6 th
MT Summit	https://amtaweb.org/mt-summit2021	August 16 th – 20 th
Interspeech	https://www.interspeech2021.org/	August 30 th – September 3 rd
The Voice of Mobile Apps	https://www.eventbrite.com/e/the-voice-of-mobile-apps-tickets-132749430119	September 29 th
The Bots & Assistants Conference	https://www.chatbotconference.com/	November 9 th – 10 th
EMNLP	https://2021.emnlp.org/	November 7 th – 11 th
Voice summit	https://www.voicesummit.ai/	December 6 th – 9 th
APP DEVELOPERS		
Developer Week Virtual	https://www.developerweek.com/global/	April 17 th – 19 th
ACCU	https://accu.org/conf-menu-overviews/conference/	April 4 th
ADDC	https://addconf.com/2021/	June 23 rd – 24 th
iCONF	https://dev.events/conferences/i-conf-21-online-8-2021	September 27 th – 28 th
SWIFT LEEDS	https://swiftleeds.co.uk/	October 7 th
DROID CON LONDON	https://dev.events/conferences/-droidcon-london-ae1ce692	October 28 th – 29 th

3.1.4. Journals and magazines

A document has been created and shared with the project partners, where everyone can contribute by adding specialised magazines, journals, or other publications they are aware of and where papers and articles related to the project's results can be published. During the final reporting period new publications were added to the list, especially online technology and science magazines and journals. Table 5 shows the new publications that were added during the final reporting period.

Table 5: Journals and magazines.

Publication	Link
Computer Speech and Language	https://www.journals.elsevier.com/computer-speech-and-language/call-for-papers/special-issue-on-voice-privacy
Hashed out	https://www.thesslstore.com/blog/write-for-hashed-out/
ERCIM News 126 (July 2021), special theme: Privacy-preserving computation	https://ercim-news.ercim.eu/about-ercim-news
Technology.org	https://www.technology.org/submit-story/
Wired magazine (U.K.)	https://www.wired.co.uk/article/contributor-guide-lines
Technology End	https://technologyend.com/write-guest-post/
Proche	https://www.theproche.com/2018/12/10/submit-guest-blogs-on-voice-recognition-technology-write-for-us/
Future Play	https://futureplay.org/contribute-guest-post/
Tech Research Online	https://techresearchonline.com/write-for-us/#
Techgenyz	https://www.techgenyz.com/become-guest-author/
Let's Talk Tech	https://www.letstalk-tech.com/author/guest-contributor/
TMC.net	https://www.tmcnet.com/usubmit/news/usubmit.aspx
Venture Beat	https://venturebeat.com/guest-posts/

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 its website AI-Hub (<https://aihub.org/about/>). During the final reporting period AIHub has kept disseminating new COMPRISE contents on its website and social media platforms. On the other hand, other online blogs have been contacted with the purpose of reaching a similar cooperation agreement, but we have received negative answers (or no answer). In all the cases the administrators contacted answered that they only accept to publish exclusive material, not material previously published in other websites as it is the case of COMPRISE materials.

3.2. Joint actions with other H2020 projects

Comprise has carried out cooperation actions with other H2020 projects. The actions carried during the final reporting period are described below.

- **ELG project** (<https://www.european-language-grid.eu/>): COMPRISE partners attended a meeting with the ELG on April 14th, 2021. Furthermore, two COMPRISE tools, the COMPRISE Voice Transformer and COMPRISE Text Transformer, were submitted in February 2021. ELG approved the COMPRISE Text Transformer which is now running as a functional service and the COMPRISE Voice Transformer which is available as a downloadable software.
- **ELRC** (<https://lr-coordination.eu/node/375>): After the meeting concerning text anonymisation in 2020, ELRC drafted an anonymisation specification for the European Commission and contacted COMPRISE for feedback on the Differential Privacy section (which is directly inspired by the work of COMPRISE). One ELRC partner (Tom Vanallemeersch) is also a member of our External Advisory Board. An article about COMPRISE was published in the ELRC Newsletter on June 2, 2021 (<https://lr-coordination.eu/index.php/node/372>)
- **SPRING** (<https://spring-h2020.eu/>): The coordinator of SPRING contacted us in January 2021 regarding ASR models for French. COMPRISE offered the use of models provided in the COMPRISE Cloud Platform. SPRING are currently evaluating the models offered.
- **LT BRIDGE** (<https://lt-bridge.eu/>): INRIA and TILDE gave presentations of speech and text anonymisation during the “Anonymisation Training for Young Researchers” workshop held on May 25th, 2021, which was co-organised by TILDE.
- **FVLLMONTI** (<https://fvllmonti.eu/>): INRIA is part of the FVLLMONTI External Advisory Board (attended a meeting on June 21st, 2021) and their coordinator has been invited to join the COMPRISE External Advisory Board (attended the 5th and 6th plenary meetings on June 28th, 2021 and October 21st, 2021).

3.3. Events

To ensure the successful dissemination of the COMPRISE results, the consortium partners have participated in various events (35 in total) between Month 1 and Month 34, both as organisers and speakers. These events included, among other, conferences and workshops.

Table 6 lists all events in which COMPRISE members have participated. The events in the green cells are the ones in which COMPRISE has participated during the final reporting period.

Table 6: Events.

Date	Name	Type
December 2018	ICT 2018 networking session	Activities jointly organised with other H2020 projects
March 2019	AI and Language for Citizens	Workshop (participation)
April 2019	Panel discussion “Positioning at the European level, an opportunity to be seized” at the General Assembly of the Materialia Competitiveness Cluster	Brokerage Event
April 2019	European Language Grid	Activities jointly organised with other H2020 projects
April 2019	Showcase H2020 and INEA	Activities jointly organised with other H2020 projects
June 2019	BDV PPP Summit 2019	Conference (organisation)
June 2019	MeetUp Intelligence Artificielle Nancy #4	Other events
July 2019	10 th edition of the French Privacy and Data Protection Workshop	Workshop (organisation)
September 2019	Interspeech 2019	Conferences (participation)
September 2019	ICT Proposers' Days	Brokerage Event
October 2019	META-FORUM 2019	Conferences (participation)
October 2019	Global Forum on AI for Humanity	Conferences (participation)
October 2019	Workshop for Riga Children's Hospital	Workshop (organisation)
October 2019	Science Festival	Other events
November 2019	Voice Tech Paris 2019	Conferences
November 2019	TPDP 2019 - Theory and Practice of Differential Privacy	Workshop (participation)
November 2019	Privacy Preserving Machine Learning at NeurIPS	Workshop (organisation)
November 2019	Privacy Preserving Machine Learning at ACM ACS	Workshop (organisation)
November 2019	Panel discussion (Rue89 Strasbourg)	Other events
April 2020	AfricaNLP / PML4DC Workshop 2020 at ICLR 2020	Workshop (participation)
May 2020	ICASSP 2020	Conference (participation)
August 2020	AISTATS 2020	Conference (participation)
September 2020	23rd International Conference on Text, Speech and Dialogue	Conference (participation)
September 2020	EAB Research Projects Conference (EAB-RPC) 2020	Conference (participation)
September 2020	Baltic HLT	Conference (participation)
September 2020	34th International Symposium on Distributed Computing	Conference (participation)
September 2020	MeetUp @CNIL	Other events
October 2020	VoicePrivacy Challenge special session at Interspeech 2020	Conference (organisation)
October 2020	Interspeech 2020	Conference (participation)
November 2020	VoicePrivacy workshop, co-located with Odyssey 2020	Workshop (organisation)
November 2020	EMNLP 2020	Conference (participation)
January 2021	Artificial intelligence seminar series, National Institute of Advanced Industrial Science and Technology (AIST)	Other events

Date	Name	Type
April 2021	Region Grand Est's webinar "Investing in digital technologies: from research to production, for daily use, Horizon Europe as an important tool"	Other events
August 2021	Privacy-preserving Machine Learning for Audio, Speech and Language Processing special session at Interspeech 2021	Conference (organisation)
August 2021	MT Summit	Conference (participation)
October 2021	CITI Laboratory workshop on privacy protection for medical data	Workshop (participation)

Furthermore, it is planned for COMPRISE to participate in the following events in the upcoming months, as listed in Table 7.

Table 7: Future events.

Date	Name	Type
November 2021	Lorentz Workshop - Speech as Personally Identifiable Information	Workshop (participation)
November 2021	Privacy Preserving Machine Learning Workshop at NeurIPS	Workshop (organisation)
November 2021	Privacy Preserving Machine Learning Workshop at ACM CCS 2021	Workshop (organisation)
November 2021	EMNLP 2021	Conference (participation)
November 2021	SPSC Symposium	Symposium (participation)
August 2022	Dagstuhl seminar on Privacy in Speech and Language Technology	Workshop (organisation)
September 2022	VoicePrivacy workshop, co-located with Interspeech 2022	Workshop (organisation)

The estimated number of people reached through the above-mentioned events is 4,158.

Table 8 categorises the people reached.

Table 8: People reached.

Category	Number of people reached
Scientific community	2,381
Industry	1,068
Civil society	121
General public	234
Policy makers	140
Media	89
Investors	20
Customers	60
Others	45

3.4. Publications

3.4.1. Scientific and specialised publications

The main purpose of this action is to publish papers and articles concerning COMPRISE results in international conferences and workshops, scientific journals, or specialised magazines. COMPRISE has published a total of 33 papers, with an additional 5 submissions that are currently under review, and another 5 that are currently in final preparation (see Table 9). The papers in the green cells have been produced, published, or submitted in the last reporting period.

Table 9: Publications.

