Proposal CTRL+V
(Control + Voice: Automation, Analytics and Acceptability)
Part B: technical description

Eligibility

This proposal builds on results (demonstrated proof of principle) achieved within eligible project(s).	**YES**
Related eligible Project(s) Acronym(s):	**ACTIVATE**
Provide the related eligible Project(s) ID Number(s):	**ERC PoC 875212**
The applicants are the owner or have any necessary agreements with owners of those results.	**YES**

1. Excellence

"Our brand is conversation. Proximity and Service are the two operative words in customer relation: the closer the operators are to the client, the better they understand what the motive is behind the call, and the better service we can deliver" (Frank Guilloteau, VP Customer relation, Mutavie)

Especially now with covid-19 distancing measures, we have all experienced the limits of remote communication: noise, "zoom fatigue", misunderstandings and, as a result, degraded human interaction. This problem has a significant impact for all professionals whose interactions mainly happen remotely. For instance, in the contact center industry, where customer contacts happen mostly by phone, only 44% of customers find phone support satisfying (Invoca, 2021). This is a yearly loss of $102 Bn, in the US only.

Figure 1: Media coverage in Science of our ERC research on voice transformation (Ponsot et al. 2018).

Project CTRL+V proposes to solve this problem by bringing to market real-time voice-transformation technologies able to improve remote communications. We take the contact-center market as a first step to prove the problem/solution fit, before tackling the global market of remote communication for professionals.

Our proposal builds on the scientific and technological results of successive ERC projects CREAM (StG 335536) and ACTIVATE (PoC 875212). Project CREAM (2014-2019) has created new algorithms to control the emotional aspects of speech in real-time conversations: for instance, transform voice to make it sound warmer, smiling and more trustworthy (Figure 1). Project ACTIVATE (2020-2021) has then created the software infrastructure to deploy these voice-transformation algorithms in real-life call-center operations, signed test agreements with two of the biggest Business Process Outsourcers (BPO) worldwide, and created a startup company, AltaVoce (https://alta-voce.tech), to explore the technology's future commercialisation.

Standing at this excellent tipping point, project CTRL+V brings together the AltaVoce startup and their historical academic collaborator CNRS, in order to address the final key technological and business/market steps needed before commercialization and investment-readiness. Importantly, our proposal is backed by a unique commercial agreement between AltaVoce and a major global BPO (Comdata, see support letter p.23), allowing our consortium to deploy project results in a real call-center environment of a potential customer, paving a direct way to commercialization at the end of the project.

1.1. Technological breakthrough

Our technology is real-time voice-transformation software which can augment the value of call-center conversations by adding control on the background noise, intelligibility and emotional quality of the interaction. For instance, our real-time voice transformations may make an angry client's voice 10% less aggressive, reducing employee fatigue at the end of the day, or make an operator's voice 10% more trustworthy, augmenting sales and customer satisfaction after the call.

This technology has a **high degree of novelty** compared to other technologies available or in development. First, because they are transformations**,** our solutions do not *synthesize* artificial sounds, but can be applied to any pre-existing voice, either human (e.g. a real call-center operator) or artificial (e.g. the vocal output of a personal assistant). Second, because they are real-time, they can be applied transparently to any ongoing conversation with no noticeable delay and can be modified to react immediately to context, e.g. the voice of an operator can become more or less smiling depending on how the user interacts with them. Finally, because they are expressive, they allow to augment conversations with appropriate, scientifically-validated emotions or attitudes that have a direct impact on their commercial or societal outcomes, e.g. customer satisfaction or transformation rate.

Compared to other voice technology, our innovation has **three critical competitive advantages**. First, contrary to solutions based on speech synthesis (e.g. Google's assistant voices WaveNet), speech transformations do not require learning voices in advance and can be applied to any vocal identity, natural or artificial, even if they are heard for the first time. Second, contrary to solutions based on pre-recorded speech (e.g. current chatbot solutions), speech transformations do not require recording voices in different emotional versions, and they can be adapted dynamically depending on conversation context. Third, contrary to all of the above, speech transformations can be applied to conversations in any language (be it French, English or Chinese), without having to redevelop a system from scratch when a new market is being targeted.

Our innovation has the **potential for several new applications and functionalities**. First, it will be used to improve B2B services that deliver voice communication over the internet (e.g. Skype-like internet voice calls, call centers, 911 telephone dispatch) allowing them not only to transmit voices but also to transform their emotions in real-time in order to optimize for conversation outcomes that are relevant for their commercial or societal bottom-line. Second, it will be used to improve consumer-electronics systems such as intelligent personal assistants (e.g. Google Duplex, Amazon Echo) or devices for patients with voice impairments (e.g. Tobii Dynavox, Acapela), adapting the output of the speech synthesizer to the user's communication needs.

1.2. Technology feasibility

Our technology is currently at technology-readiness level (TRL) 3, and the present proposal aims to take it to TRL6.

The algorithmic part of our technology was **thoroughly validated in the lab**. Our SMILE algorithm, which is able to simulate the sound of a smile in speech, was validated to operate in real-time and to trigger positive emotional reactions in more than 100 French-language participants tested in at least three different experimental studies (Arias et al. 2018; 2020; 2021). Similarly, our INTELLIGIBILITY algorithm, which is able to enhance parts of speech that makes its more understandable, is the result of 10 years of research (Zorila, Kandia & Stylianou, 2012) and has achieve the highest and second-highest intelligibility gains at two recent international software competitions (Cooke, Mayo & Valentini-Bontinhao, 2013; Rennies et al., 2020). Sound examples of both algorithms are available at https://alta-voce.tech/applications. Finally, our DENOISE algorithm, which is able to reduce background noise to improve the quality of the conversation, is based on the FullSubNet method which has been recently shown to improve on five alternative state-of-the-art methods (Hao, Su, Horaud & Li 2021).

Our recent ERC PoC project has provided **a proof-of-concept of technology deployment in the relevant environment** of the call-center. The two algorithms (smile and intelligibility) are implemented in proprietary software, and integrated in a minimal client application which can be installed remotely in batch on each of the call-center operators' machines (Figure 2). A minimal web-based dashboard allows to monitor the state of each application in the fleet (e.g. who's on and who's off, who's running what transformation, etc.), to remotely change their parameters (e.g. put 50% of the fleet off for a A/B test) and, in subsequent versions, to collect data about conversation outcomes (e.g. sales per hour, per agent and per day) from other sources within the project.

This minimal software infrastructure **has passed technical validation** by IT departments at two of the biggest international call-center companies (Webhelp and Comdata). Our startup has signed commercial agreements with both companies and preliminary tests of the smile and intelligibility effects have been conducted at more than 5 of their industrial sites (for clients incl. Bouygues Telecom, Canal+, Engie and Free; in countries incl. France, Italy, Algeria and Cote d'Ivoire). Based on these first results, AltaVoce's commercial partner Comdata has agreed to extend this test agreement to cover the technological and business/market activities of the current proposal (see support letter p.23)

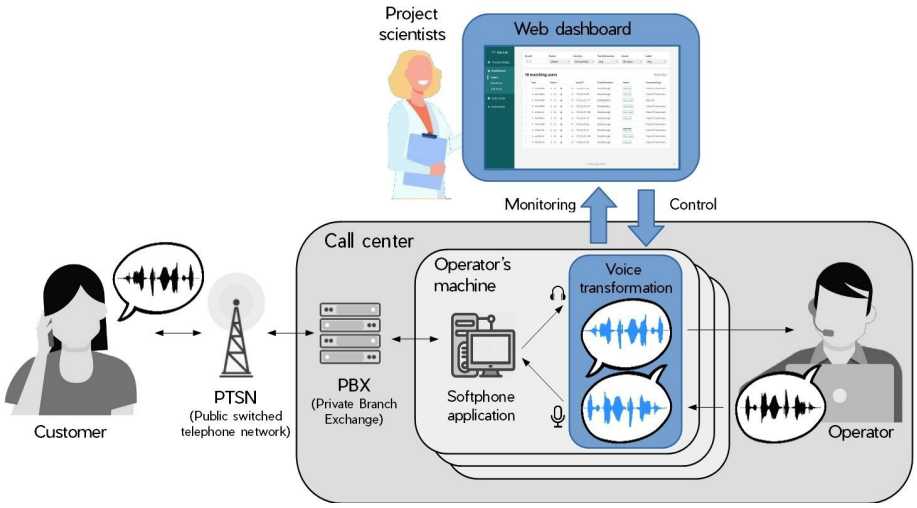
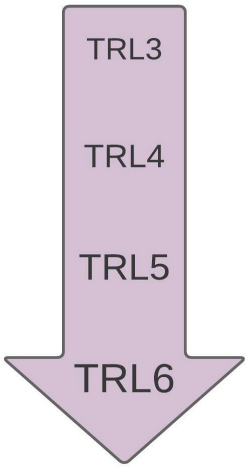


Figure 2: Proof-of-concept of technology deployment in the relevant environment of the call-center. *The application inserts itself as a virtual audio device visible to the telephone software, and replaces either the operator's microphone (to transform the out-calling operator's voice) or the operator's headphones (to transform the in-calling customer's voice). All voice signal processing occurs locally in the application (i.e. not on the cloud), so that private conversation data is not streamed outside of the call-center. The only flow of information between the client application and AltaVoce's servers concern transformation parameters and the monitoring of their activity with a web dashboard.*

Our preliminary tests have shown **potential for application**. For instance, in a recent A/B test conducted at Comdata in May 2021, we evaluated the impact of the intelligibility and smile transformations on the key performance indicators (KPIs) of 30 operators, participating in 13 different sale campaigns over 36 days, totally 5301 hours of conversation data. In roughly half of the campaigns, conversations where the operator's voice was transformed to sound more intelligible or more positive exceeded target sale-rate by up to 47% (intelligibility) and 98% (smile), compared to no-transformation. However, in other campaigns or for selected operators, the added value of the technology stayed minimal, or even slightly negative, resulting on average in a modest ROI as the innovation currently stands (AltaVoce, 2021).

