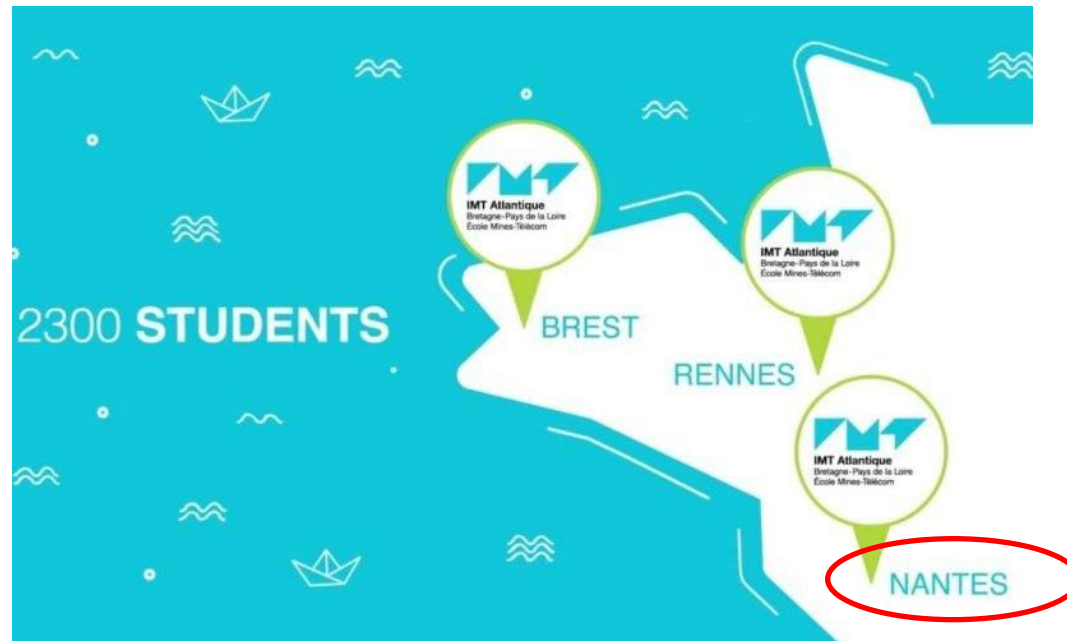




**IMT Atlantique**

Bretagne-Pays de la Loire  
École Mines-Télécom

# *Platform for Situated Intelligence (\psi): Framework pour la HM/RI*



PACCE @ LS2N (Laboratoire des Sciences du Numérique de Nantes)

Perception, Action, Cognition pour la Conception et l'Ergonomie



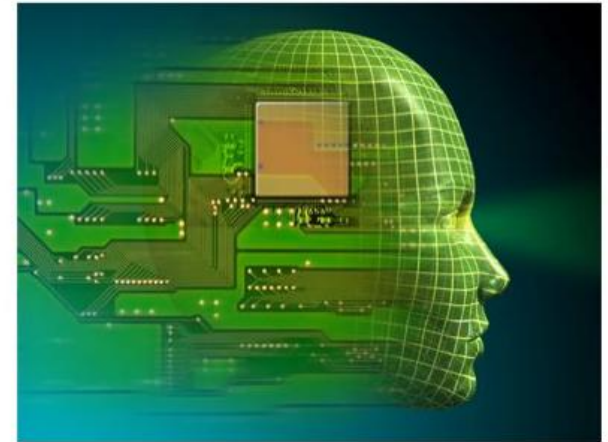
Activity Analysis  
Models

- ▶ Perception
- ▶ Motor skills
- ▶ Cognition
- ▶ Adaptation



... for a better human  
machine cooperation

- ▶ Mediated perception
- ▶ Shared haptic control
- ▶ Decision making
- ▶ Supervision



and smart systems design

- ▶ Collaborative Virtual Environments
- ▶ Medical simulation
- ▶ Co-Presence
- ▶ Human algorithm dialog