Paper title	Status
“On the effect of normalization layers on differentially private training of deep neural networks” (journal)	In preparation. Will be submitted to the Transactions of the ACL.
“TOKEN is a MASK: Few-shot named entity recognition with pre-trained language models”	In preparation. Will be submitted to ACL 2022.
“Adapting language models when training on privacy-transformed data”	In preparation. Will be submitted to LREC 2022.
“Privacy and utility of x-vector based speaker anonymization” (journal)	In preparation. Will be submitted to the IEEE/ACM Transactions on Audio, Speech, and Language Processing.
“An overview of speech privacy research” (journal)	In preparation. Will be submitted to Computer Speech and Language.
“Transformer vs. LSTM language models trained on uncertain ASR hypotheses in limited data scenarios”	Submitted to ICASSP 2022
“Enhancing speech privacy with slicing”	Submitted to ICASSP 2022
“The VoicePrivacy 2020 Challenge: Results and findings” (journal)	Submitted to Computer Speech and Language
“Training RNN language models on uncertain ASR hypotheses in limited data scenarios” (journal)	Submitted to Computer Speech and Language
“Benchmarking and challenges in security and privacy for voice biometrics”	Submitted to SPSC Symposium 2021 (1 st ISCA Symposium on Security and Privacy in Speech Communication)
“MasakhaNER: Named entity recognition for African languages” (journal)	Published in the Transactions of the ACL, MIT Press, 2021, (10.1162/tacl)
“Monolingual and cross-lingual intent detection without training data in target languages” (journal)	Published in Electronics, MDPI, 2021, 10, (10.3390/electronics10121412)
“Preventing author profiling through zero-shot multilingual back-translation”	Published in 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), Nov 2021, Punta Cana, Dominica

Paper title	Status
“The effect of domain and diacritics in Yorùbá-English neural machine translation”	Published in 18 th Biennial Machine Translation Summit, Aug 2021, Orlando, United States
“Anonymisation and re-identification risk for voice data” (journal)	Published in European Data Protection Law Review, Lexxion, 2021, 7, pp. 274–284. (10.21552/edpl/2021/2/20)
“Transfer learning and distant supervision for multilingual Transformer models: A study on African languages”	Published in 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Nov 2020, Punta Cana, Dominica
“Who started this rumor? Quantifying the natural differential privacy guarantees of gossip protocols”.	Published in DISC 2020 – 34 th International Symposium on Distributed Computing, Oct 2020, Freiburg, Germany
“On semi-supervised LF-MMI training of acoustic models with limited data”	Published in INTERSPEECH 2020, Oct 2020, Shanghai, China
“Introducing the VoicePrivacy initiative”	Published in INTERSPEECH 2020, Oct 2020, Shanghai, China
“Design choices for x-vector based speaker anonymization”	Published in INTERSPEECH 2020, Oct 2020, Shanghai, China
“A comparative study of speech anonymization metrics”	Published in INTERSPEECH 2020, Oct 2020, Shanghai, China
“Privacy guarantees for de-identifying text transformations”	Published in INTERSPEECH 2020, Oct 2020, Shanghai, China
“Data augmentation for pipeline-based speech translation”	Published in 9 th International Conference on Human Language Technologies – the Baltic Perspective (Baltic HLT 2020), Sep 2020, Kaunas, Lithuania
“Using privacy-transformed speech in the automatic speech recognition acoustic model training”	Published in 9 th International Conference on Human Language Technologies – the Baltic Perspective (Baltic HLT 2020), Sep 2020, Kaunas, Lithuania
“Investigating the impact of pre-trained word embeddings on memorization in neural networks”	Published in 23 rd International Conference on Text, Speech and Dialogue, Sep 2020, Brno, Czech Republic
“Assessing unintended memorization in neural discriminative sequence models”	Published in 23 rd International Conference on Text, Speech and Dialogue, Sep 2020, Brno, Czech Republic
“Fully decentralized joint learning of personalized models and collaboration graphs”	Published in AISTATS 2020 – The 23 rd International Conference on Artificial Intelligence and Statistics, Aug 2020, Palermo, Italy

Paper title	Status
"Private protocols for U-statistics in the local model and beyond"	Published in AISTATS 2020 – 23 rd International Conference on Artificial Intelligence and Statistics, Aug 2020, Palermo, Italy
"How can private information recorded by voice-enabled systems be identified?"	Published in European Data Protection Law Review, Lexxion, 2020, 6 (3), pp. 464–469. (10.21552/edpl/2020/3/17)
"The COMPRISE Cloud Platform"	Published in 1st International Workshop on Language Technology Platforms, May 2020, Marseille, France
"Evaluating voice conversion-based privacy protection against informed attackers"	Published in ICASSP 2020 – 45 th International Conference on Acoustics, Speech, and Signal Processing, May 2020, Barcelona, Spain
"Distant supervision and noisy label learning for low resource named entity recognition: A study on Hausa and Yorùbá"	Published in ICLR Workshops (AfricaNLP & PML4DC 2020), Apr 2020, Addis Ababa, Ethiopia
Privacy-preserving adversarial representation learning in ASR: Reality or illusion?"	Published in INTERSPEECH 2019 – Sep 2019, Graz, Austria

Papers can be found in the COMPRISE webpage: <https://www.compriseh2020.eu/papers/>

3.4.2. Popularised publications

COMPRISE has published a total of five articles in popularised publications, as listed in Table 10. The papers in the green cells have been published during the final reporting period.

Table 10: Popularised publications.

Article title	Publication
Data labelling: the hidden cost of voice-enabled technologies	Published on Technology.org, September 2021
The accent gap problem in minorities and dialect speakers	Technology.org, April 2021
Peut-on faire confiance aux IA? (Can AI systems be trusted?)	The Conversation, November 2020
On transmet beaucoup plus d'informations par la voix qu'un simple message (We transmit a lot more information with our voice than just a simple message)	Rue89 Strasbourg, November 2019
Der verschwiegene Sprachassistent der Saar-Uni (The secret language assistant at Saar-Uni)	Saarbrücker Zeitung, April 2019

3.5. Website

The COMPRISE project 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.

The website is hosted and maintained by INRIA, with the support of ROOT. All partners participate in enriching the website by providing related contents and updates that can be published.

The design and structure of the website was changed in 2020 with the purpose of improving the user experience, by making software assets more visible and accessible. This modification was made to attract software developers and final users (the main users of the COMPRISE assets) once the core software components were ready for use.

3.5.1. Website sections

We describe below the different sections within the website as well as the main changes implemented in 2020.

- **Landing page:** A landing page with an improved design has been created. The landing page includes a video introducing the COMPRISE project ("What is COMPRISE?"). Additionally, it presents all the COMPRISE software assets, which are represented by different icons that include direct links to the software section of the website.



Figure 1: Landing page (part 1/3).

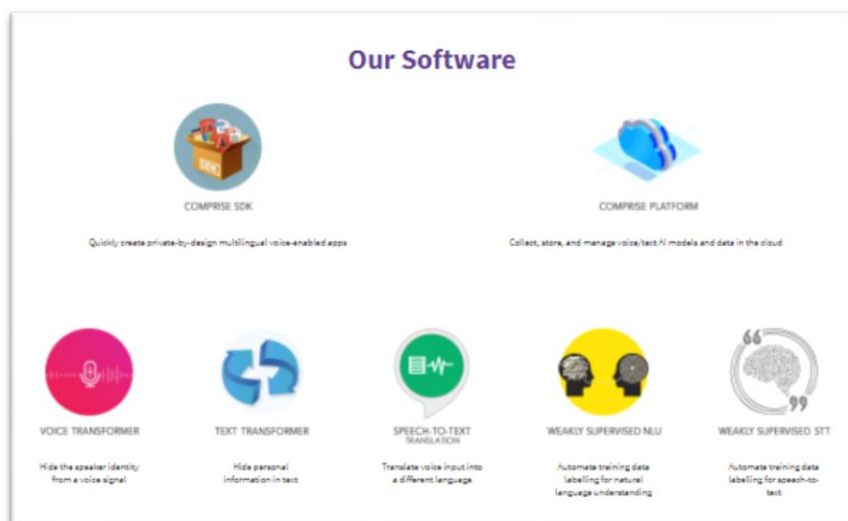


Figure 2: Landing page (part 2/3).

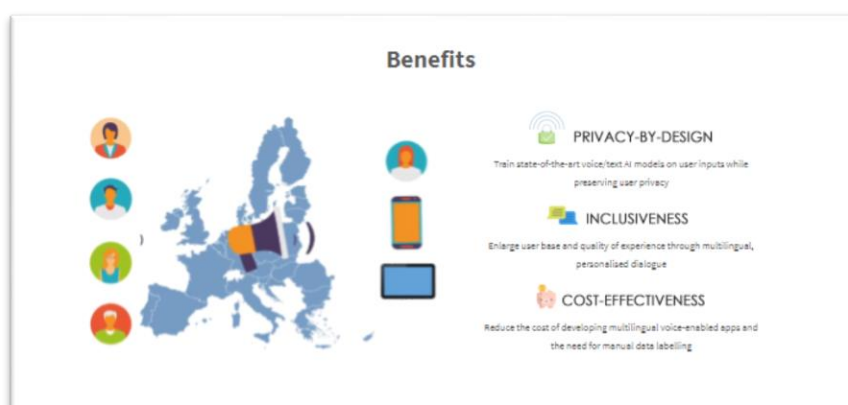


Figure 3: Landing page (part 3/3).

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

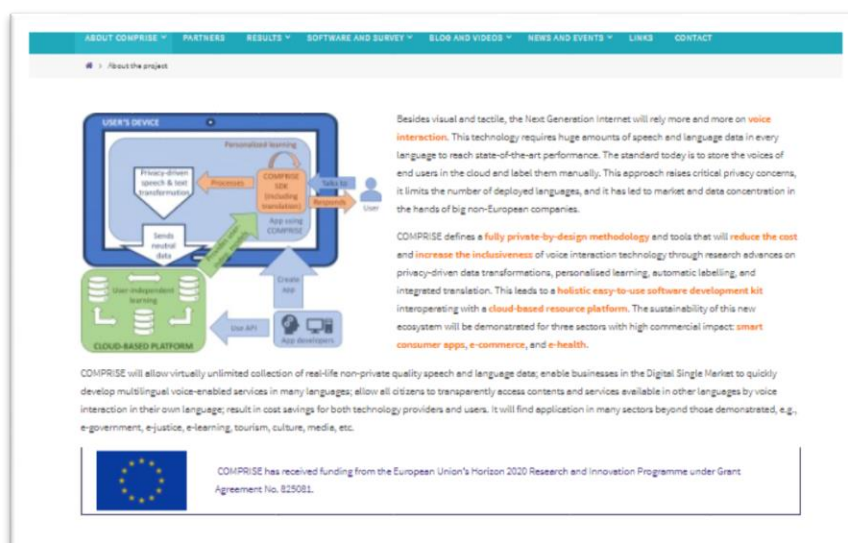


Figure 4: "About the project" page.