These preliminary tests indicate a clear series of **steps to be taken** before reaching commercialization and investment-readiness:



1. Develop a control layer to automatically adapt the parameters of the voice transformation algorithms and move the system's point of functioning into systematic positive added-value. We have recently performed an experimental proof-of-concept of the feasibility of such a control mechanism (TRL3), showing that manual adaptation of the algorithmic parameters for each operator can improve performance, and that it is possible to automate this adaptation process. Our next step is to validate the technology in the lab, by simulating its impact retrospectively on the test data we have already collected (TRL4).
2. Deploy and technically validate the automatic control mechanism in the call-center environment (TRL5), by integrating it in the software infrastructure we have already validated for the voice transformations.
3. Demonstrate the added-value of the complete system (i.e. voice-transformation algorithms + automatic control) in the call-center environment of external partner Comdata (TRL6).

1.3. Objectives

The **goal of project CTRL+V is to increase the maturity of the technology** beyond the current proof of concept (namely, that voice transformations deployed in the call-center environment can increase KPIs under certain conditions) to a viable demonstrator with demonstrated added-value for our potential customers.

These objectives include **a well-concerted mix of technology development and market/business validation activities**, in three main dimensions:

1. **Automation**: Our current proof of concept applies a single voice transformation (smile or intelligibility) throughout, with overall modest added-value but spectacular performance in some contexts. The first objective of project CTRL+V is to develop a control mechanism, based on AI / machine-learning, which capitalizes on the massive amounts of data (thousands of hours of telephone calls, coupled with voice analytics and KPIs) continuously collected during our operations in order to automatically adapt the voice transformations to the context of the call, and thus drive the system into systematically-high added-value.

When this objective is completed, our technology will be able to learn, for instance, to only apply smile transformations for certain types of sales campaigns (e.g. difficult campaigns with low target sale rate), adapt algorithmic parameters to the acoustic characteristics of the operator's voice (e.g. more intelligibility on deep, male voices) or operator's demographics (e.g. no smile on employees with more than 6-month experience on the job). This objective is plausible because we already have strong evidence of high system performance in certain conditions. It is realistically achievable within the scope of the project because we already have an experimental proof-of-concept, all the data necessary for training and validation, and because our project team is appropriately staffed for machine-learning and voice technology R&D. It will be measured and verified by comparing the added-value of the voice transformations with and without adaptation using A/B tests after that development is done and technically validated.

2. **Analytics**: The second objective of project CTRL+V is to give the system ability to automatically collect and analyse contextual information about the calls and their associated KPIs (by gathering data from the call centers) in order to support both the machine-learning process of Objective 1 and the fine-grained evaluation of the system's added-value for our potential customers.

When Objective 2 is completed, our technology will be able to continuously analyse how its performance is related, for instance, to what operator is giving the call, what category of client is being called, what type of call is being made (e.g. outbound sales or inbound customer support) or what day/week/month a given call is made. In addition, our technology will be able to evaluate its impact, not only on commercial KPIs such as sales-per-hour and transformation rate, but also on human-resources (HR) indicators such as employee fatigue and well-being. This objective is plausible because KPI data at this level of granularity is already collected and exploited by our potential customers. It is realistically achievable within the project because we already have a minimal infrastructure with a web dashboard to collect monitoring data, because a separate commercial agreement with potential customer Comdata enables privileged access to these sources of data, and because our project team is appropriately staffed for software and IT operations (DevOps) and for data analytics/data science. It will be measured and verified by offering our potential customers transparent continuous analytics about system performance in a web dashboard.

3. **Acceptability**: The third objective of project CTRL+V is to validate user adoption and the societal acceptability of the innovation. Because our innovation is completely novel, questions such as *"If given the choice, when do operators decide to turn on the voice transformation?", "do customers find acceptable that their own voices are transformed without their knowing?"*, *"are voice transformations more socially accepted in the content of outbound sales or inbound customer support?"* are impossible to answer without empirical study, especially in global markets with high cultural diversity and rapidly changing attitudes towards new technologies. To do so, project academic partner CNRS will, first, conduct inductive ethnographic survey and interviews with users of the technology (operators, managers and clients of commercial partner Comdata), to document how it transforms their work activity, but also how employee activity can introduce in return changes in the technology (Akrich, 2006). Second, we will conduct empirical ethics experiments on call-center employees and customers, in which participants judge the moral acceptability of text vignettes describing different scenarios of applications (see Bonnefon et al. 2016 for a similar approach on autonomous vehicles). The vignettes will be developed with a graphical designer to improve their clarity, and will be illustrated with real conversation examples and audio examples of voice transformations.

When Objective 3 is completed, we will be able to provide our customers with independent ethical and sociological recommendations into the most socially acceptable way to deploy the innovation in their organization: one that is adapted to their context (e.g. sales or support) and culture (e.g. for their newly opening site in *Côte d'Ivoire*), that is in line with their legislation and public policy, and that is most empowering for their workforce (in a logics of "adoption", rather than simple "adaptation" to the technology). This objective is plausible because it will be conducted by an independent academic partner (CNRS) rather than by the stakeholders of the technology itself (AltaVoce) or potential customers such as Comdata. It is realistically achievable within the project because the academic partner already has extensive experience conducting ethnographic studies of technology adoption (Rey, Simoncini & Triclot, 2021) and empirical ethics experiments (Guerouaou, Vaiva & Aucouturier, 2021) on similar technologies. It will be measured and verified, first, with academic publications and reports and, second, by incorporating results into the Ethics Charter of the AltaVoce company, linked to the company's General Terms and Conditions.

The **potential applications** of these activities are, in the short-term, mainly focussed on the market of customer relation and call-centers, and are two-fold: first, to increase our customers' KPIs by augmenting operator performance (e.g., by increasing sales-per-hour in outbound sale campaigns, or customer satisfaction in inbound customer support) ; second, to reduce our customers' HR/turnover costs by increasing employee well-being and reducing emotional fatigue (e.g. by supporting positive emotional displays in the operator's voice, or by reducing aggressivity in the customer's voice). These applications are **plausible**, because they were designed in concert with major players of the call-center industry, and because they are already being tested in real-operations at the startup's commercial partners Comdata and Webhelp.

Our three objectives are **appropriate for the planned technology development and validation of the innovation :**

1. Objective 1 (Automatization) is the direct answer to the proof-of-concept testing of the technology, which has shown the need for adapting transformations to the call context before commercial value can be created. Software development and R&D activities in Objective 1 will lead to the deployment and technical validation of the automatic control mechanism in the call-center environment (TRL5), by integrating it in the software infrastructure we have already validated for the voice transformations.

2. Objective 2 (Analytics) responds to both the need to collect data for the machine-learning activities of Objective 1, and to validate the system's added-value for our potential customers. DevOps and data-science activities in Objective 2 will lead to demonstrating the added-value of the complete system (i.e. voice-transformation algorithms + automatic control) in the call-center environment (TRL6).

3. Objective 3 (Acceptability) is needed to ensure that the innovation's added-value (created with Objective 1, and validated with Objective 2) is compatible with the societal/ethical imperatives of maintaining diverse, meaningful and empowering work for the call-center workforce, as well as efficient and humane service for call-center customers. Social science activities in Objective 3 will lead to establishing in what work situation the system's added value is societal desirable, and thus to increasing product adoption in potential customers.

At the same time, our three objectives are also ideally positioned **to deal with technical and innovation uncertainties and to enable alternative directions** if needed. First, the machine-learning activities of Objective 1 make the system highly resilient to contextual changes in the industry. Instead of being stuck with a fixed strategy (e.g. increasing smile), the system will be able to quickly adapt to, for instance, changes in audio conditions, e.g. by increasing call intelligibility when the workforce moves to massive telecommuting as we have seen in the covid-19 pandemic. Second, the analytics activities of Objective 2 will make the system more quickly adaptable and broaden the scope of potential KPIs on which it can create value. Instead of being stuck with a single outcome (e.g. sales per hour), the system will be able to monitor its performance in real-time and identify new sources of revenue by, e.g., demonstrating impact on associated HR costs (thus opening further opportunities e.g. in markets which are not primarily revenue-driven, such as public-sector call-centers for 911/emergency services). Third, the social science activities of Objective 3 directly address unknowns in product acceptability, and create a methodology to quickly discover the most acceptable usage scenarios for every new potential customer. Finally, beyond the call-center industry, all three objectives create a more versatile product which can target alternative markets (e.g. B2C digital voice assistants), should the company need to pivot or diversify its activities in the future.

The project's academic partner CNRS is well-experienced in all aspects of **Open Science and research data/output management.** As part of the successive CREAM and ACTIVATE ERC projects, the team has published more than 20 journal articles in gold open access ; published 3 research datasets under creative commons licences ; and published 3 open-source scientific software under the MIT licence, which have been downloaded more than 4000 times (http://cream.ircam.fr). For the current project, we will implement Open Science practices for the activities of Objective 3. In particular, the early sharing of research data and output for experiments in empirical ethics will contribute to the openness of the project's societal objectives and ensure that the company's Ethics Charter fully reflects them. In more details: **Type of data/research outputs:** observational (ethnography), experimental (empirical ethics), est. size: 10Mb. **Findability:** persistent d.o.i. provided by University of Bourgogne Franche-Comté open data portal institutional; data and code stored in github directory. **Accessibility:** MIT Licence/ Creative Commons. **Interoperability:** Raw experimental data and verbatims shared as .csv files; analysis software shared as Python notebooks. **Reusability:** allowed by data-sharing licence; all analysis tools open-source; **Curation and storage:** responsible: academic partner PI; quality assurance by University Open Data portal.

