SUMATOSOFT

PROJECT TYPE
Web application

TECHNOLOGIES
Ruby 2.7.2, Rails 6.0.3.4, PostgreSQL 14.2, Kurento media server, rabbitmq, React 17.0.1, react-scripts, eslint, stylelint, prettier
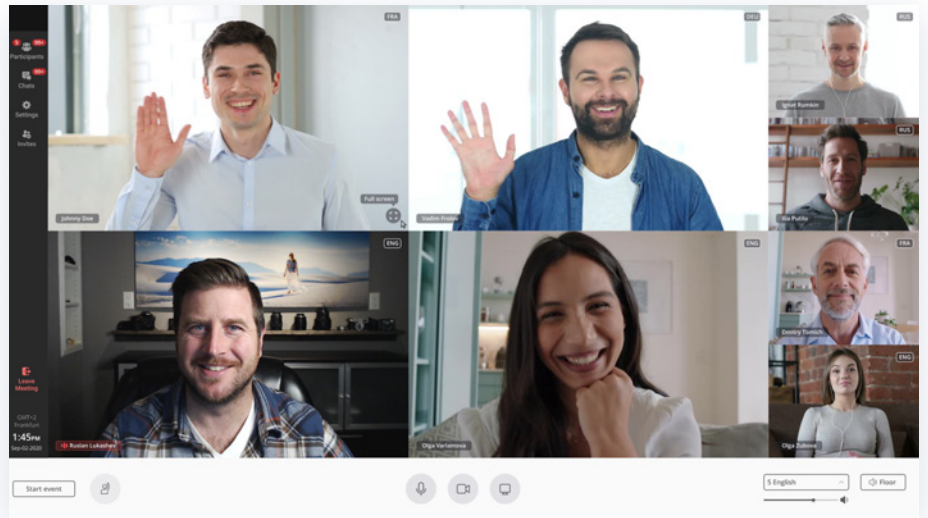
DURATION
1 year

METHODOLOGY
Scrum

TEAM
1 PM/BA
1 Tech Lead/Back-end Developer
2 Back-end Developers
1 Front-end Developer
1-2 QA
1 Designer

Remote simultaneous interpretation web platform for conferences

Platform designed and developed as a tool by interpreters, for interpreters, promising a unique and tailored experience to its users.

Project Distinctive Features

1. **Event creation and management**

2. **Multichannel system** – the ability to switch between multiple audio channels which correspond to languages.

3. **Unlimited participant volume**

4. **Role-based functionality** – unique features for each user role – Moderator, Speaker, Listener, and Interpreter. These features range from microphone/video control, screen sharing, and waiting room management for Moderators, to audio channel switching and handover functions for Interpreters.

5. **Video streaming different modes** – grid, speaker mode, screen sharing mode, multiple screen sharing mode.

6. **Integrated chat system** – an integrated chat feature with file sharing, group and private messaging capabilities.

7. **Participants search**

Business challenge

Our Clients, a group of professional interpreters operating in Belgium were forced to move online during the COVID pandemic. They reviewed all major online tools on the market that were suitable for interpreting at big conferences and were left unsatisfied by their features and overall user experience. Recognizing the limitations of these existing solutions, they decided to pioneer their own. This led to the inception of a project aimed to build a platform tailored for interpreters, by interpreters.

Main Challenge

Our main business challenge was to create an online system that would feel familiar to interpreters accustomed to offline work. Offline, interpreters sit in pairs in a soundproof booth, listen to the speaker through headphones, and speak into a microphone, all while managing a system interface to adjust settings, select channels, or control other features. Our task was to translate this experience to the digital world, designing interfaces and channels that resemble those of physical devices used by interpreters.

Our main technical challenge was to build a platform architecture so it could handle large-scale online conferences. This meant managing a high volume of participants and handling multiple audio channels for different languages concurrently, all without any sound overlapping or form of interference.

Our solution

During the Discovery phase, **we undertook a detailed study** of offline interpreters' equipment and existing online tools, leading to the identification of four user roles:

- ✓ Moderator
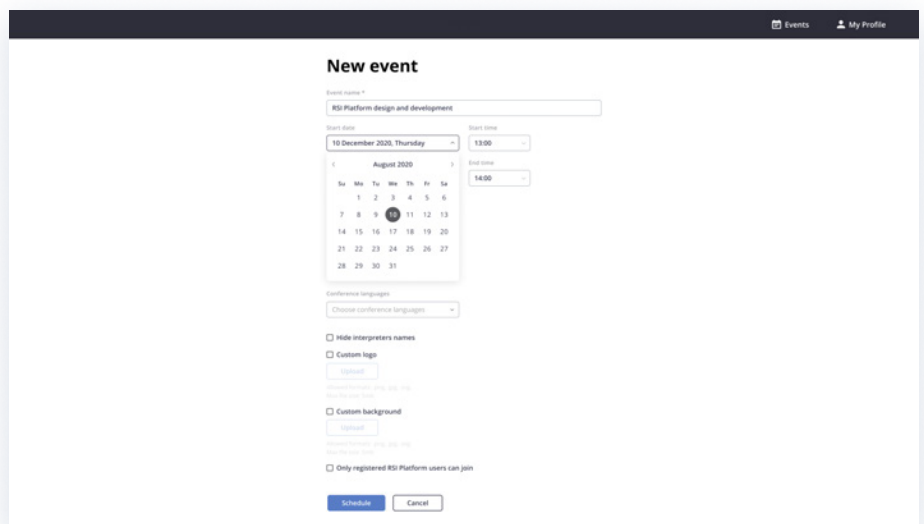- ✓ Speaker
- ✓ Listener
- ✓ Interpreter
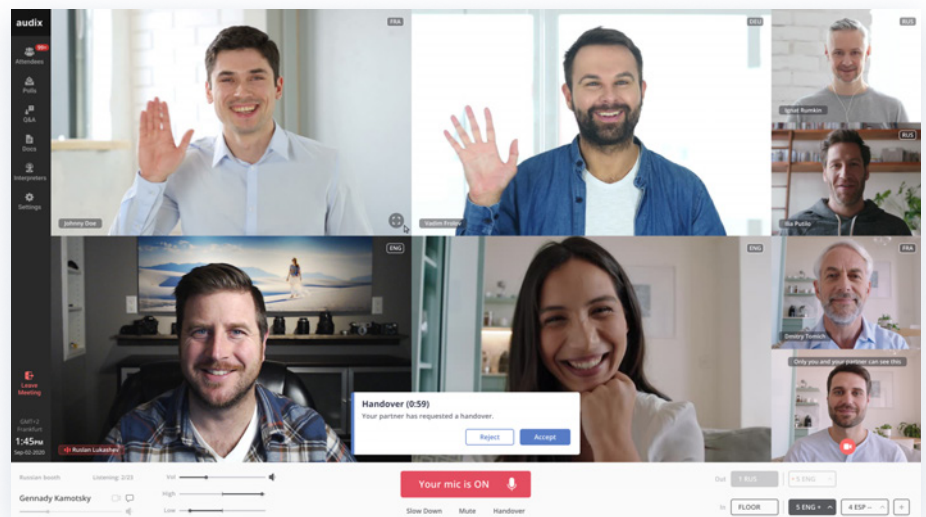
Understanding these roles allowed us to map out the necessary features for each user type, and lay the groundwork for the minimum viable product (MVP) design.

After defining MVP, we **conducted a proof of concept** to ensure that it was possible to construct a multichannel system for interpreting considering the technical requirements of the Client. This phase helped us validate the idea and establish the feasibility of creating such a system online.

Next, we **entered the prototyping stage**. Here, we focused on adapting the offline system familiar to interpreters to the web realm. This process involved not only creating a prototype that combined elements of traditional interpreting systems with web functionality but also educating the Client about online design standards. Regular demo meetings with the Client were conducted throughout the prototyping stage to align their expectations from using the offline interpreting tool with web realms. Finally, we designed a friendly and intuitive user interface for each role:

✓ **For Moderator** – we integrated tooltips for non-obvious functionality, ensuring a smooth user experience even with complex controls.

✓ **For Speakers** – we designed interfaces that closely resemble existing web conference solutions speakers use, reducing the learning curve and enhancing usability.

✓ **For Interpreters** – we built a convenient interface that resembles a familiar offline interpretation tool.

✓ **For Listeners** – we developed an intuitive interface with only the necessary features for listening and communication.

Then, we built a platform that can host online conferences with an unlimited number of listeners from a multinational audience without any language barriers.

Here's how it works

The Moderator first creates the event (which is the conference) and sends distinct invitation links to the Speaker, Interpreters, and Listeners. When entering the conference, all users input their names, and interpreters select a language channel. Listeners can switch between channels after entering the conference. The system identifies the user's role based on the link used to join the conference and either redirects the user to the Speaker channel, places interpreters into a booth for Interpreters, or situates them in the virtual space for Listeners. The conference can include an unlimited number of participants; however, the price for use increases depending on the number of participants since more resources are required, providing a **way for the project to monetize**.
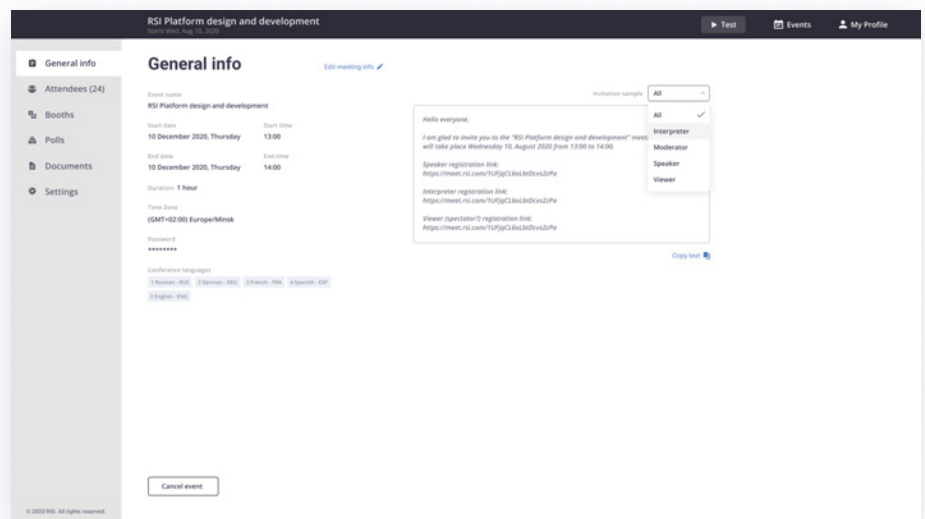
Let's dive deeper into the system functionality per user role:
 1. **Moderator** – responsible for event coordination.