- **Partners:** In this section, visitors can find information about 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 here.

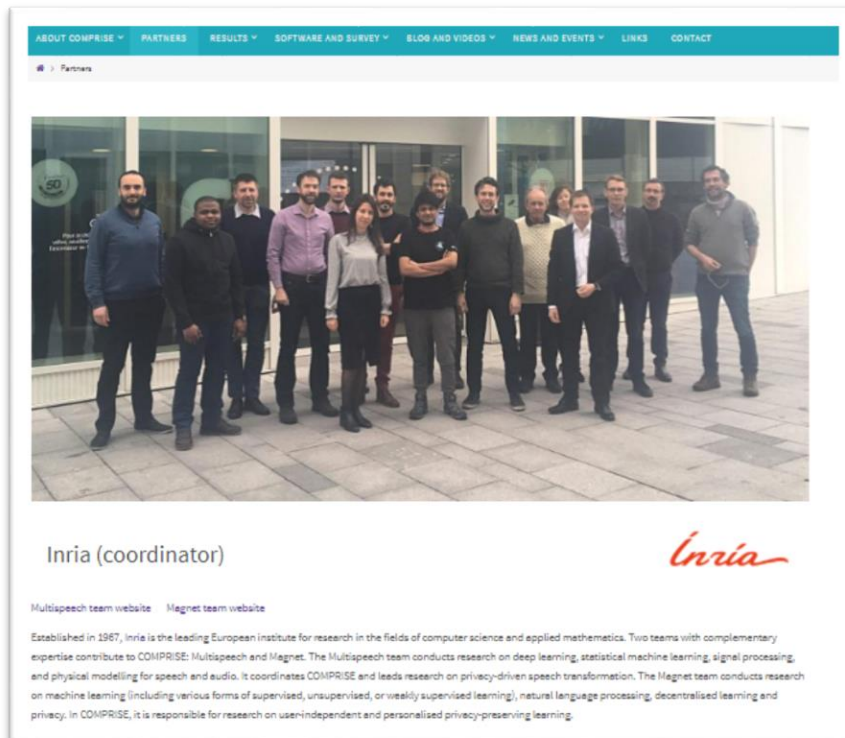


Figure 5: Partners page.

- **Results.** This section is divided into two subsections. Before implementing the modifications to the website there was a third “Software” subsection that has been moved to a new section dedicated exclusively to COMPRISE software solutions. The current subsections are:
 - **Deliverables:** Visitors can access and download public deliverables that have already been submitted. Currently there are 21 deliverables available, with an expected total of 23 by the end of the project.

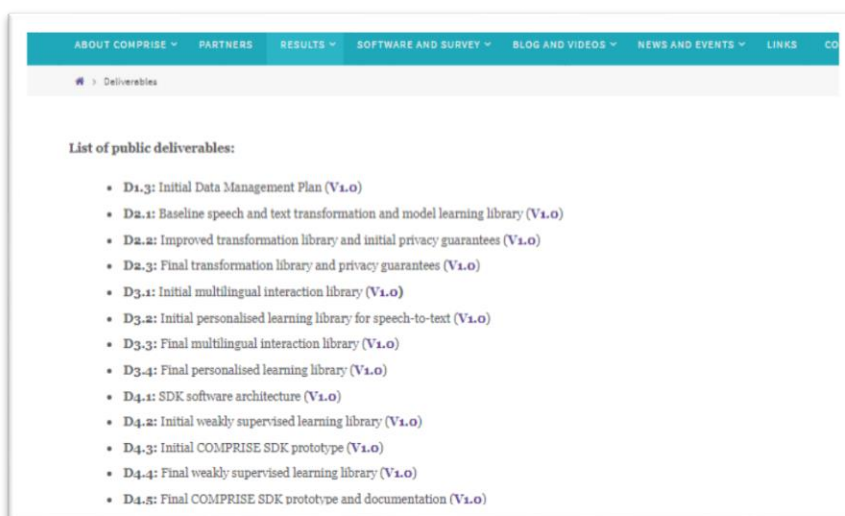


Figure 6: Deliverables page.

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

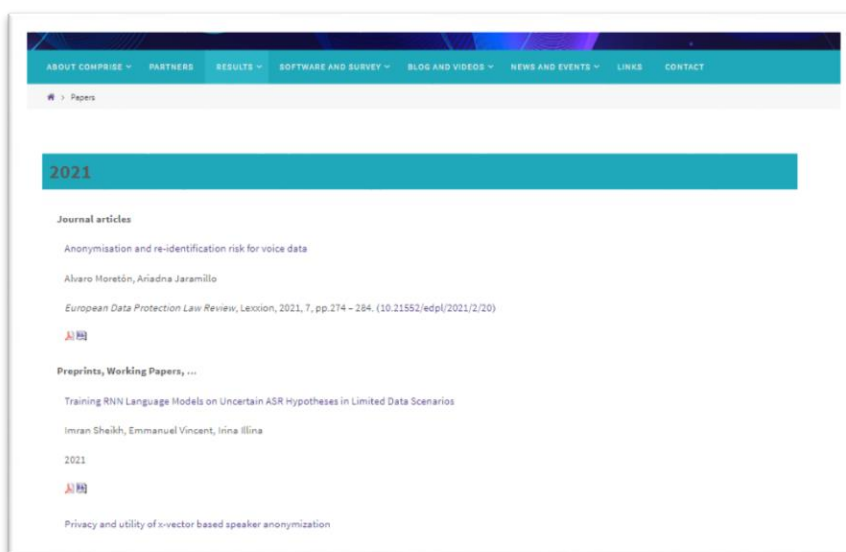


Figure 7: Papers page.

- **Software and survey:** A new section exclusively dedicated to the software assets produced in COMPRISE has been created to make these assets more visible and enhance their accessibility. This section is divided into the following sub-sections:
 - **Developer survey:** This section includes a 12-question survey for software developers with the purpose of better understanding the difficulties they encounter when integrating voice-enabled features into their apps.

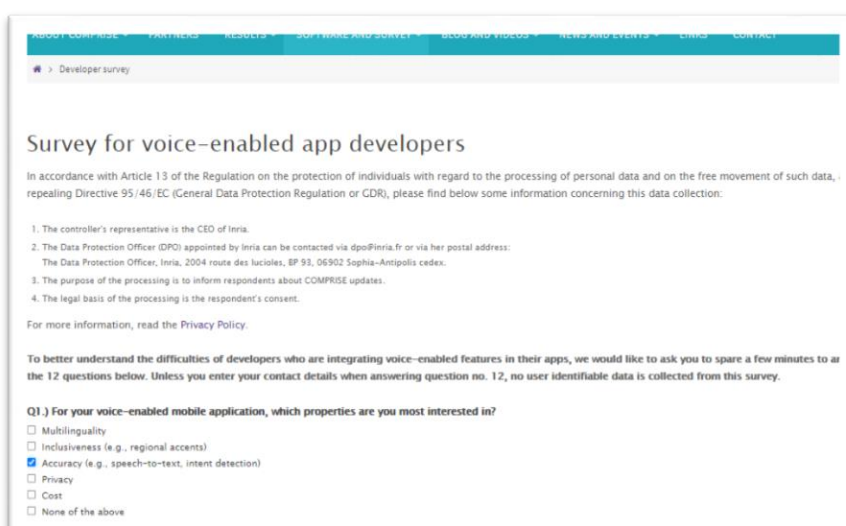


Figure 8: Developer survey page.

- **Software:** All the COMPRISE software assets are described in this section. The description of each software asset is accompanied with a link to

the Gitlab repository, if available, where it can be downloaded. In addition, we provide a direct link to a video which explains, in a user friendly manner, the functionalities and the problems that each asset is addressing.

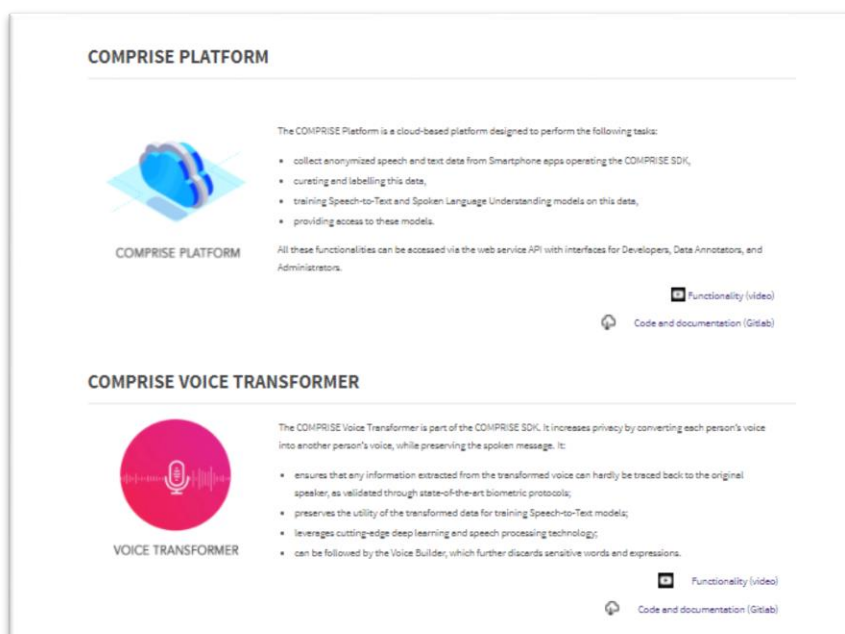


Figure 9: Software page.