Finally, because 71% of workers in the global call centre industry are female (Belt, 2002), any innovation that introduce changes in call-center work practices can have positive or negative impacts on **sex/gender diversity**. Positively, using voice transformations to support employee performance and reduce emotional fatigue can help break flat career structures and glass ceilings that limit progression. Negatively, by assisting workers in their emotional communication and potentially depriving them of autonomy of deciding e.g. when to smile, the innovation may disproportionately install female workers in routinized and low-status clerical roles (Scholarios & Taylor, 2011). Project Ctrl+V addresses this issue frontally, by including independent and substantial social-science research into how the technology will impact the work, and the sex/gender dimension in particular (Objective 3). The sex/gender dimension will notably be addressed in the company's Ethics Charter, to be drafted at the end of the research. In addition, all recruitments made in the project will conform to the CNRS Gender Equality Plan (Plan d'Action pour l'Egalité, 2021-2023 https://mpdf.cnrs.fr/roadmap/)

2. Impact

The global call-center market is worth $ 340 Bn in 2020, growing at 5% and employing 23 Mn workers (Global Industry Analysts, 2021; Magellan Solution, 2020). The software segment, which is the size of the market tackled by our innovation, is estimated to be worth 20% of the call-center market, i.e. $ 68 Bn. In this market, we estimate that **real-time voice technologies can target a market share of 3%, i.e. $ 18 Bn**, using two alternative methods of calculation: First, given the 15 billion hours of phone calls each year in call centers and a price of $ 1.2 per hour, we estimate that revenues would be $ 18 billion if we reach 100% penetration rate. Alternatively, the same amount is given by considering the average time of an agent speaking over the phone (about 60 hours per month, 11 months), times the 23 Mn of working agents globally.

To reach this market of $18Bn, the most important strategy of Project CTRL+V is **to implement a series of pilot deployments (A/B tests)** in the contact-centers of our commercial partner and potential customer Comdata. During a pilot, call-center operators use the product prototype for their daily calls in e.g. sales campaigns, according to an A/B test protocol: some days the product is ON, some other days it is OFF, based on a random distribution. To avoid any bias, the operators are informed about the test period, but not about the ON/OFF status of the prototype. At the end of the pilot, production KPIs (e.g. sales per hour) are measured. Their values are compared between days ON and OFF of the prototype, in order to quantify the value brought by the technology.

Feedback from the managers and local teams is collected at every step of the test. First, a kick-off meeting is held before the beginning: the main KPIs and characteristics of the campaign are defined, as well as all the elements needed to optimize the experimental design of the A/B test. Second, weekly (remote) meetings with the local teams serve to identify technical issues and check that data collection is properly running. Finally, a test debrief with teams and account managers allows to validate whether the use of the prototype had a significant impact on the KPIs during the campaign, to quantify the ROI, as well as to keep records of key learnings and valuable input for product design.

2.1. Business and Market fit

Pilot deployments will form the bulk of project activities from M6-M18. By providing constant real-world feedback from the contact-center staff and customers, these activities drive the validation and outcomes of the project roadmap. First, pilot deployments **support the development of a business model**, by validating the financial scheme of the commercial product. Given that KPIs are compared with and without the prototype, activities in WP5 will be able to elaborate a pricing scheme which takes into account the return-on-investment (ROI) related to the improvement of the KPIs that is measured during the pilot. Importantly, ROI measured during the project is not only relevant to subsequent commercialization at tested sites in Comdata, but also highly relevant to the global industry, because Comdata employs more than 50,000 employees and operates more than 100 call-centers in 22 countries.

Second, because they use the prototype in actual working conditions, pilot deployments also **support the development of high-value product features**, and generally improve product design, from its overall conception to practical aspects such as installation, security needs, configuration and everyday use. For instance, during our previous proof-of-concept tests at commercial partners Comdata and Webhelp, we were able to improve the usability of the product's web dashboard, e.g. by allowing call-center managers to assign labels to groups of users/operators, or to filter results by sales campaign. Such improvements are essential to increase product value, because they address direct client demands ("what additional features would make you more likely to pay for this product"), rather than our (possibly misguided) predictions of what such demands could be.

Another key insight of project Ctrl+V is that pilot deployments, as implemented here with commercial partner Comdata, also serve as **a customer-approach strategy**. Because our technology is novel, we approach new clients by proposing to test and demonstrate product value by implementing a pilot, without requiring a subsequent commitment to purchase. Pilot deployment is invoiced to the new client with a flexible pricing scheme taking into account the client ROI at the end of the A/B test (i.e. potentially very cheap if the product does not add substantial value). At the end of the pilot, client-specific ROI is used to design a viable pricing scheme, close the sale and extend product deployment to the whole production workforce.

Because they streamline the process of launching and conducting pilots at new clients, project activities in WP2 (automated adaptation) and WP3 (real-time analytics) therefore **directly improve our capacity to commercialize the product**. Commercialization activities (WP5) will start as soon as product milestones are validated during pilot deployments at Comdata (M18), and target not only Comdata but other potential customers in the industry. Our principal customers are BPO (Business Process Outsourcers), 25% of the market. We will primarily target leaders (accounting e.g. for 72% of the market share in France), such as Webhelp (with whom we already have commercial agreements for tests similar as Comdata's) or Teleperformance. In-house call centers are also our customers, in selected industries such as banking, insurance, telecom, luxury goods, energy, polling institutes. Product commercialization will be measured with two metrics: pilot conversion rate (i.e., the number of pilot deployments that are converted in licence agreements) and active users growth rate (i.e., the monthly growth of the number of active users). First, because the duration of a pilot is typically limited to 3 months, conversion rate is directly related to market penetration rate and speed. Second, because every pilot converted in production provides a varying number of active users, active user growth is important to measure and validate the ideal range for the target account size (number of operators taking part in a pilot, and in production).

Finally, because they implement the same series of steps and data exchanges than the product's final commercial deployment, pilot deployments also serve to **support full compliance to the GDPR**, the only regulation impacting our industry and this, ahead of exploitation. In particular, it is to be noted that, because voice transformation are done locally, in client software, and not in the cloud, identifying and confidential conversation data do not need to leave the call-center IT infrastructure in order to be transformed (Figure 1).

The project's **strategy for managing intellectual property** mainly relates to the functional software bricks implemented by the AltaVoce team in WP2 and WP3. The multidisciplinary skills and experience necessary for the development and mastery of project technologies are rare and difficult to replace, and constitute a first barrier to natural entry. However, to strengthen the existing barriers to entry, a key evaluation of innovative results will be performed at MILESTONE 2 (M6) and, if successful, will trigger a patent writing and application (as part of WP5). In addition, we will set up regular monitoring of innovative results throughout the project (e.g. innovations in analytics as part of WP3 during M6-M18), with reports on work progress and results obtained. These reports as well as the coded solutions will be formalized and timestamped by a trusted, third-party IP law consultancy. Note also that, ahead of the project, we have commissioned freedom-to-operate searches on methods related to our principal technologies, which confirmed that the intelligibility and smile enhancement method proposed by AltaVoce do not reproduce claims by known third-party patents.

In addition, all AltaVoce employees recruited on the project will sign an employment contract which includes several clauses protecting the company, such as an obligation of confidentiality, a non-competition clause and an inventive mission clause for people likely to invent in the field of AltaVoce. We also make sure to obtain all the intellectual property rights on any creations of the interns. Finally, confidentiality agreements will be systematically concluded between the three project partners. As part of these partnerships, the source code is never shared among partners.

One critical element of impact for the project is that **it directly involves potential customers to test product value and acceptability.** Technical validation (WP2.2 & WP3.1) and functional tests (WP2.3) will be conducted in the actual call-center environment, during its real operations, and involve hundreds of call-center operators testing the technology. Economic value (impact on KPI) will be directly relevant to the call-center's direct operations. In addition, technology deployment for the tests will involve Comdata's IT services, and provide direct feedback about product acceptability and usability. Finally, academic partner CNRS will also separately assess product's ethical and societal acceptability, using social science experiments involving call-center employees as participants and interviewee. The fact that Comdata is both an external partner in these activities, and in-fine a potential customer, maximizes the commercialization potential of project's results.

In sum, pilot deployments conducted in the project are designed to directly **validate the problem / solution fit**. The main bricks of the prototype are tested individually to get feedback by the operators and management about their ergonomy. Problem/solution fit will be measurable/verifiable within the project by the deliverables of the acceptability activities of WP4.

At the same time, pilot deployments **validate the product/market fit** by quantifying the ROI that the technology guarantees in a production campaign ("is the solution bringing value to the contact center production?"), by automatically adapting them to call characteristics: for instance, amplifying operator's smiles may fit best for a telemarketing campaign, while reducing customers' arousal could be more useful for customer support. Product/market fit will be measurable/verifiable within the project by the ROI measured at the end of the testing period (WP2.3, M18), and by market reactions to the subsequent investment-readiness activities (WP5, M18-M24).

2.2 Economic and/or societal benefits

Our preliminary results deploying voice transformations in real call-center operations have suggested that the technology can, in some contexts, improve KPIs such as sales-per-hour by an average 5-10%. Through controlled experiments and subsequent optimizations, the project should confirm and improve these numbers. This suggests that, upon completion, project Ctrl+V will provide a compelling new opportunity for a +10% increase of revenue for the call-center industry, creating substantial demand and new EU/global markets. Projected to a potential customer such as Comdata France, our impact could scale up to a €30 million increase in revenue; projected to the global market, this is an increase of $1.9 Billion.