**i. INTRODUCTION : LE  
RÉFÉRENTIEL COMMUN**

**ii. PROJETS**

**iii.SAAC**

**iv.\PSI**

**i. WHO IS USING**

**ii. CASE D'USAGE**

**iii. PSISTUDIO**

**iv. CODE**

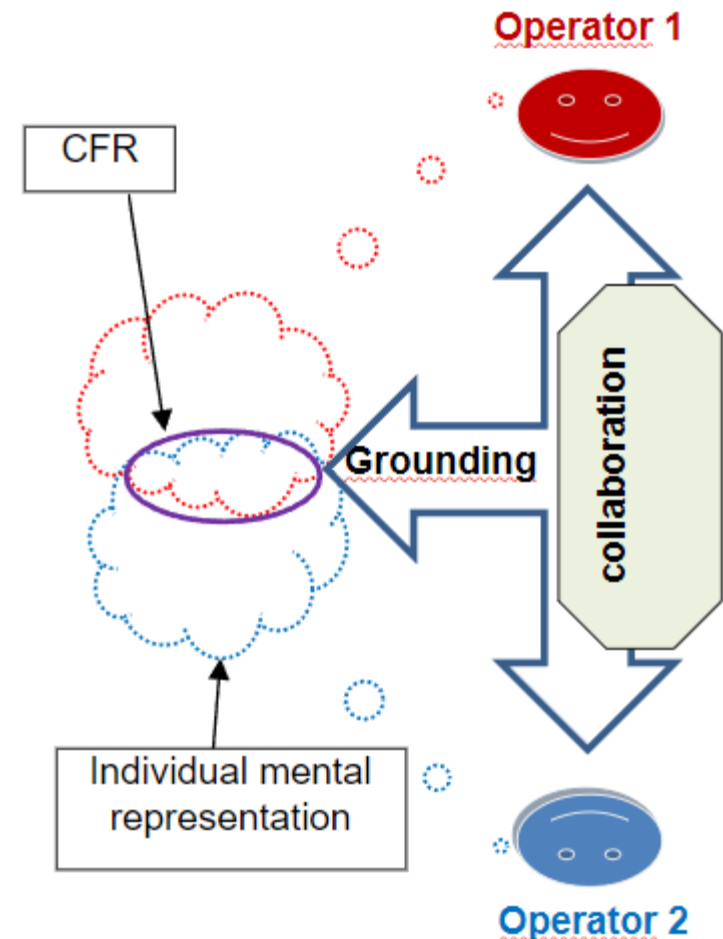
**v. DATASETS**

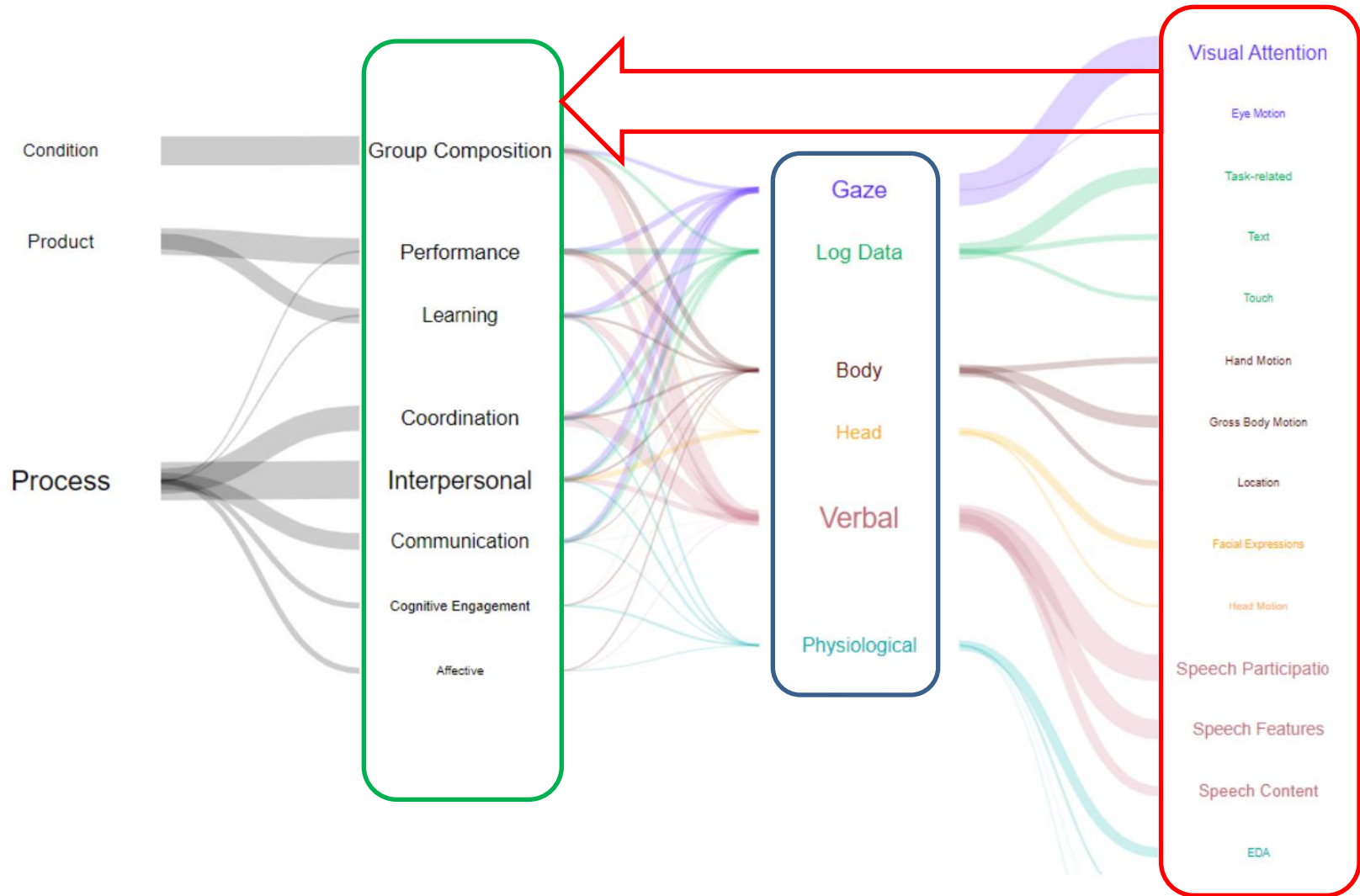
**vi.DISCUSSION**



**IMT Atlantique**  
Bretagne-Pays de la Loire  
École Mines-Télécom

- ▶ Représentations partagées cohérentes (Loiselet & Hoc, 2001)
- ▶ Compréhension, connaissance, croyances...
- ▶ Exemple : chirurgie
  - ▶ Équipe avec un objectif commun
  - ▶ Tâches différentes
  - ▶ Activités en parallèle
  - ▶ Événements indésirables !
- ▶ La communication pour construire et actualiser le référentiel commun (ancrage ; Clark, 1996)
- ▶ Communication verbale et non verbale