- Blog and videos: This section was newly created in 2020 so that visitors to the website could easily read about the problem COMPRISE is solving. This section is divided in the following subsections:
 - **Blog:** In this section all blog entries that have been created by all of our project partners, and infographics related to COMPRISE activities are published. Currently there are 37 published blog entries. This subsection was previously included in the “News and events” section of the website.

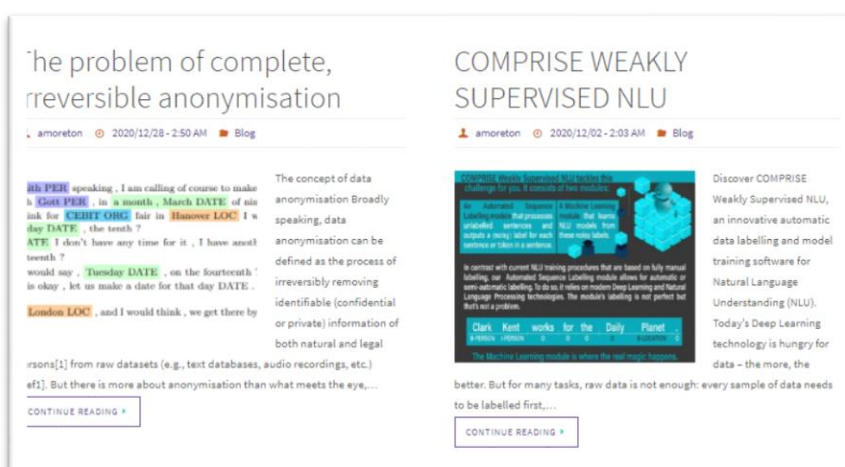


Figure 10: Blog page.

- **Videos:** In this section, visitors can access 25 project-related videos. This section has been reorganised with the videos categorised by groups (“Problems”, “Solutions”, “Interviews and media”, “Research Results”). This subsection was previously included in the “News and events” section of the website.

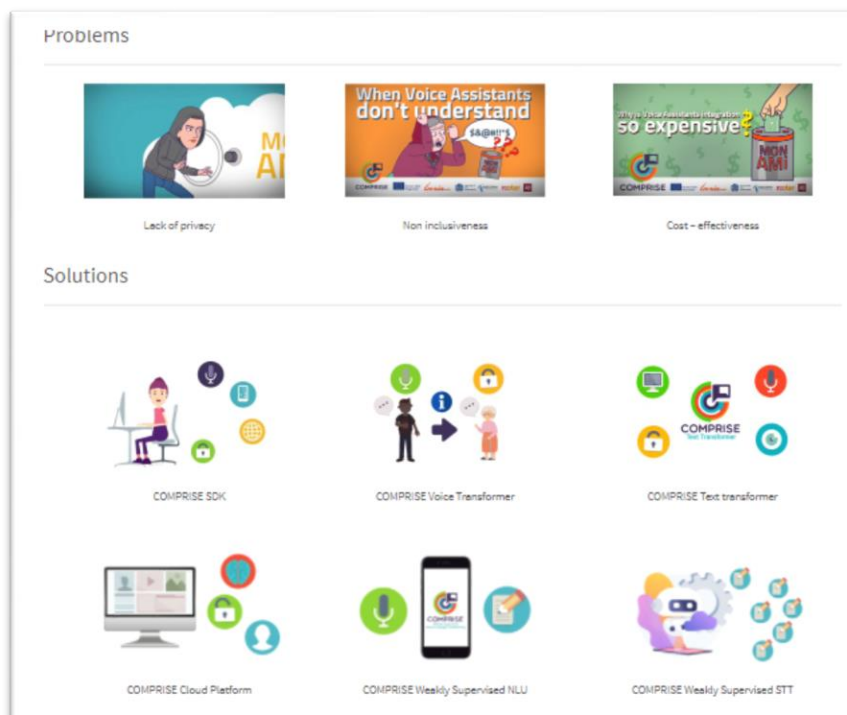


Figure 11: Videos page.

- **News and events.** This section is divided into two subsections. Previously it had two additional subsections “Blog” and “Videos” that were moved to a new section.
 - **Project news:** In this section news related to the project’s progress are published.

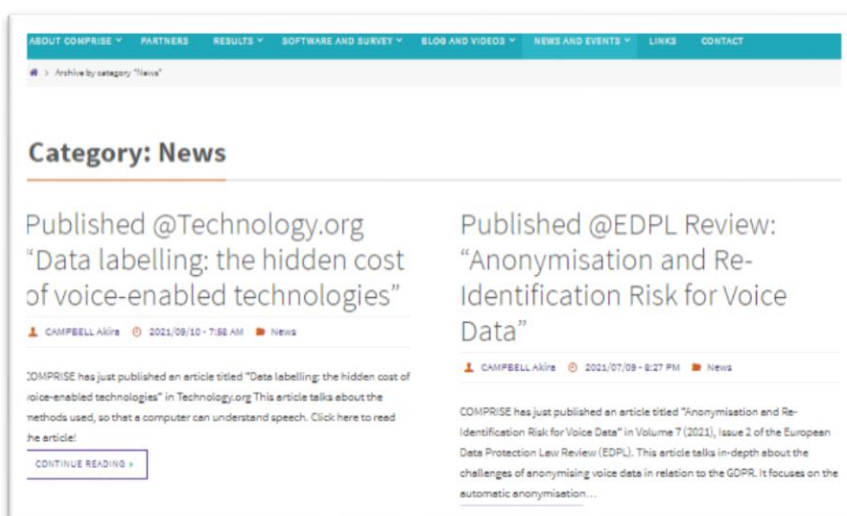


Figure 12: Project news page.

- **Events:** In this section news related to events in which partners have participated or will participate to disseminate the project's activities and goals are published.

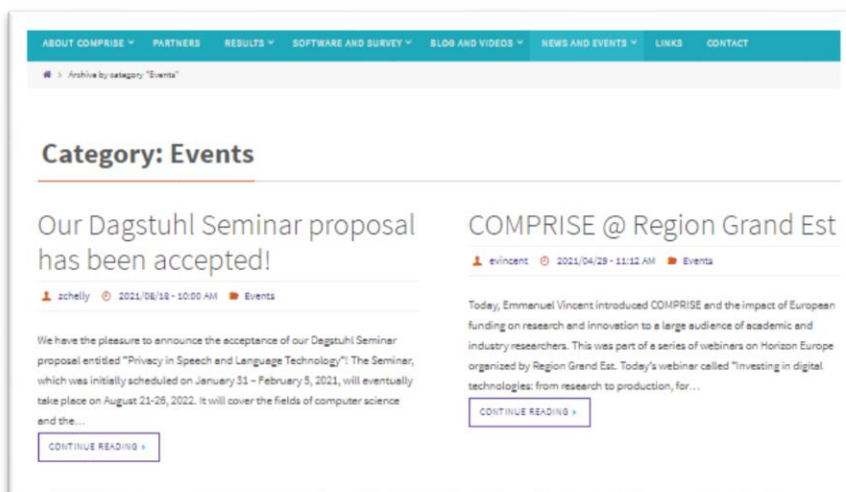


Figure 13: Events page.

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



Figure 14: Links page.

- **Contact:** In this section the contact details of the Project Manager and COMPRISE's social networks are listed.

3.5.2. COMPRISE website impact

To analyse the flow and interests of the visitors to our website, we are using MATOMO, a free and open-source application that tracks web visitors in real-time and display reports for analysis.

Table 11 lists all relevant data obtained from the website's launch until September 30, 2021.

Table 11: Overview of visits.

Overview of visits	
Visits	11,661
Unique visitors	3,792
Average length of visit	2 min 31 s
Bounced visits (i.e., left the site after viewing a single page)	58%
Visitor actions (i.e., page views, downloads, outgoing links and internal searches within the site)	2,7
Average generation time	0.32 s
Page views	26,805
Unique page views	19,828
Total searches on your website	4
Downloads	1,734
Unique downloads	1,547
Outbound links	2,795
Unique outbound links	2,475
Maximum number of shares in one visit	181

As we can see in Table 11, the number of unique visitors is 3,792 (1,708 in the previous reporting period). The average time spent on the website is 2 min 31 s.

The bounce rate, that is the percentage of users who enter directly from www.com-priseh2020.eu and do not interact on the page, is 58%, which is considered an average rate by website marketing specialists.⁵

MATOMO also analyses the number of unique downloads on the website. By September 30, 2021, there have been a total of 1,547 unique downloads (596 in the previous reporting period). These downloads have been mostly deliverables of the project, but we can also highlight one of the infographics uploaded to the website. The documents with the highest number of downloads are listed in Table 12.

⁵ <https://www.brafton.com/blog/strategy/brafton-2017-content-marketing-benchmark-report/>

Table 12: Most downloaded documents.

Document	Downloads
D5.1 “Data Protection and GDPR Requirements”	214
D7.1 “Dissemination and Communication Action Plan”	139
D2.1 “Baseline speech and text transformation and model learning library”	112
10 things you should know about voice technologies	68
D1.3 “Initial Data Management Plan”	62
D7.3 “First dissemination and communication report”	45
D4.1 “Initial weakly supervised learning library”	38
D5.2 “Platform hardware and software architecture”	33

Table 13 contains the metrics of the pages within the COMPRISE website (content, categories, where visitors click).

Table 13: Metrics.