Beyond these compelling economic benefits, the proposed innovation is also **expected to generate positive impacts in the employment and societal domains**. The global call-center industry employs more than 23 million workers, in what is often perceived as "white collar sweatshops" where employees are poorly paid and work under stressful conditions. Call-center workers have turnover rates between 30-45%, more than double the average for other occupations, and are absent 8.2 days per year, which is 0.8 day higher than other industries (Doellgast & O'Brady, 2020). By generating novel managerial insights on how voice parameters impact workers physical and emotional fatigue, project Ctrl+V will improve employee well-being. In addition, by issuing socio-technical recommendations on the best deployment of voice technologies, our project will help organizations focus on their capacity to augment rather than merely replace employees. Perhaps even more importantly, because we will involve real call-center employees and customers, our work will also build a method for incorporating ethical and sociological insights into organizational decision-making, legislation and public policy. The project's legacy will be nothing short of fulfilling the promise of a "technical democracy" (Callon, Lascoume, Barthes, 2001), the enrichment of traditional representative democracy with new procedures for consultation and representation involving both experts and non-experts, ordinary citizens and politicians.

2.3 Entrepreneurship

AltaVoce already boasts a **strong and motivated team of founders**, with skills in **technology** (CTO Gilles Degottex, PhD), **product engineering** (CPO Marco Liuni, PhD) and **business development** (CEO Nicolas Martinois, MSc). In addition to the founding team, Project Ctrl+V will fund the hiring of the startup's first two employees, Agate Berriot (DevOps lead) and Mohamed Trabelsi (Application lead).

Transitioning from academic research (Liuni, Degottex) to entrepreneurship requires a shift in mindset, and solid accompaniment. The team is grateful to have benefited from two innovation support programs: Deeptech Founders (https://deeptechfounders.com; Degottex, Liuni): a 6-month entrepreneurial learning program taught by entrepreneurs and major players in Deeptech; RISE (https://www.cnrsinnovation.com/rise; Degottex, Liuni, Martinois): a 12-month startup support program set up by CNRS Innovation, tailored to entrepreneurial projects with the participation of CNRS researchers. Such mentoring will continue throughout the project as, since September 2021, AltaVoce is part of the incubation program Agoranov supporting selected Tech & Science startups (https://www.agoranov.com; more than 460 companies incubated, 2,1B€+ funds raised, 12K+ jobs created, 6 IPO).

2.4 Partnerships and investment-readiness

Project CTRL+V aims to take the technology beyond the current proof-of-concept to **become investment ready**. To do so, a key step is to define a solid product roadmap, according to AltaVoce strategy: bringing to market voice technologies able to improve remote communications. Project activities will demonstrate the problem/solution fit with the global call-center market, with a product in a finished/near-finished state which brings value, generates

interest, attention, and is perceived as such by potential customers. By M18 in the project (Milestone 3), we will have a technically-validated product with both client and web interfaces (WP2 and 3), tested its acceptability in a real-call center environment (WP4), measured its economic value in terms (KPIs) relevant to its potential customers, protected its assets with IPR management (WP5) and widely communicated about it in key industry channels (e.g. trade conventions), all this is in close partnership with potential customers. The last period of the project (M18-24) is devoted to preparing investment-readiness materials (financial plan, pitch deck), in view to start commercial discussions between AltaVoce and companies such as Comdata and Webhelp, and seek VC/BA investment.

For reference, consider the funding trajectory of US-based company Krisp (https://www.crunchbase.com/organization/krisp-af3e). Krisp raised $17.5M from 2018 to 2021, with a product strategy that is close to AltaVoce's, i.e. making remote workers more professional by removing background noise in online conversations. AltaVoce technologies have a significant advance, being able not only to remove noise, but also to control the intelligibility and emotional aspects of speech in real-time. By proving that such technologies bring value to call-centers, Project CTRL+V takes an essential step towards tackling the global remote workers market, and will provide evidence to investors that our ambitions to mimic and exceed the success of start-ups like Krisp are well-founded.

At the end of the project, AltaVoce will **seek seed capital** in exchange of equity stake from business angels (BAs) and venture capitalists (VCs). We will seek to raise an initial 1.5M€, possibly in two stages with an initial 500k€ leveraged by debt. The founding team is in regular contact with VCs and BAs (e.g. XAnge, Karot Capital, Kima Ventures), whom we are keeping informed of our product milestones and trajectory towards investment readiness, and we will be able to activate these contacts at the end of the project.

Because AltaVoce's product can be used simultaneously with other solutions improving efficiency in call-center operations, such as Afiniti ($100M funding) or AlloMedia ($12M), we may be able to involve **key partners in product commercialisation**. However, it is not our current strategy to rely on a third-party distributor to sell our product to their existing client base. In order to grow and improve our product, we need to collect direct feedback from our clients and users, and to have the leeway to update the product's technical architecture or economic model without being tied with a commercial agreement with a distributor.

3. Quality and efficiency of the implementation

3.1 Quality of the team

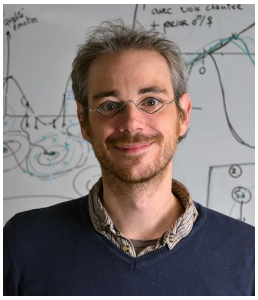
Project Ctrl+V is coordinated by the voice-tech startup AltaVoce, in partnership with the startup's historical academic collaborator (CNRS/FEMTO-ST Institute).

Partner 1, AltaVoce, in Paris (http://alta-voce.tech) is an emerging voice-tech startup specializing in real-time voice transformations for the customer relation and communication industries. Founded in Feb. 2020, AltaVoce originated in research from two successive ERC projects CREAM and ACTIVATE. Project PI **Marco Liuni**, PhD, is Chief Product Officer for AltaVoce. A computer-science researcher turned entrepreneur, Liuni will supervise project management activities (WP1) as well as supervise AltaVoce's involvement in the project in collaboration with co-founders Nicolas Martinois (CEO) and Gilles Degottex, PhD. (CTO).

Together, the AltaVoce team brings to the project **high-quality expertise** in voice and machine-learning R&D (Objective 1), web-service development/operations (Objective 2) and business development. Liuni and Degottex both hold Computer Science PhDs from IRCAM in Paris (one of the largest research centers for music and voice technology), have more than 10 years of research experience in computer audio and have been recipients of some of the world's most competitive postdoctoral scholarships (US Fulbright in University of California Los Angeles for Liuni, EU Marie Skłodowska-Curie in Cambridge University for Degottex). Martinois holds a MBA from EMLyon (France), is a national entrepreneurial contest winner (Campus Creation) and complements the skills of the cofounders' team with 5-year consultancy experience in corporate strategy and business development.
In addition to Liuni, Degottex and Martinois, project Ctrl+V will fund the company's first two employees, Agate Berriot (DevOps lead) and Mohammed Trabelsi (Product development lead).

Because it owns the technology and controls 100% of its share, AltaVoce's team is in **capacity to move decisively** towards the innovation. Its **capacity to attract market interest** and create commercial value from the technology is demonstrated by the fact that the startup has already signed commercial test agreements with two of the biggest

global call-center companies – one of which, Comdata, agreeing to support the current proposal's activities. Finally, its **capacity to develop an attractive business and investment proposition** is demonstrated by the fact that the startup was already selected for support/funding by the CNRS Innovation program, BPI French Tech and the AGORANOV incubator in Paris.



Partner 2, the FEMTO-ST Institute in Besançon and Belfort, France (https://www.femto-st.fr/en) is one of the country's largest engineering and system science research unit, with a research program spanning manufacturing systems, industry 4.0 and the sociology/philosophy of technological innovation. CNRS Senior researcher **Jean-Julien Aucouturier**, PhD., will supervise FEMTO-ST's involvement in the project, together with Assistant Profs. Bénédicte Rey, Mathieu Triclot and Nicolas Simoncini, PhDs. FEMTO-ST's operative institution CNRS (Centre National de la Recherche Scientifique) is France's largest public research operator and will coordinate administrative and financial aspects of FEMTO-ST's participation. (Aucouturier's salary declared by CNRS; Rey, Triclot and Simoncini declared by UTBM, CNRS's linked third-party in co-operating FEMTO-ST).

The FEMTO-ST team brings to the project **high-quality expertise** in cognitive science (Aucouturier), sociology (Rey, Simoncini) and philosophy of innovation (Triclot) which are necessary for the social science activities of Objective 3. A computer-science PhD turned cognitive-science researcher, Aucouturier is the two-time recipient of competitive ERC funding CREAM and ACTIVATE, which led to publications in PNAS (2016,2018), Current Biology (2018, 2021) and Nature Communications (2021). He was awarded the 2018 Prix d'Emergence Scientifique from the French Hearing Foundation. Aucouturier is co-founder with Liuni, Degottex and Martinois of project coordinator AltaVoce, for which it currently acts as scientific advisor (French "loi PACTE"). His team's recent research has included the study of the ethical perception of deep-fakes (Guerouaou, Vaiva & Aucouturier, 2021), which is directly relevant to the ethics activities of Objective 3.

Aucouturier's colleagues Rey, Triclot and Simoncini are members of FEMTO-ST's social-science team RECITS, and are highly experienced in in-situ ethnographic studies of innovation usage and adoption (Rey, Simoncini & Triclot, 2021). They have conducted field work in several technological innovation projects including transition to hydrogen energy in industry and the adoption of health-tech wearables (Al Masry et al., 2021), which are directly relevant to the ethnographical activities of Objective 3. Simoncini is the recipient of the 2020 Turriano Prize of the International Committee for the History of Technology (ICOHTEC).

Because the FEMTO-ST institute and its operating institution CNRS are fully experienced in conducting public-private collaborations, and currently have IPR and data-sharing agreements specifically related to AltaVoce's technical development, the academic team is in **capacity to move decisively** towards the innovation. Because project Ctrl+V includes budget for compensating Comdata employees for non-productive time while answering the surveys and interview of Objective 3, the FEMTO-ST team will be in **capacity** to produce key acceptability insights into the deployment of the technology in the call-center organization, and thus contribute **to develop a societally-appropriate business and investment proposition**.