 ✓ There might be several moderators.

 ✓ Create events, look through previous events, and send invitation links.

 ✓ Manage the waiting room.

 ✓ Control the microphones/videos/screen-sharing of participants.

 ✓ Exclude participants.

 ✓ Letting interpreters enter the main speaker's channel to facilitate communication.

 2. **Interpreter** – responsible for real-time translation of the conference.

 ✓ One booth contains two interpreters that can see each other and chat.

 ✓ Working within two groups of channels: 'In' which is the incoming sound channel and 'Out' which reproduces the translated text in one of the declared languages of the conference.

 ✓ The Handover function, designed to warn the second interpreter in the booth that their turn for translation is coming.

 ✓ The Slow down function signals the speaker that they need to slow down, which provides a useful feedback loop.

3. **Listener** – general participant with limited control.

✓ Listen to an audio conference in a language of their choice.

✓ Switch between 'In' channels to select the language of their preference.

✓ Raise a hand if you have anything to say.

4. **Speaker** – holds a conference.

✓ Turn on/off the camera/microphone/screen sharing.

Each participant can change the display layout by switching between the speaker mode, gallery, and screen sharing. Additionally, all participants can access a chat with file sharing, group and private messaging capabilities. The option to search a participant by name is also available.
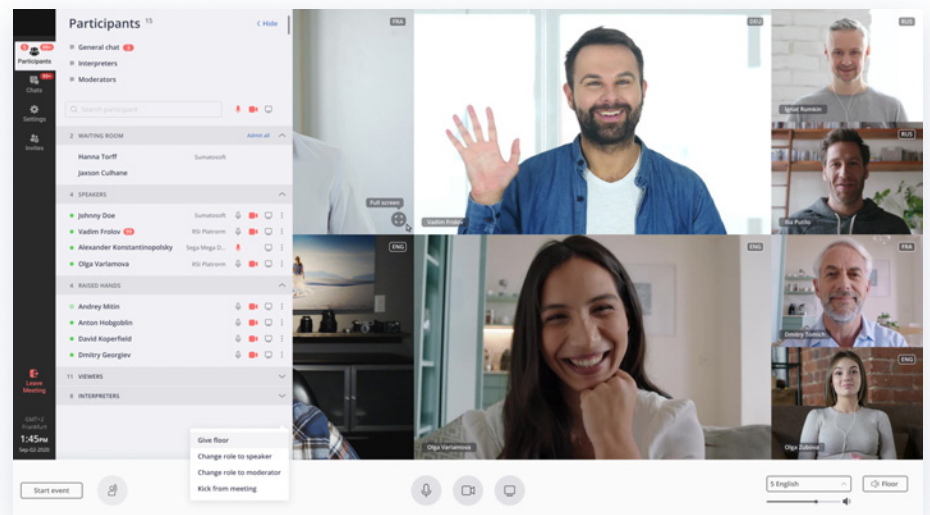
We developed a responsive web version catering specifically to the Listener role. The mobile version offered limited features, focusing mainly on the ability to listen and watch the conference.

Finally, our quality assurance process involved an extensive testing phase. Besides standard functional and non-functional testing, we orchestrated a large-scale test scenario using our own company to ensure the platform could accommodate a high volume of users. Instead of our usual video conferencing tool, we conducted regular corporate meetings with all our employees using the platform. This internal testing process gave us real user feedback and helped us fine-tune the platform before its final release. The Client also tested the app from their side using a simulator that adds thousands of fake users.

Customer's benefits

Our solution addressed the main business challenge of creating an online system reminiscent of offline interpreters' equipment with multiple audio channels for different languages. The platform was successfully launched to the market. Other professional interpreters have used the tool and appreciated its intuitive design and functional effectiveness.

What's happening with the project right now?

The Client is promoting the product on the market.