URL of the page	Page views	Unique page views	Bounce rate	Average time on page	Output percentage	Average generation time
/index	7,488	5,743	48%	44 s	60%	0.31 s
comprise	2,818	1,891	51%	59 s	44%	0.20 s
category	2,383	1,502	39%	57 s	26%	0.25 s
deliverables	1,362	1,146	49%	1 min 38 s	65%	0.15 s
partners	1,221	1,015	70%	1 min 21 s	54%	0.21 s
personal-data-concept-and-categorisation	1,098	960	69%	1 min 6 s	96%	0.41 s
videos	891	661	66%	1 min 52 s	67%	0.21 s
papers	858	332	10%	35 s	47%	0.85 s
software	835	486	56%	1 min 48 s	68%	0.19 s
voice-based-applications-for-e-health	690	612	84%	1 min 6 s	95%	1.98 s
10-things-you-should-know-about-voice-based-technologies	539	500	76%	58 s	91%	0.97 s

As indicated in Table 13, the pages of the website related to project results (deliverables, software, and papers) are receiving the most views. Also, the high position of the videos page could be linked to the visitor's interest in project results, as the videos explain in a

user friendly way the functionalities of the COMPRISE assets as well as the problems they are trying to solve.

Finally, the entry points of visitors to the COMPRISE website are listed in Table 14.

Table 14: Channel type.

Channel type	
Direct entries	6,052; 51.09% of visits
From internet sites	954; 8.05% of visits
From search engines	3,501; 29.56% of visits
From social networks	1,140; 9.62% of visits
From campaigns	4; 0.03% of visits
Different search engines	9; 0.08% of visits
Different social networks	8; 0.07% of visits
Different internet pages	156; 1.32% of visits
Different keywords	19; 0.16% of visits
Different campaigns	2; 0.02% of visits

As indicated in Table 14, the highest proportion of visits are direct entries to the website (51%). This is followed by approximately 30% visiting via search engines results, using keywords such as Comprise H2020, videos 2020, etc.

Approximately 10% of the visits have been made via social networks. Table 15 provides the details of which social network visitors came from and their actions while on our website.

Table 15: Social networks.

Social network	Visits	Actions	Actions per visit	Average time on site	Bounce rate
LinkedIn	416	775	1.9	1 min 55 s	70%
Twitter	392	1,088	2.8	2 min 10 s	50%
Facebook	261	388	1.5	39 s	85%
YouTube	34	126	3.7	7 min 32 s	41%
Hacker News	25	41	1.6	26 s	84%
GitHub	9	20	2.2	2 min 6 s	56%

3.6. Social media

The aim of activities carried out on social media are to attract visitors to the COMPRISE website, disseminate the project's results, contact stakeholders, and inform the wider

audience about the project's progress and events in which the partners participate to disseminate our project results.

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

The actions carried out and the results obtained in each of the COMPRISE social network platforms, from the beginning of the project until Month 34 (September 2021), are as follows.

3.6.1. Twitter

A total of 426 tweets with 274,771 impressions (the number of times that COMPRISE's 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 521 followers and has been visited 18,774 times.

3.6.2. LinkedIn

A total of 210 posts with 13,000 estimated impressions (the number of times that COMPRISE's LinkedIn posts have shown up in someone's timeline) have been published through the COMPRISE LinkedIn account (<https://www.linkedin.com/company/comprise-h2020/>). Additionally, the COMPRISE profile has 172 followers.

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

Currently, there are 25 videos produced by COMPRISE and uploaded to the COMPRISE YouTube channel (with a total of 2,394 views). Additionally, three videos from other channels that feature COMPRISE can be found in the "PLAYLISTS" section of the channel. These additional videos are from "France 3", "Rue89 Strasbourg" and "CNIL" (with a total of 543 views) where COMPRISE's coordinator Emmanuel Vincent has participated to disseminate the project.

3.7. Generation of content

A variety of material has been created to publicise the project. This section lists all the items that have been created to disseminate the COMPRISE project.

3.7.1. Videos

25 videos have been produced during the project (until Month 34):

- 1 introduction video providing an overview of the COMPRISE project (produced by ROOT)
- 3 videos showing the problems addressed by COMPRISE (produced by [Vikki Academy](#))
- 7 videos explaining the functionalities of COMPRISE assets and how they solve these problems (produced by ROOT)
- 1 video interview with Marc Tommasi, who explains how privacy is ensured by COMPRISE.
- 11 videos of partners' presentations about COMPRISE results in different conferences
- 3 "how-to" videos explaining how to install and run three assets: the COMPRISE Text Transformer, the COMPRISE Voice Transformer, and COMPRISE Weakly Supervised STT.

Table 16 list the videos produced to date and their corresponding links. The videos in the green cells have been produced during the final reporting period.

Table 16: Videos.

Video title	Produced by	Link
What is COMPRISE?	ROOT	https://www.youtube.com/watch?v=asHD8hBcPPA
Problem 1: Lack of Privacy	Vikki Academy	https://www.youtube.com/watch?v=gm2cC8za8Us
Problem 2: Non-inclusiveness	Vikki Academy	https://www.youtube.com/watch?v=-HvADcfEOuE
Problem 3: Cost-effectiveness	Vikki Academy	https://www.youtube.com/watch?v=5LQb9X3RtUs
How will COMPRISE ensure privacy	ROOT/ INRIA	https://www.youtube.com/watch?v=_LE7XbNPjSs
COMPRISE SDK	ROOT	https://www.youtube.com/watch?v=bg0HNFo-25k
COMPRISE Voice Transformer	ROOT	https://www.youtube.com/watch?v=kh8no66BSDM
COMPRISE Text Transformer	ROOT	https://www.youtube.com/watch?v=q6XGwqdLC8c
COMPRISE Cloud Platform	ROOT	https://www.youtube.com/watch?v=VhCZ4BIUUJ0
COMPRISE Weakly Supervised NLU	ROOT	https://www.youtube.com/watch?v=J0I5vx3q6Nc
COMPRISE Weakly Supervised STT	ROOT	https://www.youtube.com/watch?v=JpVWDNYAzOQ
COMPRISE Speech-to-Text Translation	ROOT	https://www.youtube.com/watch?v=nifbgM1kOhA
Evaluating voice conversion-based privacy protection against informed attackers (ICASSP 2020 paper)	INRIA	https://www.youtube.com/watch?v=i17FMead47M

Video title	Produced by	Link
Privacy guarantees for de-identifying text transformations (Interspeech 2020)	USAAR	https://www.youtube.com/watch?v=DHoIUtEJ7oM
Introducing the VoicePrivacy initiative (Interspeech 2020)	INRIA	https://www.youtube.com/watch?v=wMwK6BTw_N4
On semi-supervised LF-MMI training of acoustic models with limited data (Interspeech 2020)	INRIA	https://www.youtube.com/watch?v=TxbQ65SgVyk
A comparative study of speech anonymization metrics (Interspeech 2020)	INRIA	https://www.youtube.com/watch?v=FqEZV2XVxbA
Design choices for x-vectors based speaker anonymization (Interspeech 2020)	INRIA	https://www.youtube.com/watch?v=9CibEQrBAmI
Achieving multi-accent ASR via unsupervised acoustic model adaptation (Interspeech 2020)	INRIA	https://www.youtube.com/watch?v=-11kt1OFdFU
Assessing unintended memorization in neural discriminative sequence models (TSD 2020)	USAAR	https://www.youtube.com/watch?v=ODwBOulvb5Y
Investigating the impact of pre-trained word embeddings on memorization in neural networks (TSD 2020)	USAAR	https://www.youtube.com/watch?v=_8SFtgJFI8k
Distant supervision and noisy label learning for low resource named entity recognition (AfricaNLP / PMLD4C 2020)	USAAR	https://www.youtube.com/watch?v=zPOdiuvjIGc
Preventing author profiling through zero-shot multilingual back-translation (EMNLP 2021)	USAAR	https://www.youtube.com/watch?v=cPYjL7ZKC3Q
COMPRISE Text Transformer Installation + Use	USAAR	https://www.youtube.com/watch?v=VES0iizdv4w
COMPRISE Voice Transformer Installation + Use	INRIA	https://www.youtube.com/watch?v=fj4cH3Rotns
COMPRISE Weakly Supervised STT: Installation + Use	INRIA	https://www.youtube.com/watch?v=s0Hw07qzw8c

It is planned that partners responsible of each of the software assets produce at least one “How-To Video” for each of these assets.

3.7.2. Leaflet

A leaflet has been designed to present the project in an attractive and understandable way. 300 units of the leaflet have been printed and handed to the project partners (depending on the number requested by each of them) to be distributed at events.



Figure 15: Leaflet outer page (part 1/2).

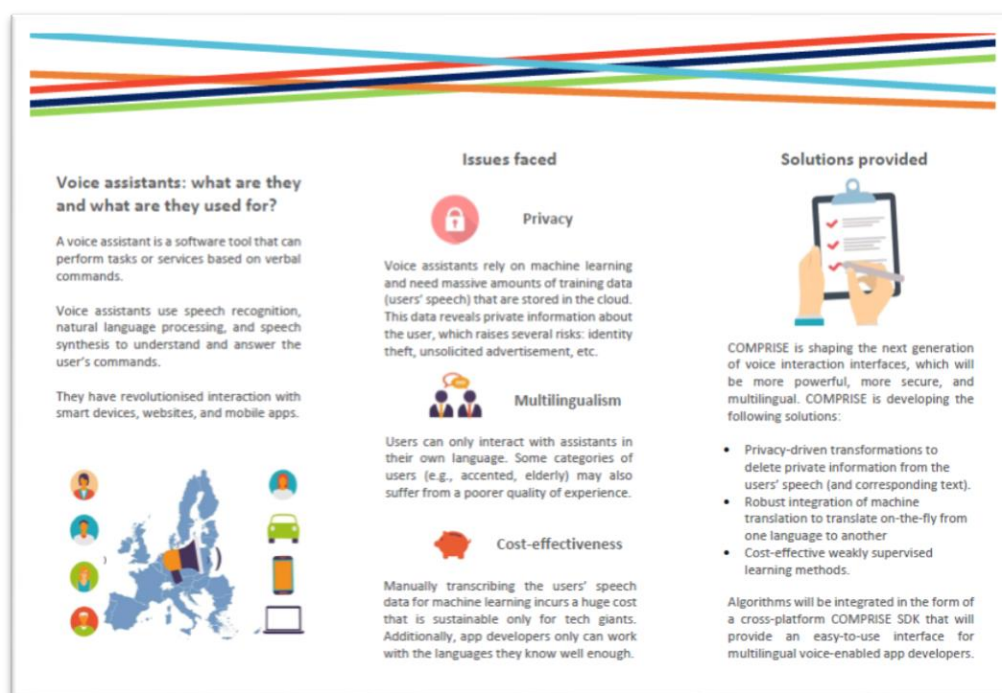


Figure 16: Leaflet inner page (part 2/2).