Comdata

Finally, beyond the project, our proposal is also backed by an existing commercial agreement between AltaVoce and a major global BPO (Comdata, see letter of support p.23), allowing our consortium to deploy project results in a real call-center environment, and paving a direct way to commercialization at the end of the project. The Comdata Group (https://www.comdatagroup.com) is a leading global service provider in customer Interaction, employing more than 50,000 employees and operating more than 100 call-centers in 22 countries. Its local branch Comdata France, represented in our agreement by Sales Manager Vincent Premat, is the country's third biggest BPO, with customer accounts including service providers such as Orange and Free, Electricité de France and Canal Plus. External support from Comdata is a major asset for the project's **capacity to co-develop an attractive business and investment proposition**. Comdata already has a successful work history with AltaVoce through proof-of-concept tests conducted in the scope of our previous ERC PoC project and, being a potential customer for the technology, will bring unique insights into both the operational and the financial/ROI requirements of a successful business proposition.

3.2 Work plan and resources

Our proposed **pathway towards deployment** and investment-readiness includes:

1. The development by AltaVoce of (1) an AI-driven controller to adapt the parameters of the voice transformations and (2) the web analytics front-end and back-end to support its operations, and its validation in the lab (MILESTONE 1: TRL4); the deployment of the system in Comdata for (1) technical validation (MILESTONE 2:TRL5) and (2) at least three A/B tests in the actual call-center environment for economic validation (MILESTONE 3: TRL6)
2. In parallel to the tests, two independent academic study by CNRS to establish acceptable moral/societal usage scenarios, and draft the company's Ethics Charter (for MILESTONE 3)
3. Finally, AltaVoce will use these measures and conclusions to prepare investment-readiness materials, including an updated pricing proposal, update business plan, IPR and scalability plan; engage with commercial discussion with Comdata and other potential customers, and start approaching investors (MILESTONE 5).

The project's milestones are defined to **track essential progress along this pathway**:

MILESTONE 1, at M4: corresponds to the in-house validation of the project's new technological development, and our innovation reaching TRL4. This milestone will trigger IP protection measures, and deployment for technical validation at commercial partner Comdata.

MILESTONE 2, at M6: technical validation obtained from Comdata, and our innovation reaching TRL5. This milestone will trigger the start of A/B tests for economic validation, as well as the start of the field ethnography of product acceptability.

MILESTONE 3, at M18: economic validation obtained, and our innovation reaching TRL6. Conclusions of the ethics study, and v1 of the company's Ethics Charter. This milestone will trigger the preparation of investment-readiness materials and commercial discussions with Comdata and other potential customers.

Work package No	Work Package Title	Lead Participant No	Lead Participant Short Name	Person-Months	Start Month	End month
1	Project management	1.1	Liuni (AltaVoce)	24	M1	M24
2	Automation	1.2	Degottex (AltaVoce)	44	M1	M18
3	Analytics	2.1	Berriot (AltaVoce)	37	M1	M18
4	Acceptability	3.1	Aucouturier (CNRS)	59	M1	M24
5	Investment-readiness	1.3	Martinois (AltaVoce)	19	M4	M24
				Total: 183		

Our workplan is structured in **5 workpackages** (WPs):

WP1 ("Project management") is devoted to elaborating and implementing the project's Tech-to-market transition plan, Data management plan and Dissemination and exploitation plan according to the guidelines for "proactive management" of the EU Directorate-General for Research and Innovation (EU, 2020). It runs throughout the project.

WP2 ("Automation") is devoted to the development (MILESTONE 1), technical validation (MILESTONE 2) and economic validation via A/B testing (MILESTONE 3) of the project's main technological development, an AI-driven controller for voice transformations.

WP3 ("Analytics") is devoted to the development and technical validation of the data monitoring back-end and front-end necessary for the deployment and economic validation of the WP2 technology (MILESTONE 2), as well as front-end development for the product's commercial readiness (MILESTONE 3).

WP4 ('Acceptability") is devoted to the independent empirical study of acceptable product usage, from the point of view of ethics and social science. It precedes and accompanies the A/B tests starting MILESTONE 2, and produces material for investment-readiness at MILESTONE 3, notably a draft of the company's Ethics Charter.

Finally, WP5 ('Investment-readiness') is devoted to the preparation of investors' material following all project activities, including IPR (starting at MILESTONE 1), pricing, updated business plan, and the initiation of commercial and investment discussions (starting at MILESTONE 3)

The project duration is 24 months, divided in two 12-month reporting periods:

Project duration	Number of periods	RP1 duration	RP2 duration
24	2	12	12

The timing of the different workpackages, and how they inter-relate, is illustrated in Figures 3 and 4.

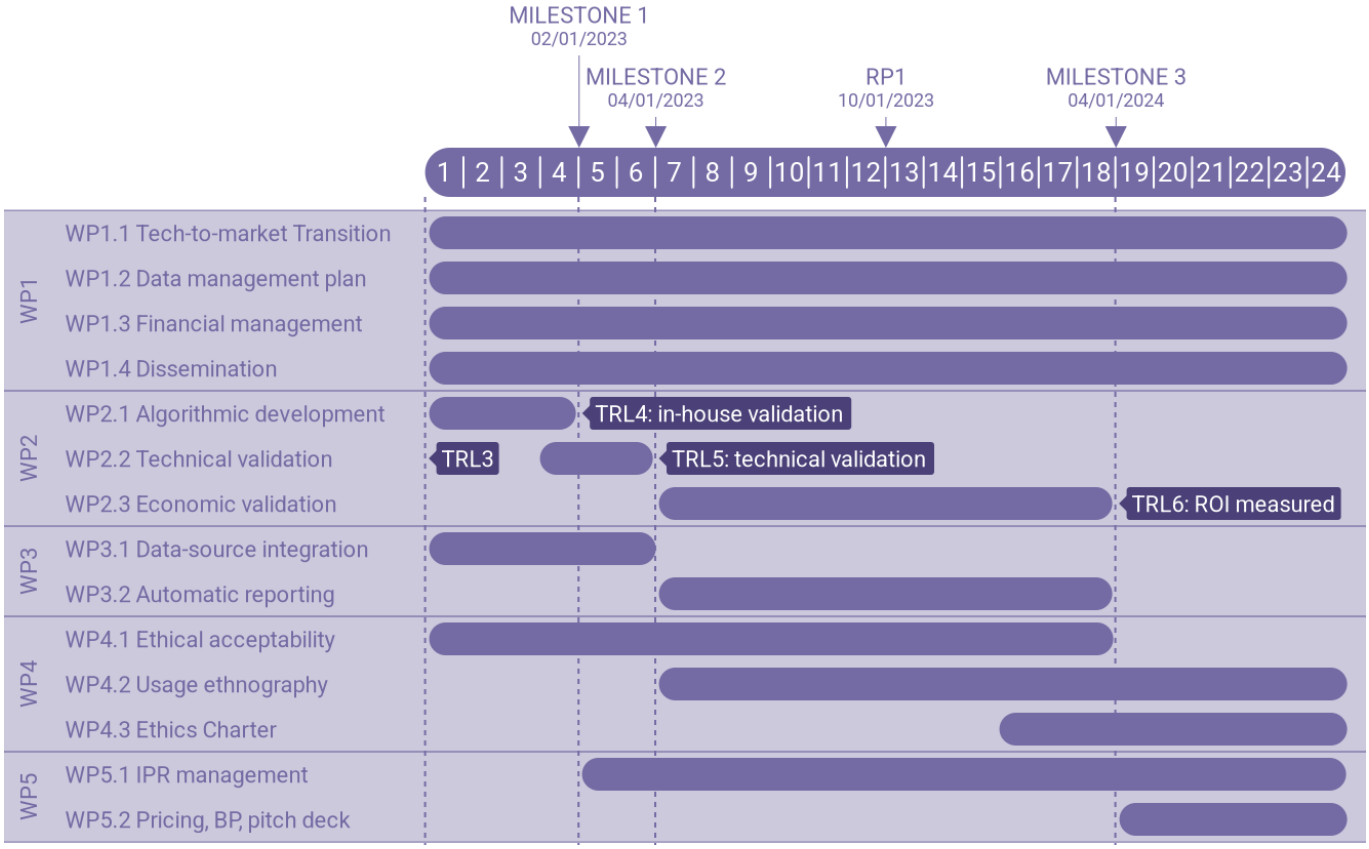


Figure 3 - Project Gantt diagram. WPs 2 (automation), 3 (analytics) and 4 (acceptability) contribute to reaching milestones 1,2 and 3, which track essential progress going from TRL3 (the current proof of concept) to TRL6 (functional product in the customer's environment) and investment readiness.