## Chirurgie

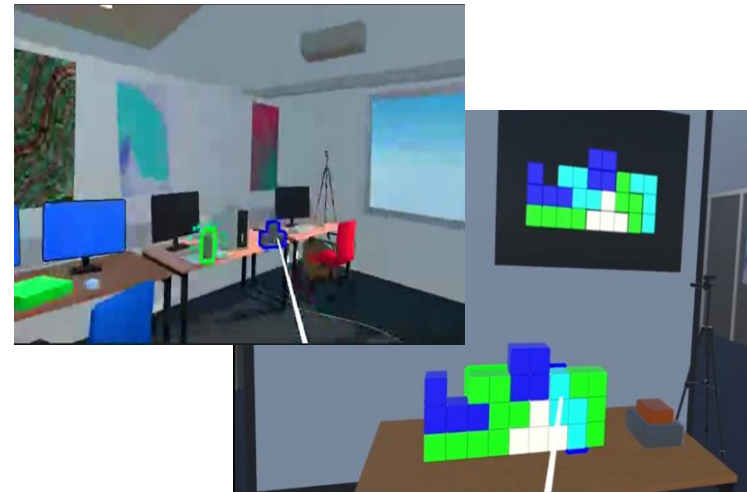


OR n°3 Centre Hépatobiliaire Paul Brousse

A. Allemang--Trivalle  
ICMI 2023



## VR



A. Léchappé  
ICMI 2023



## Machine-readable

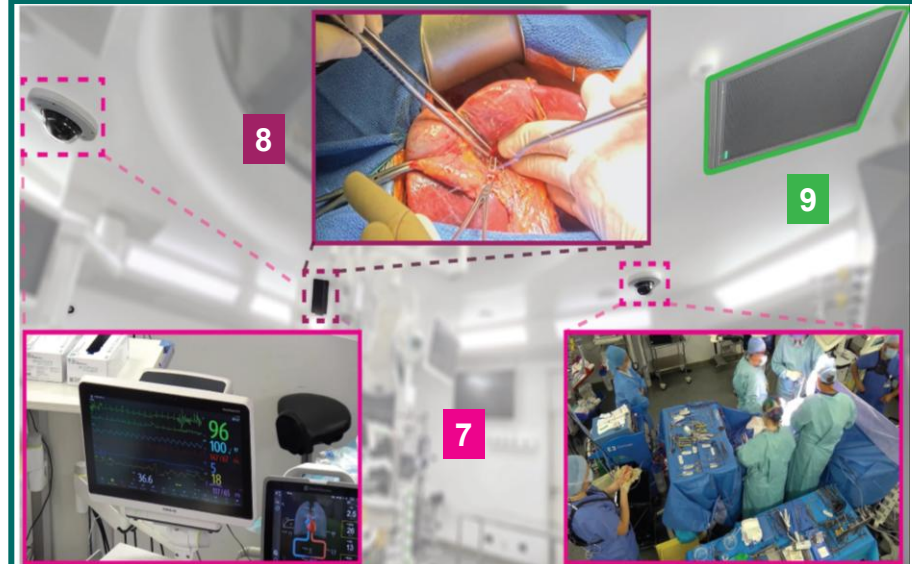
## Human-readable

### Wearable sensors



- 1** Lavalier microphone
- 2** Physiological garment
- 3** Tracking tag
- 4** Audio transmission system
- 5** Wi-Fi router
- 6** Computer (Microsoft PSI)

### Operating room sensors

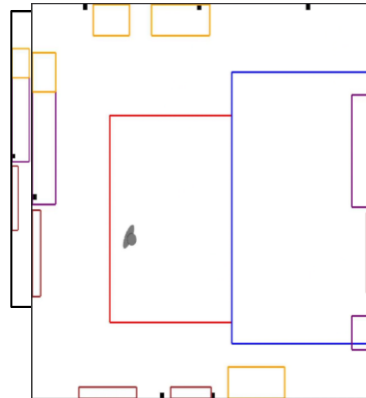


OR n°3 Centre Hépatobiliaire Paul Brousse

- 7** PTZ cameras
- 8** Surgical camera
- 9** Ceiling microphone

## Metrics

- 1 Speech
- 2 Physiological data
- 3 Localization
- 4 Postures
- 5 Gaze



### User

- Speech flow
- Stress and mental workload
- Proximity to areas of interest
- Gaze on peers

### Team

- Overlapping speech
- Turn-taking
- Dominance
- Proximity to team members
- Physiological synchrony
- Joint visual attention
- Mutual gaze