3.7.3. Roll-up

A roll-up has been designed and produced to be used at different public events.



Figure 17: COMPRISE roll-up.

3.7.4. Poster

A poster has been designed to disseminate COMPRISE at various public events, and it has been updated during the project lifetime. The version below was shown at META FORUM 2020.

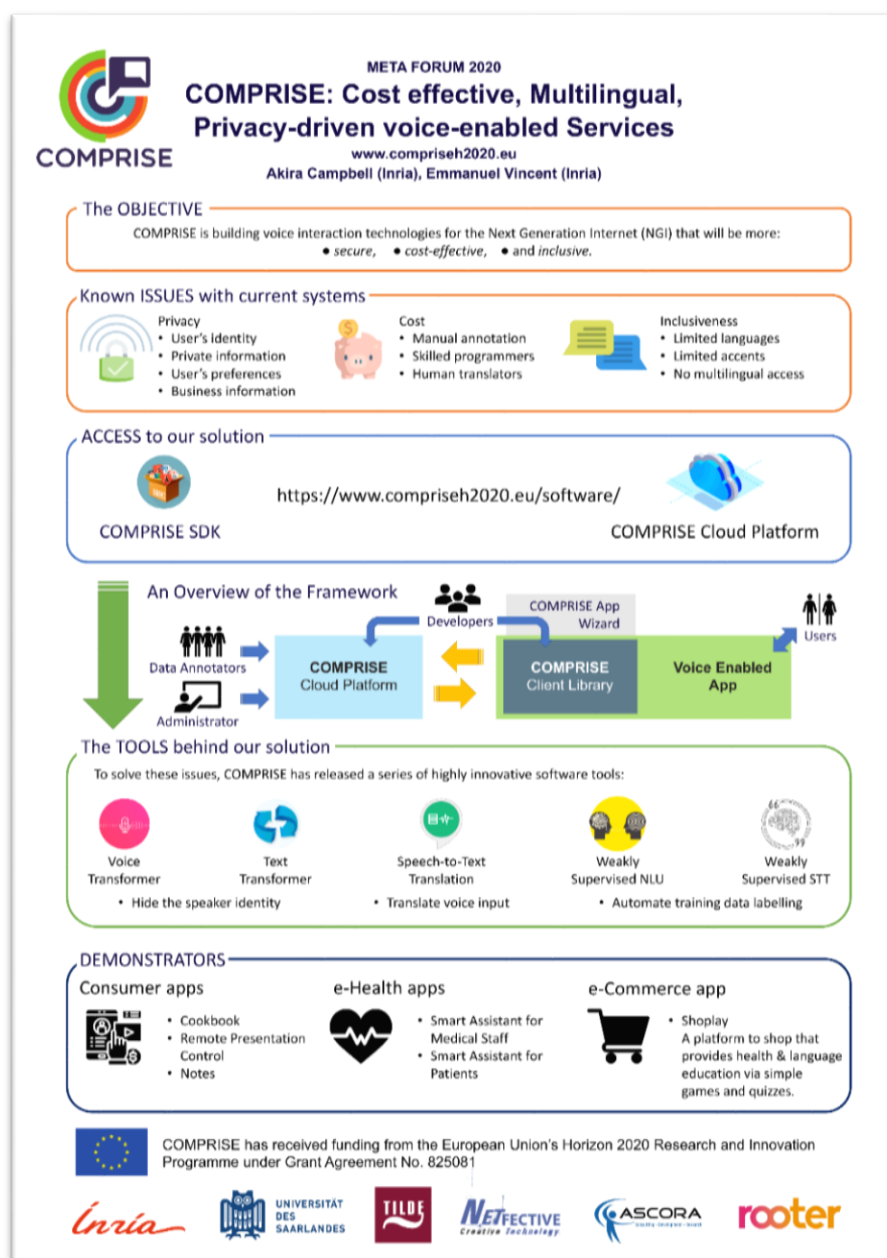


Figure 18: COMPRISE poster.

3.7.5. Infographics

In total 10 infographics have been published in the Blog section of the COMPRISE website. Between Month 1 and Month 16 three infographics were produced:

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

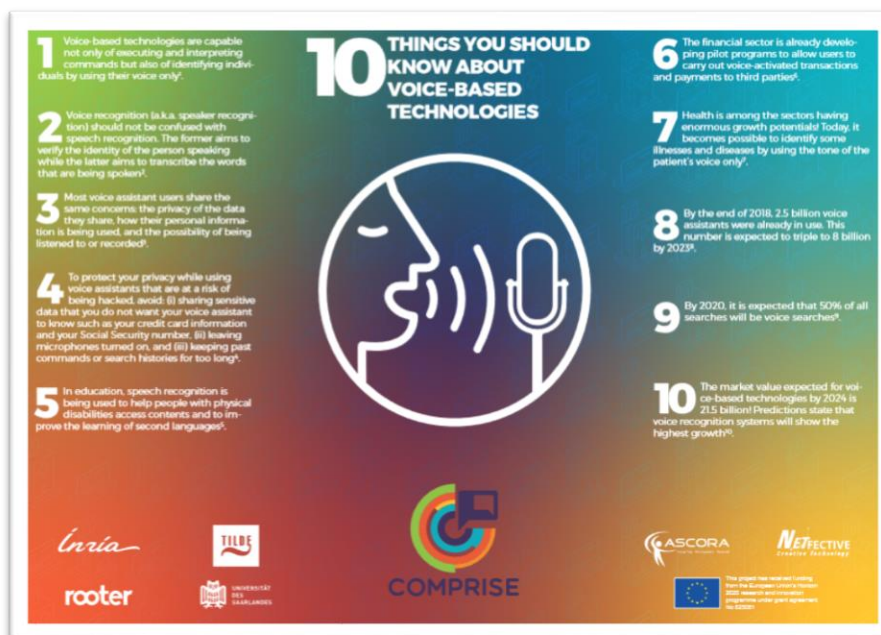


Figure 19: Infographic “10 things you should know about voice-based technologies”.

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



Figure 20: Infographic “10 privacy risks associated with voice-enabled technologies”.

- Business opportunities around voice enabled technologies (<https://www.com-priseh2020.eu/business-opportunities-around-voice-enabled-technologies/>)

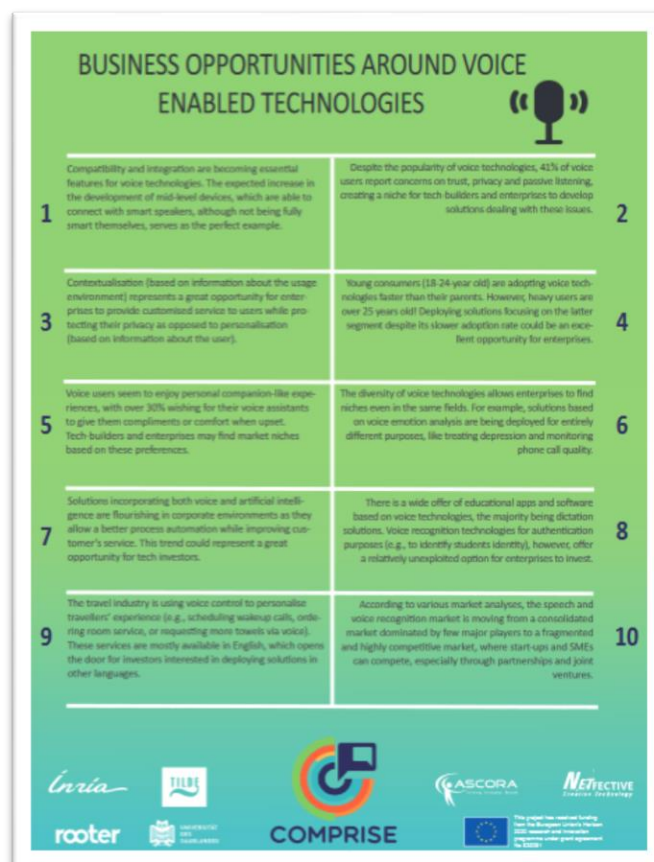


Figure 21: Infographic “Business opportunities around voice enabled technologies”.

During the final reporting period 7 additional infographics have been produced:

- COMPRISE Voice Transformer and Text Transformer (<https://www.com-priseh2020.eu/comprise-features-two-main-branches/>)



Figure 22: Infographic “COMPRISE Voice Transformer and Text Transformer”.

- COMPRISE SDK (<https://www.compriseh2020.eu/comprise-sdk/>)



Figure 23: Infographic “COMPRISE SDK”.

- COMPRISE Weakly Supervised NLU (<https://www.compriseh2020.eu/comprise-weakly-supervised-nlu/>)

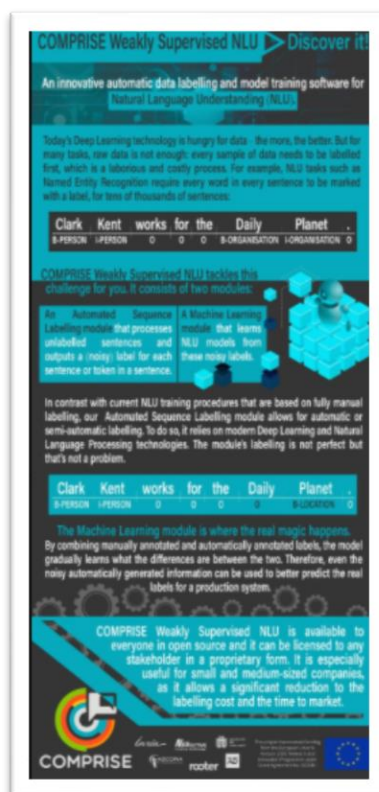


Figure 24: Infographic “COMPRISE Weakly Supervised NLU”.