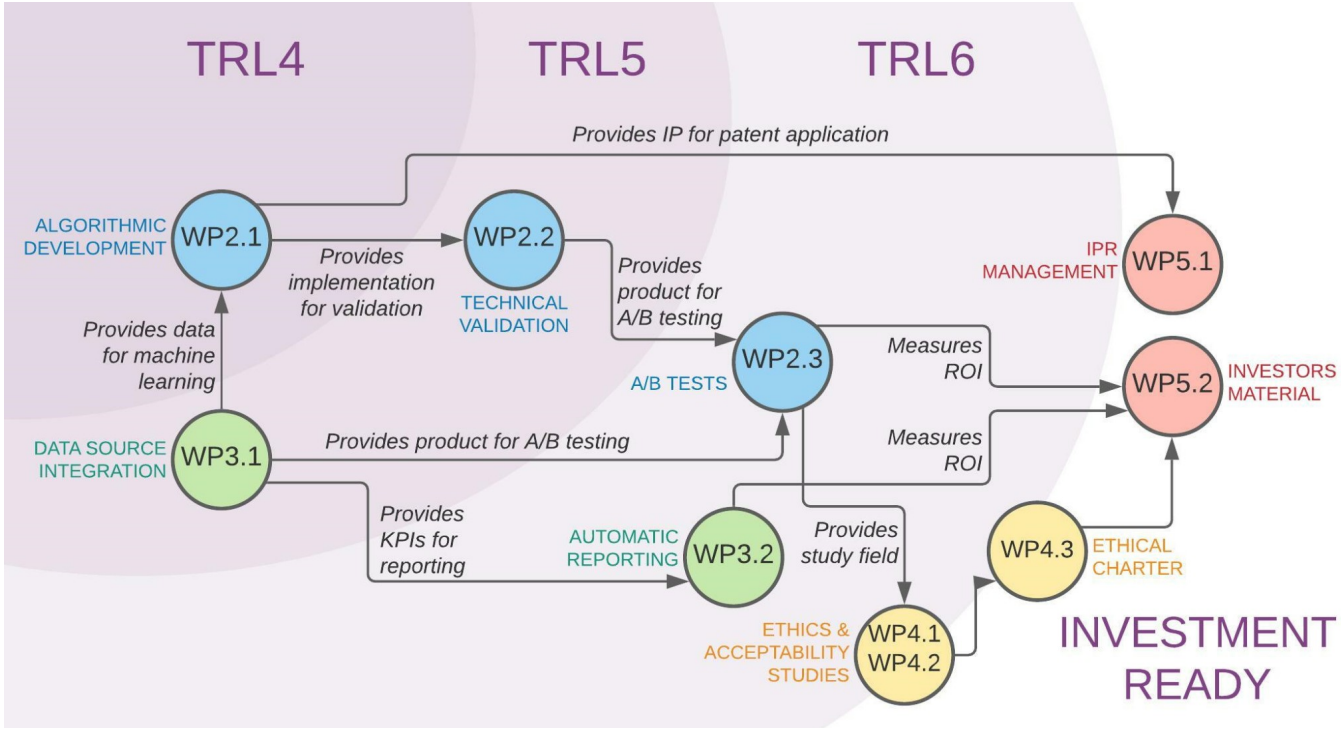


Figure 4. Relations between WPs.

WP1. Project Management

Work package number	1	**Lead beneficiary**				AltaVoce	
Work package title	Project management						
Participant number	**AltaVoce**					CNRS (*)	UTBM (*)
Short name of participant	**Liuni**	Degottex	Martinois	Berriot	Trabelsi	Aucouturier	Rey
Person months per participant:	**5**	2	8	2	2	3	2
Start month	M1			**End month**		M24	

(*) CNRS and UTBM are linked third-party co-operating the FEMTO-ST Institute

Objectives

Work in WP1, led by partner 1 (AltaVoce; WP lead: Liuni), is concerned with the project's proactive management, incl. elaborating and implementing the project's Tech-to-market transition plan, Data management plan and Dissemination and exploitation plan.

Description of work

Task 1.1 Tech-to-market Transition management

Objectives: Draft the project's tech-to-market transition plan in coordination with the project's PM; ensure transitions activities start immediately; monitor the project's progress

Work programme: The transition plan will be drafted at M3, and revised at M12. Coordination will be ensured by periodic consortium meetings yearly, open to all project participants, during which activity reports (at M12 and M24) will be presented and adopted.

Partner's contributions: Led by partner 1 (AltaVoce), with a steering committee composed of one nominated person ("technical lead") for each WP (Degottex, Berriot, Aucouturier & Martinois).

Task 1.2. Data management

Objectives: Draft and implement the project's data-management plan; interface with PM

Work programme: Data Management Plan (DMP) will be drafted at M6 and updated at M12.

Partner's contributions: Led by partner 1 (AltaVoce)

Task 1.3. Financial management

Objectives: Monitor project budget in consortium meetings; write cost statements and financial reports

Work programme: Cost-statements issued with activity reports at the time of yearly consortium meetings.

Partner's contributions: Led by partner 1 (AltaVoce); reports from each WP's technical lead.

Task 1.4 Dissemination and exploitation

Objectives: Draft a dissemination and exploitation plan; implement actions that allow the project team to share insights created and milestones reached during the project.

Work programme: Dissemination plan will be drafted at M6, and updated a M12.

Partner's contributions: All partners will participate in the promotion of the project and its results. Project's results (demonstration of value, commercial milestones reached, etc.) will be disseminated, when possible, as research results in general-science journals (e.g. PNAS), as market-research pieces in call-center trade magazines (e.g. En-Contact) and trade conventions (e.g. VoiceTech, Call & Contact Center Expo) as well as in the general, technical and economic press via press-releases issued by the partner's press offices. Additionally, the project will also offer briefings to governmental agencies such as ANACT (working conditions) and CNPEN (ethics).

Deliverables (brief description and month of delivery)

D1.1. Tech-to-market transition plan (at M4, revised at M12)

D1.2. Data Management Plan (DMP) (at M6, revised at M12)

D1.3. Dissemination and exploitation plan (at M6, revised at M18)

D1.4. Activity and financial report, at the end of each RP (M12 and M24)

WP2. Automation

Work package number	2	**Lead beneficiary**			AltaVoce
Work package title	Automation				
Participant number	**AltaVoce**				
Short name of participant	Liuni	**Degottex**	Martinois	Trabelsi	Berriot
Person months per participant:	8	**12**	2	18	4
Start month	1			**End month**	18

Objectives

Work in WP2, led by partner 1 (AltaVoce; WP2 lead: Degottex) consists in developing procedures to adapt the parameters of voice transformations to the real-time state of the call-center, incorporating real-time voice automation in the telephony system used in the call-center, and conducting A/B testing to validate the economic value of the innovation in the actual call-center environment.

Description of work

Task 2.1 Algorithm development and in-house validation

Objectives: Develop an AI-driven control algorithm able to automatically adapt the AltaVoce voice transformations to the context of the call; put the prototype through functional and performance tests by simulating its impact retrospectively on the test data we have already collected..

Work programme: Algorithmic development to be conducted from M1-M4, with in-house validation to be conducted at M4 (MILESTONE 1), measured as increase of KPI compared to no automation, and certified in the form of a confidential technical report (DELIVERABLE D2.1). Completion of Task 2.1. triggers the IPR activities (e.g. patent application) of WP5.1.

Partner's contributions: Led by partner 1 (AltaVoce). The work will exploit the high capacity of neural networks to handle context-dependent predictions. First developments will use time-independent prediction (simple Feed-Forward layers), and then units and architectures that are aware of time constraints (Convolutional and Recurrent layers).

Task 2.2. Deployment and technical validation

Objectives: Deploy the new algorithm in the call-center environment, by integrating it in the software infrastructure we have already validated for the voice transformations, and obtain technical validation from commercial partner Comdata's IT Services.

Work programme: Technical validation to be obtained by M6 (MILESTONE 2), verified in the form of a technical report (D2.2) and a software release (client software v2.0, D2.3)

Partner's contributions: Led by partner 1 (AltaVoce), with feedback from the Comdata IT services deploying the technology. Work will include speed optimization of the machine learning algorithm, updates of client GUI and data/network access for distant communication between software client and server. Client's behavior analysis and design iterations will be also included to ease and speed up the deployment of the software.

Task 2.3. Economic validation by A/B testing

Objectives: Validate the economic value of the innovation by conducting A/B experiments which test the impact of automated voice technologies on measures of performance in the actual call-center environment.

Work programme: We will conduct at least 3 data collection campaigns, including typically N=50 agents, followed longitudinally for 3 months, resulting in 10,000s hours of real-business operations. Voice transformations will be assigned to call-center agents on a daily basis, with random rotations among possible configurations (e.g. automated vs no-transform). Participating operators and customers will be blind to which transformation is active at any given time. Performance indicators (KPI) will be measured at the level of individual calls (e.g. monetary value, customer satisfaction) or aggregated over a day (e.g. sales per hour). Results from one campaign will be included in the RP1 activity report at M12, and data from all 3 campaigns will be available by M18 (MILESTONE 3), verifiable in the form of a confidential technical report (D2.3). Test results will serve to construct an objective measure of added commercial value (e.g. a +10% increase in sales per hour), which will be the basis for the preparation of investment-readiness materials and commercial discussions from M18 onwards (WP5).

Partner's contributions: Led by partner 1 (AltaVoce) in collaboration with external commercial partner Comdata. Staff from Comdata Sales Management will be involved to select the appropriate B2B client and campaign for each test, presenting the initiative and obtaining client approval. Comdata operators will be compensated for possible loss of performance due to transformation rotations, e.g. in the no-transform conditions. Metrics of added commercial value will be analysed and agreed upon jointly by AltaVoce and Comdata data science teams.

Deliverables (brief description and month of delivery)

D2.1. Technical Report about in-lab validation (at M4, MILESTONE 1)

D2.2. Technical Report about technical validation (at M6, MILESTONE 2)

D2.3. Software release: client software v2.0 including automatic adaptation (at M6, MILESTONE 2), bundled with web services v2.0 (see WP3/D3.2 below)

D2.4. Technical Report about test campaign results (at M18, MILESTONE 3)

WP3. Analytics

Work package number	3	**Lead beneficiary**			AltaVoce	
Work package title	Analytics					
Participant number	**AltaVoce**					CNRS
Short name of participant	Liuni	Degottex	Martinois	Trabelsi	**Berriot**	Aucouturier
Person months per participant:	4	6	2	4	**18**	3
Start month	M1				**End month**	M18

Objectives:

Work in WP3, led by partner 1 (AltaVoce; WP2 lead: Berriot) consists in developing and validating the back-end and front-end monitoring services necessary for collecting data for training and testing the WP2 technology.

Description of work

Task 3.1. Data-source integration

Objectives: Extend the existing web dashboard/services (Figure 2) to add the capacity to link with the client call-center data sources (e.g. Microsoft Power BI) and to collect input and output data for the machine learning technology of WP2. Input data consists of data about call context, such as operator's demographics, experience, type of call, etc. which serves as features for predicting voice transformation performance. Output data consists of KPIs (e.g. sales per hour), which are used to train and evaluate the technology.

Work programme: Service development to be conducted from M1-M6, with technical validation in the call-center environment to be conducted at the same time as Task 2.2 (MILESTONE 2). Completion of Task 3.1. triggers the A/B tests of the complete system (Task 2.3). Verifiable with a report on technical validation by Comdata (D3.1), and with software release as part of deliverable D2.3. (web services v2.0, D3.2)

Partner's contributions: Led by partner 1 (AltaVoce), in collaboration with Comdata IT services. Access to client data sources may use either a push (offering the call-center a data format and API with which they can send their data) or a pull strategy (directly calling from the call-center's data API, if available).

Task 3.2. Automatic reporting of product performance

Objectives: Extend the web dashboard/services to add the capacity to automatically compute/report on value created by the product (e.g. +9.2% sales per hour in call campaign #2).

Work programme: Development to be conducted from M6-M18, in parallel to the A/B tests of Task 2.2. Verifiable with software release of web services v3.0 (deliverable D3.2, at M18)

Partner's contributions: Led by partner 1 (AltaVoce), with data-science expertise from partner 3 (CNRS). Work will include automatizing data-science tasks such as statistical analyses, effect size calculation and visualisation.

Deliverables (brief description and month of delivery)

D3.1. Technical Report about technical validation (at M6, MILESTONE 2)

D3.2. Software release: web services v2.0 including data monitoring (at M6, MILESTONE 2), bundled with client software v2.0 (see WP2/D2.3 above)

D3.3. Software release: web services v2.0 including automatic reporting (at M18, MILESTONE 3).

WP4. Acceptability

Work package number	4		**Lead beneficiary**			CNRS		
Work package title	Acceptability							
Participant number	AltaVoce		**CNRS (*)**			UTBM (*)		
Short name of participant	Liuni	Degottex	**Aucouturier**	Postdoc 1	Postdoc2	Rey	Simoncini	Triclot
Person months per participant:	2	2	**6**	18	18	6	3	4
Start month	M1				**End month**	M24		

(*) CNRS and UTBM are linked third-party co-operating the FEMTO-ST Institute

Objectives:

Work in WP4, led by partner 3 (CNRS; WP4 lead: Aucouturier) consists in conducting an independent empirical study of product usage, from the point of view of ethics and social science, and establishing in what work situation the system's added value is societal desirable.

Description of work

Task 4.1. Ethical acceptability

Objectives: Collect experimental data on the perceived moral acceptability of text vignettes describing different scenarios of application of voice transformations in the context of the call-center (see Bonnefon et al. 2016 for a similar approach on autonomous vehicles); asking questions such as "*Do operators find acceptable that their own voices are transformed without their knowing*", or "*do customers find acceptable that operators' voices are optimized to sound more agreeable*".

Work programme: Ethical approval for the study sought for M4. The vignettes will be developed from M1-M6; studies will be conducted online and analysed from M6-M18. Completion of Task 4.1. will trigger the drafting of the product's Ethics Charter (Task 4.3). Verifiable with ethical approval letter (D4.1., M4), publicly shared data and analysis code (deliverable D4.2) as well as a report/scientific publication (deliverable D4.3, see e.g. Guerouaou, Vaiva & Aucouturier, 2021).

Partner's contributions: Led by partner 3 (CNRS, one 18-month postdoctoral fellow), in collaboration with external commercial partner Comdata management for vignette development, and participation in experiments from Comdata employees. All participants (agents and customers) will give informed consent for participating. No identifying data (incl. no speech recording) will be collected. All experiments will be subjected to the prior approval of the Internal Review Board (IRB) of the Université de Bourgogne Franche Comté.

Task 4.2. Usage ethnography

Objectives: Conduct a situational observation of the deployment of the voice transformation among call-center employees, using the methodology of ethnography to document the modalities of appropriation of the technology by its users, identify tensions and propose recommendations (Rey, Simoncini & Triclot, 2021); asking questions such as "*If given the choice, when do operators decide to turn on the voice transformation*".

Work programme: Fieldwork (inductive ethnographic surveys and interviews) will accompany the deployment of the technology during the A/B tests of WP2.3, from M6-M18, and data (verbatims) will be analysed from M18-M24. Verifiable with publicly shared data (verbatims; deliverable D4.4) as well as a report/scientific publication (deliverable D4.5, at M24). Completion of Task 4.2. will trigger the drafting of the product's Ethics Charter (Task 4.3).

Partner's contributions: Led by partner 3 (CNRS, one 18-month postdoctoral fellow), with participation of Comdata employees. The project includes a budget to compensate call-center employees for non-productive time while answering surveys.

Task 4.3. Ethics Charters

Objectives: Use the ethical and sociological recommendations of Tasks 4.1 and 4.2 to draft an Ethics Charter, to be appended to the company's *General Terms and Conditions*. The Charter will enforce that the product is deployed in socially acceptable ways, and provide safeguards against usage that increase discrimination (incl. in the sex/gender dimension, cf. gender statement p.6) or diminish the value of human work.

Work programme: Version 1 of the Charter to be drafted at M18 with preliminary results from 4.1 and 4.2, and updated at M24 after the final results of 4.2. Verifiable with publicly shared Ethics Charter (D4.6).

Partner's contributions: Led by partner 3 (CNRS), in consultation with AltaVoce and commercial partner Comdata's legal teams.

Deliverables (brief description and month of delivery)

D4.1. IRB approval for ethics experiment (at M4)

D4.2. Experimental data and analysis code for empirical ethics experiments (at M18, MILESTONE 3)

D4.3. Report/Scientific publication about empirical ethics results (at M18, MILESTONE 3)

D4.4. Data (verbatims of interviews) for ethnographic study (at M24)

D4.5. Report/Scientific publication about usage ethnography (at M24)

D4.6. Ethics Charter, published publicly on AltaVoce's website (at M24)

WP5. Investment-readiness

Work package number	5	**Lead beneficiary**		AltaVoce
Work package title	Investment readiness			
Participant number	**AltaVoce**			
Short name of participant	Liuni	Degottex	**Martinois**	
Person months per participant:	5	2	12	
Start month	M4			M24

Objectives:

Work in WP5, led by partner 1 (AltaVoce; WP2 lead: Martinois) consists in preparing investment-readiness material following all project activities, including IPR, pricing, updated business plan, and the initiation of commercial and investment discussions.

Description of work

Task 5.1. Intellectual Property Rights Management

Objectives: Protect IP from the machine-learning development of WP2.

Work programme: Starting at MILESTONE 1 (M4). Analysis of the innovative characteristics of the invention, and writing the patent application (M4-M6, application due at MILESTONE2). Analysis/response to patent search report (M12-M18). International application (translation, application), M18-M24. Verifiable with publicly available application at M24 (D5.1).

Partner's contributions: Subcontracted to third-party IPR consultancy by partner 1 (AltaVoce), with consultation of AltaVoce's technical and legal staff.

Task 5.2. Pricing, updated business plan and investment-readiness material

Objectives: Use the measure of commercial value established in WP2.3 to propose a pricing model and update the company's business plan, in view to start commercial discussions between AltaVoce and Comdata, and seek

VC/BA investment. Update pitching material/deck for investors.

Work programme: Preliminary version of pricing/BP/investor's deck at M12, based on performance predictions from in-house validation. Updated version at M24, based on metrics from WP2.3. Verifiable with pitch deck (deliverable D5.2, at M24)

Partner's contributions: Led by partner 1 (AltaVoce), in consultations with the Comdata sales team.

Deliverables (brief description and month of delivery)

D5.1. Patent application, upon public disclosure (M24)

D5.2. Pitch deck (MILESTONE 3, M18)

List of Deliverables

Number	Name	WP	Lead participant	Type	Dissemination level	Delivery date
D1.1.	Tech-to-market transition plan	WP1	AltaVoce/ Martinois	R	SEN	M4
D1.2	Data Management Plan (DMP)	WP1	AltaVoce/ Liuni	DMP	SEN	M6
D1.3	Dissemination and exploitation plan	WP1	AltaVoce/ Martinois	R	SEN	M6
D1.4	Activity and financial report	WP1	AltaVoce/ Liuni	R	SEN	M12,M24
D2.1	Technical Report about in-lab validation	WP2	AltaVoce/ Degottex	R	SEN	M4
D2.2	Technical Report about technical validation	WP2	AltaVoce/ Degottex	R	SEN	M6
D2.3	Software release: Client software v2.0 including automatic adaptation	WP2	AltaVoce/ Degottex	OTHER (Software)	PU	M6
D2.4	Technical Report about test campaign results	WP2	AltaVoce/ Liuni	R	SEN	M18
D3.1	Technical Report about technical validation	WP3	AltaVoce/ Berriot	R	SEN	M6
D3.2.	Software release: Web services v2.0 including data monitoring	WP3	AltaVoce/ Berriot	OTHER (Software)	PU	M6
D3.3	Software release: web services v2.0 including automatic reporting	WP3	AltaVoce/ Berriot	OTHER (Software)	PU	M18
D4.1	Ethical approval for ethics experiment	WP4	CNRS/ Aucouturier	ETHICS	PU	M4
D4.2	Experimental data and analysis code for empirical ethics experiments	WP4	CNRS/ Aucouturier	DATA	PU	M18
D4.3	Report/Scientific publication about empirical ethics results	WP4	CNRS/ Aucouturier	R	PU	M18

D4.4	Data (verbatims of interviews) for ethnographic study	WP4	CNRS/Rey	DATA	PU	M24
D4.5	Report/Scientific publication about usage ethnography	WP4	CNRS/Rey	R	PU	M24
D4.6	Ethics Charter, published publicly on AltaVoce's website	WP4	AltaVoce/ Liuni	DEC	PU	M24
D5.1	Patent application, upon public disclosure	WP5	AltaVoce/ Liuni	DEC	PU	M24
D5.2	Updated Pitch deck and business plan	WP5	AltaVoce/ Martinois	DEC	SEN	M24