- COMPRISE Cloud Platform (<https://www.compriseh2020.eu/the-comprise-cloud-platform/>)

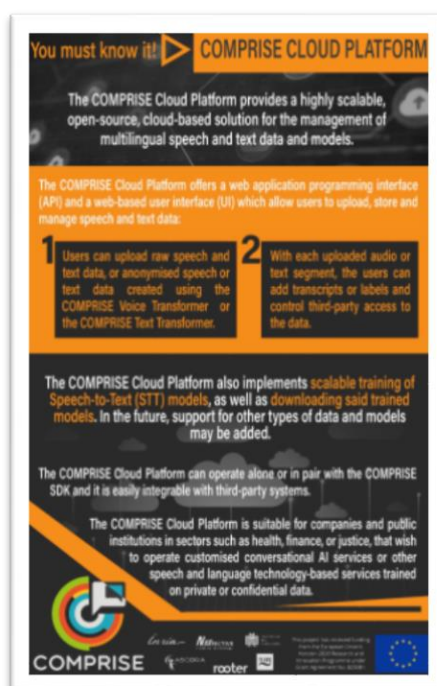


Figure 25: Infographic “COMPRISE Cloud Platform”.

- COMPRISE Speech-to-Text Translation (<https://www.compriseh2020.eu/discover-the-benefits-of-comprise-speech-to-text-translation/>)

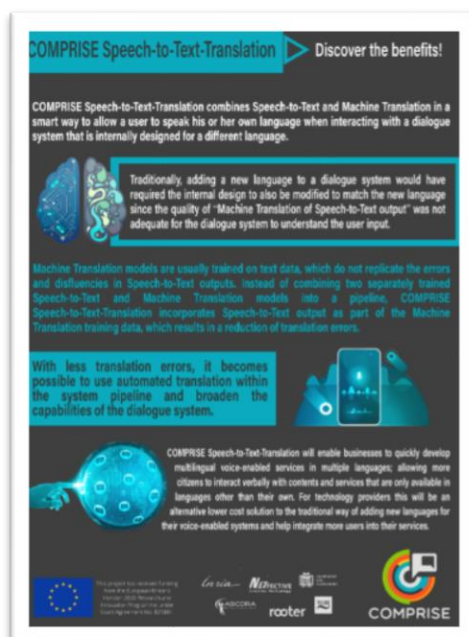


Figure 26: Infographic “COMPRISE Speech-to-Text Translation”.

- The accent gap: What it is, and how to tackle it (<https://www.compriseh2020.eu/the-accent-gap-what-it-is-and-how-to-tackle-it/>)



Figure 27: Infographic “The accent gap”.

- Reducing the cost of training models with Weakly Supervised Learning (<https://www.compriseh2020.eu/reducing-the-cost-of-training-models-with-weakly-supervised-learning/>)



Figure 28: Infographic “Reducing the cost of training models with Weakly Supervised Learning”.

3.7.6. Blog articles

Besides the infographics, 27 articles have been published in the Blog section of the COMPRISE website. Each blog post has been written by partner, taking into consideration the author's (partner) expertise and role within the project. During the first period of the project (between Month 1 and Month 16), 12 blog posts were published:

- 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/>)

During the final reporting period 15 additional blog posts have been published:

- How AI enabled language technologies can help citizens in a global pandemic (<https://www.compriseh2020.eu/how-ai-enabled-language-technologies-can-help-citizens-in-a-global-pandemic/>)
- 4 reasons why you should adopt the COMPRISE SDK (<https://www.compriseh2020.eu/4-reasons-why-you-should-adopt-the-comprise-sdk/>)
- Usage of speaker embeddings for more inclusive Speech-to-Text (<https://www.compriseh2020.eu/usage-of-speaker-embeddings-for-more-inclusive-speech-recognition/>)
- Sources of personal information in voice technologies (<https://www.compriseh2020.eu/sources-of-personal-information-in-voice-technologies/>)
- “Could you please take a (secret) note?” – an example for the usage of the COMPRISE ecosystem (<https://www.compriseh2020.eu/could-you-please-take-a-secret-note-an-example-for-the-usage-of-the-comprise-ecosystem/>)
- Is the long-awaited worldwide communication without borders a distant dream or reality? (<https://www.compriseh2020.eu/is-the-long-awaited-worldwide-communication-without-borders-a-distant-dream-or-reality/>)
- Voice-driven features – Advanced technologies for added-value services (<https://www.compriseh2020.eu/voice-driven-features-advanced-technologies-for-value-added-services/>)
- Cost effective Speech-to-Text with Weakly and Semi Supervised Training (<https://www.compriseh2020.eu/cost-effective-speech-to-text-with-weakly-and-semi-supervised-training/>)
- The problem of complete, irreversible anonymisation (<https://www.compriseh2020.eu/the-problem-of-complete-irreversible-anonymisation/>)
- Are voice-based technologies still important in 2021? (<https://www.compriseh2020.eu/are-voice-based-technologies-still-important-in-2021/>)
- DevOps & COMPRISE to deliver best-in-class apps! (<https://www.compriseh2020.eu/devops-comprise-to-deliver-best-in-class-apps/>)
- Ethics in voice technologies (<https://www.compriseh2020.eu/ethics-in-voice-technologies/>)
- Protecting user privacy with voice conversion (<https://www.compriseh2020.eu/protecting-user-privacy-with-voice-conversion/>)
- Reducing the cost of localized Named Entity Recognition (<https://www.compriseh2020.eu/reducing-the-cost-of-localized-named-entity-recognition/>)
- Speech recognition in Estonian courts (<https://www.compriseh2020.eu/speech-recognition-in-estonian-courts/>)

3.8. Press releases and media activities

3.8.1. Press releases

7 press releases have been produced and published on the COMPRISE website as well as on the websites of the project partners. The press releases have been disseminated from our social media accounts and in some digital journals.

Between Month 1 and Month 16, 2 press releases were published:

- Towards the next generation of voice interaction technology (https://www.compriseh2020.eu/files/2019/01/PR_COMPRISE_110419_for-consortium.pdf)

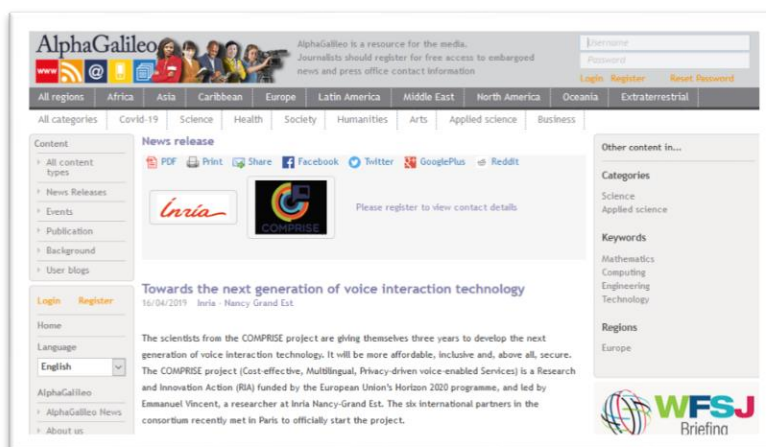


Figure 29: First press release.

- Speech data privacy: a real challenge! (<https://www.compriseh2020.eu/speech-data-privacy-a-real-challenge/>)



Figure 30: Second press release.

During this reporting period 5 additional press releases have been produced and published:

- Privacy-driven speech and text transformation tools (<https://www.compriseh2020.eu/our-first-privacy-preservation-tools-have-been-released/>)



Figure 31: Third press release.

- COMPRISE SDK has been released! (<https://www.compriseh2020.eu/comprise-sdk-has-been-released/>)

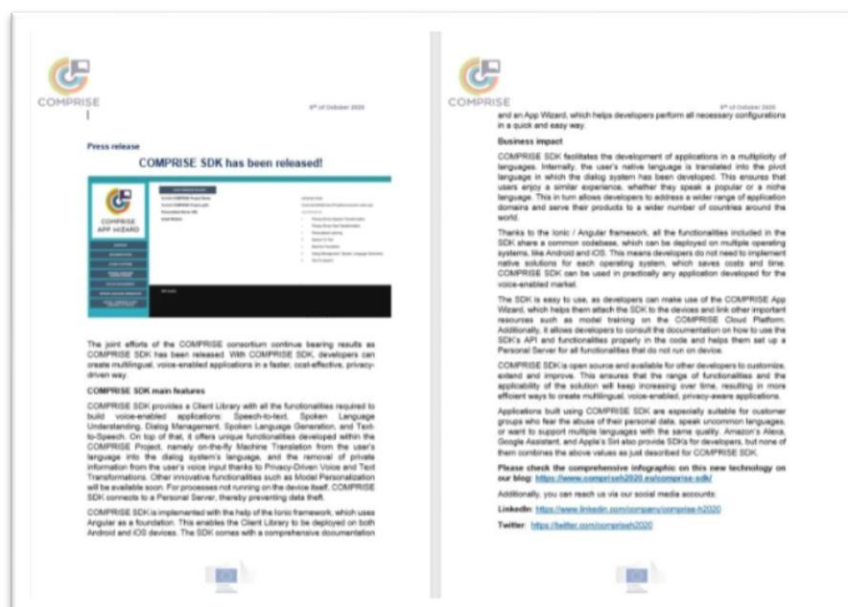


Figure 32: Fourth press release.

- COMPRISE Cloud Platform has been released! (<https://www.compriseh2020.eu/comprise-cloud-platform-has-been-released/>)



Figure 33: Fifth press release.

- The COMPRISE Weakly Supervised Speech-to-Text and COMPRISE Weakly Supervised Natural Language Understanding Tools have been released! (<https://www.compriseh2020.eu/comprise-weakly-supervised-speech-to-text-and-comprise-weakly-supervised-natural-language-understanding-tools-have-been-released/>)



Figure 34: Sixth press release.

- The COMPRISE Speech-to-Text Translation has been released!
(<https://www.compriseh2020.eu/the-comprise-speech-to-text-translation-has-been-released/>)

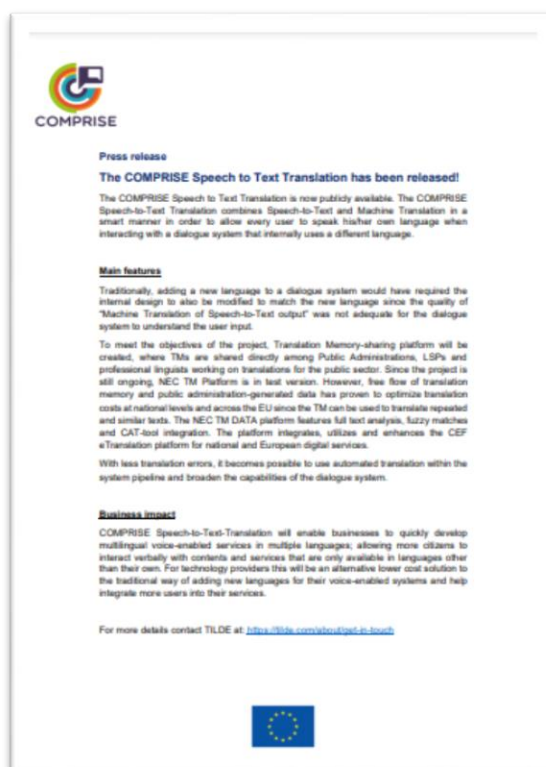


Figure 35: Seventh press release.

It is planned that three additional press releases (each of them about one of the demonstrators) will be produced in November 2021.

3.8.2. Other media activities

During the project, the coordinator of COMPRISE has communicated the project's activities and results through different media channels. These activities that were broadcast though traditional media channels can be found in Table 17 below. The items listed in the green cells have been carried out during the last reporting period.

Table 17: Other media activities.

Communication activity	Media channel	Related link
Report/ Interview	France 3 TV	https://www.compriseh2020.eu/comprise-france-3-tv-channel/
Press article	L'Usine Nouvelle	https://www.usinenouvelle.com/l-usine-nouvelle-du-27-juin-2019-n3617,3340
Broadcasted debate	Rue89 Strasbourg	https://www.compriseh2020.eu/comprise-rue89-strasbourg/
Interview	Sapience (radio)	https://radiovillageinnovation.com/broadcast/85422-Protection-de-la-vie-priv%C3%A9e-2-outils-de-transformation-de-la-voix-et-de-texte

Communication activity	Media channel	Related link
Interview	Sapience (radio)	https://radiovillageinnovation.com/broadcast/85436-PROTECTION-DE-LA-VIE-PRIV%C3%89E-2%C3%A8me-PARTIE-LE-TRANSFORMATEUR-DE-TEXTE-LE-CONSORTIUM-COMPRISE

4. KPIs

The status of the KPIs set for communication and dissemination activities are listed in Tables 18 and 19.

Table 18: Dissemination KPIs.

Dissemination KPIs	Status
<i>KPI-DISS-1: At least 3 journal papers per year on average</i>	Currently, 4 journals papers have been published. A further 2 have been submitted and 3 are in final preparation (see Section 3.4).
<i>PI-DISS-2: At least 8 conference or workshop papers per year on average</i>	Currently, 19 conference papers have been published. A further 3 have been submitted. and 2 are in final preparation (see Section 3.4).
<i>KPI-DISS-3: At least one collaboration with another H2020 project under this call.</i>	Collaborations have been established with ELRC, ELG, LT BRIDGE, and SPRING to different degrees (see Section 3.2).
<i>KPI-DISS-4: At least 500 downloads for public deliverables and prototypes</i>	So far, we have registered 1,547 unique downloads for COMPRISE's released public deliverables and prototypes.
<i>KPI-DISS-5: At least 1 public communication in a networking event organised as part of relevant associations, for example, META-NET, BDVA.</i>	KPI-DISS-5 has been fulfilled as COMPRISE has been presented in two networking events organised as part of relevant associations (the META Forum and the BDV Summit).
<i>KPI-DISS-6: At least 1 public communication in an EC event on data protection and privacy</i>	COMPRISE has not participated in an EC event on data protection and privacy as these events are typically "invitation only" and cover a disjoint scope from COMPRISE. Therefore, it was decided to submit data protection related articles to specialised legal journals. Two such articles have already been published in the EDPL journal (see Section 3.4).
<i>KPI-DISS-7: 2 workshops organised</i>	8 workshops or conferences have been organised, and 4 more will be organised in the near future, including a Dagstuhl seminar on Privacy in Speech and Language Technology (see Section 3.3).

Table 19: Communication KPIs.

Communication KPIs	Status
<i>PI-COMM-1: Web presence: at least 5000 new visitors per year to project website:</i>	So far 3,792 unique visitors totalling 11,661 visits (see Section 3.5.2).
<i>PI-COMM-2: Project story factsheets: at least 10 story factsheets</i>	10 factsheets have been published (see Section 3.7.5).
<i>KPI-COMM-3: Press releases: at least 10 press release</i>	7 press releases have been published and another 3 will be published in November 2021 (see Section 3.8.1).
<i>KPI-COMM-4: Media materials: at least 5 project videos</i>	12 project videos, as well as 11 videos regarding project results presented at various conferences, have been produced (see Section 3.7.1).
<i>KPI-COMM-5: Twitter activity: at least 4 tweets per month</i>	On average 14 tweets per month have been published.

5. Specific actions to attract software developers and final users

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, during the reporting period, several actions have been reinforced to reach the goal of creating a community of software developers to exploit COMPRISE results.

5.1. Website

As explained in Section 3.5 the COMPRISE website has been restructured and redesigned to make it more attractive and user friendly to software developers. In the new website, COMPRISE software assets have been made more visible and accessible so software developers can easily find the information they need to understand and use the assets created by COMPRISE.

5.2. Re-organisation of Gitlab

The Gitlab repository has been reorganised to offer a better experience to developers. It now includes a separate top-level public entry for every COMPRISE asset:

<https://gitlab.inria.fr/comprise>.

5.3. Proper documentation

Proper documentation on how to install COMPRISE components, how to use them and their functionalities has been provided in GitLab:

<https://gitlab.inria.fr/comprise>.

5.4. How-to videos

How-to videos have been recorded and published (see Section 3.7.1). Currently three such videos have been prepared. The remaining videos that present how to install and use the components will be made available by the end of the project.

5.5. Developer survey

As mentioned in Section 3.5 a survey addressed to software developers has been uploaded to the website with the purpose of better understanding the difficulties faced when integrating voice-enabled features in their apps.

5.6. Websites for COMPRISE assets

Each COMPRISE partner has created a website to promote and exploit the software assets (or products that incorporate these assets) for which they are responsible, as listed in Table 20. These websites provide useful information for software developers and/or final users about the software assets, including the functionalities, benefits, and/or how to use and install them.

Table 20: Exploitation websites.

Software Asset	Website
Overall COMPRISE solution	https://www.bluage.com/solutions/comprise-offer
COMPRISE SDK	https://comprise-sdk.ascora.eu/
COMPRISE Cloud Platform	https://tilde.com/research/Tilde-ASR-Training-platform
COMPRISE Voice Transformer	https://team.inria.fr/multispeech/software/comprise-voice-transformer/
COMPRISE Weakly Supervised STT	https://team.inria.fr/multispeech/software/comprise-weakly-supervised-stt/
COMPRISE Text Transformer COMPRISE Weakly Supervised NLU	https://www.lsv.uni-saarland.de/research/comprise-cost-effective-multilingual-privacy-driven-voice-enabled-services/the-comprise-text-transformer/
COMPRISE Speech-to-Text Translation	https://tilde.com/research/Tilde-Speech-Translation
TILDE Hospital Concierge	https://tilde.com/research/Hospital-Concierge

5.7. Materials

The communication materials produced during this reporting period aim to promote the different COMPRISE software assets and explain in a user-friendly way their functionalities. In this sense, for each software asset one video and one infographic have been produced and uploaded to the COMPRISE website (e.g., when a visitor is checking an asset in the Software section of the website they can directly access a video that explains the asset's functionalities) and they have also been disseminated from our social media platforms and in various different discussion groups mentioned in Section 3.1.2.

5.8. Other actions to attract developers

Other actions also have been carried out to attract software developers and final users during this reporting period. Some of them are listed below:

- identifying events for software developers where COMPRISE results can be disseminated;
- placing special attention on developers' and potential final users' profiles when targeting potential followers on social media;
- joining software developer discussion groups on LinkedIn to disseminate the results of COMPRISE.

6. Conclusions

This deliverable provides a summary on the dissemination and communication actions carried out from Month 1 (December 2018) to Month 34 (September 2021) of the project, highlighting the activities carried out in the last reporting period (from April 2020 to September 2021).

Most of the KPIs defined in the DOA regarding communication and dissemination activities have already been fulfilled, and some of them have clearly been exceeded. The KPIs related to conference and journal publications are expected to be fulfilled by the end of the project, as several papers have already been submitted and several others are in preparation. The KPI set for the website (number of unique visitors per year) is the only one which we do not expect to reach. Although several actions have been carried out to increase the number of visitors (new design and structure of the website, strong production and dissemination through several channels of contents directly linked to the website, etc.), technical limitations of the website have prevented additional SEO actions from being carried out.

During the last reporting period particular emphasis has been made on creating a community of software developers around COMPRISE to facilitate the exploitation of the results. This includes the communication materials created during this period aiming to promote the COMPRISE software results and help developers understand their functionalities and which problems they can solve.

Lastly, there are some pending communication and dissemination actions that will be carried out during the last two months of the project (e.g., press releases, submission of new papers, and activities in social media).