List of milestones

Milestone number	Milestone name	Related work package(s)	Due date (in month)	Means of verification
1	In-house validation(TRL4)	WP2	M4	In-house validation of the project's new technological development (Deliverable D2.1)
2	Technical validation (TRL5)	WP2-WP3	M6	Software release incl. automation and data monitoring (Deliverables D2.3 and D3.2); Technical validation (Deliverables D2.2, D3.1)
3	Economic validation (TRL6)	WP2-WP3-WP4	M18	Economic validation obtained (D2.3), software with automatic reporting (D3.3), conclusions of ethics study (D4.3)

Critical risks for implementation

Description of risk (indicate level of (i) likelihood, and (ii) severity: Low/Medium/High)	Work package(s) involved	Proposed risk-mitigation measures
Can't deploy technology at Comdata. Low likelihood (existing commercial agreement; a preliminary version of the technology, without automation and analytics, has been validated prior to the project by Comdata and is already deployed for tests). High severity (undermines the measure of market fit in WP2.3)	WP2	If deployment is impossible at Comdata specifically, deployment (WP2.2, 2.3, 3.1) will be sought at other commercial partners (e.g. Webhelp) where the technology is also being tested. Comdata will keep involvement for the co-definition of acceptability (WP4) and commercial elements (WP5)
Can deploy automation but with restricted access to data sources. Low likelihood (a proof of concept of KPI access, requiring manual operation by the customer, is already deployed for tests), medium severity (impacts the quantity of data on which training can be made, thus potentially limits product added value)	WP3	If access to data-sources cannot be automated, manual access will be provided on a slower, possibly weekly basis, still allowing testing in WP2.3. If it is impossible to access KPI data sources altogether, automation will be conducted on the sole basis of the acoustics of the operator's voice (e.g. add smile if voice is too low), for which data is readily available in the transformation client.

Difficulty to include a sufficient number of operators for acceptability studies. Low likelihood (risk alleviated by provisioning budget to compensate call-center employees for lost business hours), low severity (impacts the accurate evaluation of societal acceptability, potentially limiting product adoption in the longer-term)	WP4	Online participation in ethics experiments (WP4.1) can be sought from employees at other commercial partners, and from the general public. In case enough participants cannot be recruited for on-site ethnography, user traces (use duration, hours, etc.) can be collected from the software and data-mined for patterns without requiring participant interviews.
KPI impact measured from A/B tests less than expected/insufficient to seek commercialization. Low likelihood (our proof of concept shows high-to-very-high performance in some contexts, in a way that is unlikely to be due to statistical/sampling noise, so it is unlikely that automation cannot provide value), high severity (would require a shift in the company's business strategy).	WP5	Measuring low KPI impact is not stricto-sensu a risk, but high-value feedback in the process of commercialization. In case results at M18 (Milestone 3) do not warrant prospects of commercialization in the call-center industry, activities in WP5 will be reoriented to define a suitable shift in business strategy, e.g. exploring other markets for which all developments made in the project (automation, analytics, etc.) remain valuable assets.

Table 3.2f: Summary of staff effort

	WP1 - project management	**WP2 -** Automation	**WP3 -** Analytics	**WP4 -** Acceptability	**WP5 -** Investment-readiness	**Total Person-Months per Participant**
AltaVoce / **Liuni**	**5**	8	4	2	5	24
AltaVoce / **Degottex**	2	**12**	6	2	2	24
AltaVoce / **Martinois**	8	2	2	-	**12**	24
AltaVoce / **Berriot**	2	4	**18**	-	-	24
AltaVoce / **Trabelsi**	2	18	4	-	-	24
CNRS / **Aucouturier**	3	-	3	**6**	-	12
CNRS / **Postdoc1 (*)**	-	-	-	18	-	18
CNRS / **Postdoc2 (*)**	-	-	-	18	-	18
UTBM / **Rey**	2	-	-	6	-	8
UTBM / **Simoncini**	-	-	-	3	-	3
UTBM / **Triclot**	-	-	-	4	-	4
Total Person Months	24	44	37	59	19	183

(*) to be recruited for the project

Table 3.2g: 'Subcontracting costs' items

AltaVoce		
	Cost (€)	**Description of tasks and justification**
Press-release writing and distribution (WP1.4)	3,000	Externalized PR agency for the communication activities of WP1.4. Includes service of a PR writer + distribution to agencies in EU and US, for 2 releases at M1 (at project start) and M18 (at end of test campaign)
Patent application	7,000	Externalized patent agency for the IPR activities of WP5.1. Includes writing, application, response to search report
CNRS		
	Cost (€)	**Description of tasks and justification**
Graphic/web design	2,000	External graphic/web design agency for the design of the web experiments of WP4.1 (ethics).

Table 3.2h: 'Purchase costs' items (travel and subsistence, equipment and other goods, works and services)

All purchase costs are < 15% of the personnel costs for that partner (AltaVoce: <181,800€; CNRS: <38,413€).

Travel and subsistence include coordination meetings between Paris and Besançon (WP1), technical visits at Comdata sites (WP2-3), field studies at Comdata sites (WP4) and, for AltaVoce, presentations of the project (incl. Travel, speaker fees, stands) at professional customer-relation conventions in Europe and the US. Total AltaVoce: 30,000€. CNRS: 8,600. **Other goods, works and services** include the rental of GPU and web/data server resources to improve AltaVoce service response time and reliability (WP2-3), compensating for call-center operator's non-productive tome while answering surveys (WP4), open access manuscript charges (WP1) and general-public participant fees for the ethics studies (WP4). Total AltaVoce: 55,000; CNRS:15,000. No major equipment to be purchased in the project.

References

Akrich, M. (2006) La description des objets techniques. Sociologie de la traduction. Paris: Presses des Mines.

Al Masry Z., Bazzaro F., Cabaret K., Dembinski O., Devalland C., Gay C., Rey B., Simoncini N., Triclot M. & Zerhouni N. (2021) Etude de l'impact socio-technique d'un soutien-gorge intelligent pour la détection précoce du cancer du sein, Santé Publique.

AltaVoce (2021). Comdata test campaign analytics, March-May 2021. Confidential Report.

Arias, P., Belin, P. & Aucouturier, JJ. (2018) Auditory smiles trigger unconscious facial imitations. Current Biology. Vol. 28 (4), PR782-R783. https://www.cell.com/current-biology/fulltext/S0960-9822(18)30752-8

Arias, P., Bellmann, C. & Aucouturier, JJ (2021) Facial mimicry in the congenitally blind. Current Biology.

Arias, P., Soladié, C., Bouafif, O., Roebel, A., Séguier, R. & Aucouturier, JJ. (2020) Realistic manipulation of facial and vocal smiles in real-world video streams. IEEE Transactions on Affective Computing, Vol. 11(3)

Belt, V (2002) A female ghetto? Women's careers in call centres. Human Resource Management J. 12(4): 51–66.

Bonnefon, J. F., Shariff, A. & Rahwan, I. (2016). The social dilemma of autonomous vehicles. Science, 352(6293), 1573-1576.

Callon, M., Lascoumes, P. & Barthe, Y (2001) Agir dans un monde incertain. Essai sur la démocratie technique, Paris, Le Seuil (collection "La couleur des idées").

Cooke, M., Mayo, C. and Valentini-Botinhao, C. (2013) Intelligibility-Enhancing Speech Modifications: the Hurricane Challenge, in Proc. Interspeech, pp. 3552–3556.

Doellgast, V., & O'Brady, S. (2020). Making call center jobs better: The relationship between management practices and worker stress.

European Commission, Directorate-General for Research and Innovation (2020) Implementing the pro-active management of the EIC pathfinder for breakthrough technologies & innovations: Lessons from the ARPA model & other international practices. KI-03-20-470-EN-N, doi:10.2777/593467

Guerouaou, N., Vaiva, G. & Aucouturier, JJ (2021) The shallow of your smile: the ethics of expressive vocal deep-fakes. Philosophical Transactions of the Royal Society B, in press.

Global Industry Analysts (2021) Global Call Centers Industry report. Retrieved 21/09/2021 from https://www.reportlinker.com/p087286/World-Call-Centers-Markets.html

Invoca (2021) Buyer Experience Benchmark Report, retrieved 21/09/2021 from https://explore.invoca.com/buyer-experience-benchmark-report

Magellan Solutions (2020) Call center benchmarking report. Retrieved 21/09/2021 from https://www.magellan-solutions.com/studies/call-center-benchmarking-report

Ponsot, E., Burred, JJ., Belin, P. & Aucouturier, JJ. (2018) Cracking the social code of speech prosody using reverse correlation, Proceedings of the National Academy of Sciences 115 (15) 3972-3977. https://www.pnas.org/content/115/15/3972

Rennies, J., Schepker, H., Valentini-Botinhao, C. and Cooke, M. (2020) Intelligibility-Enhancing Speech Modifications — The Hurricane Challenge 2.0, in Proc. Interspeech, pp. 1341–1345.

Rey B., Simoncini N., Triclot M., « « Les sciences humaines et sociales en recherche technologique : vers une démarche de conception fondée sur l'ethnographie », in Lamard P., Lequin Y. (éd.), La démocratie technique en travail, Presses de l'UTBM, 2021.

Scholarios, D., & Taylor, P. (2011). Beneath the glass ceiling: Explaining gendered role segmentation in call centres. Human Relations, 64(10), 1291-1319.

Zorila, T.-C. , Kandia V. and Stylianou, Y. (2012) Speech-in-noise intelligibility improvement based on spectral shaping and dynamic range compression, in Proc. Interspeech, 2012, pp. 635–638