## Questionnaires

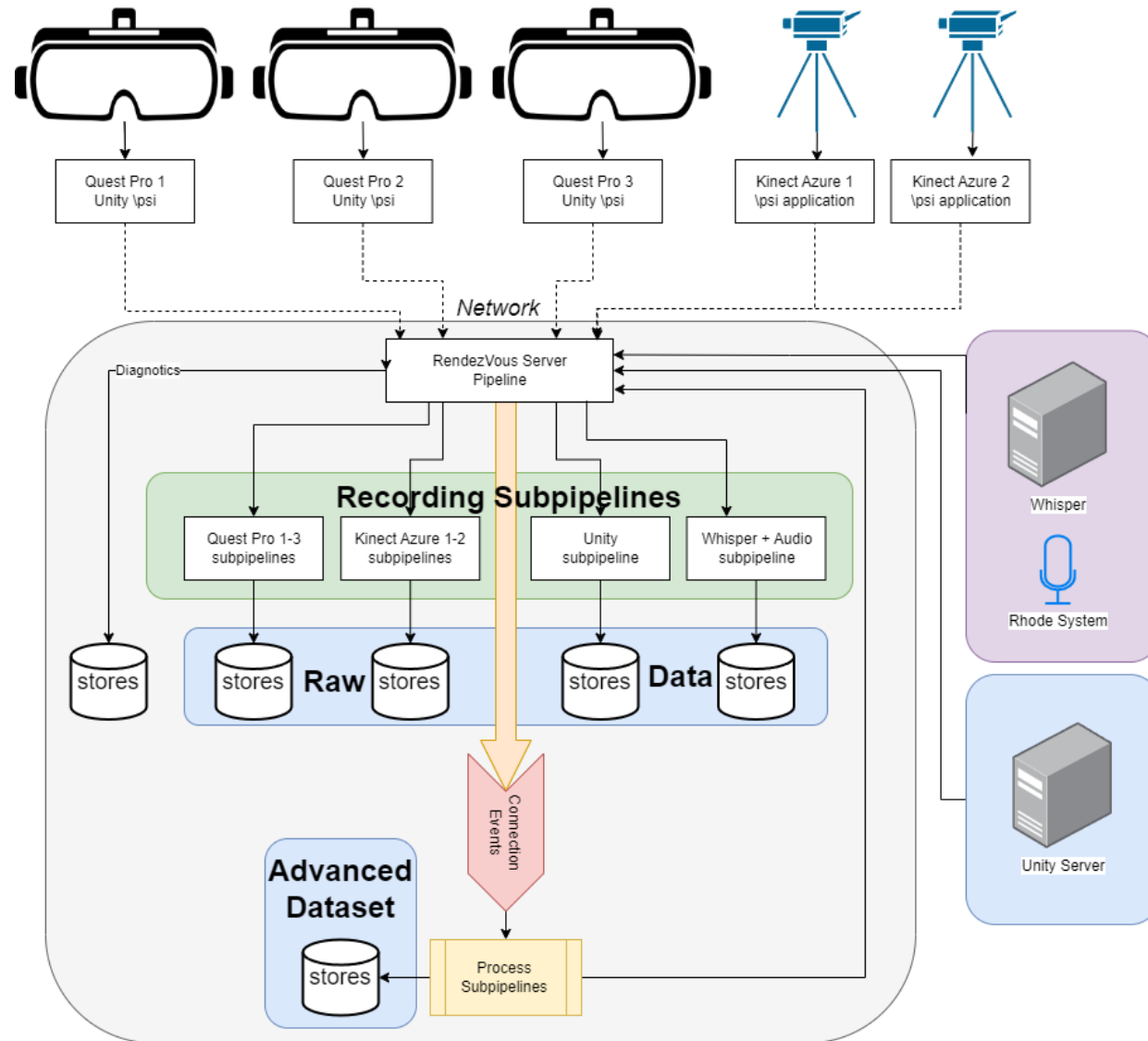
Get to know the people:

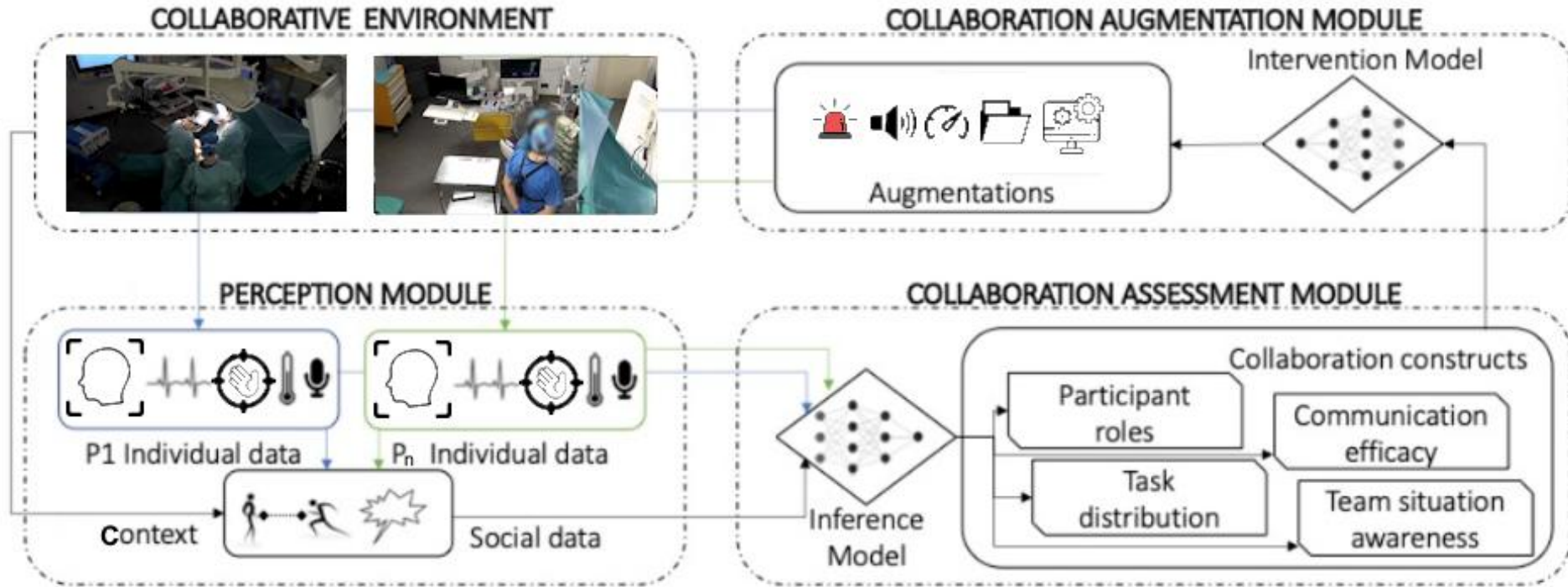
- **Background, Conditions, Relationships**

## Context

Annotations made by expert:

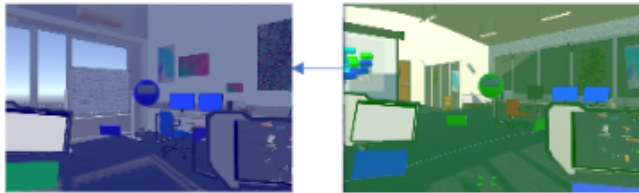
- **Task steps and phases**
- **Errors**
- **Adverse events** (surgery)
- **Collaboration quality**



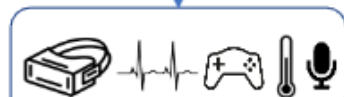
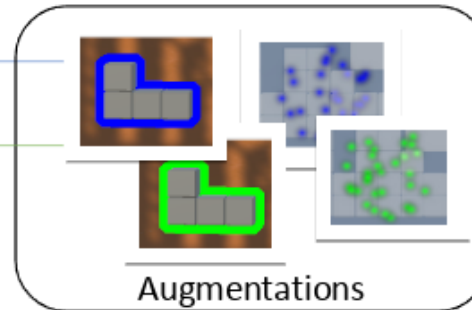


<https://www.imt-atlantique.fr/fr/recherche-innovation/collaborer/projet/saac>

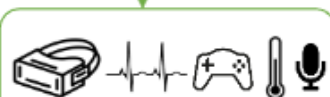
## COLLABORATIVE VIRTUAL ENVIRONMENT



## COLLABORATION AUGMENTATION MODULE



P1 Individual data



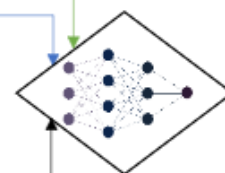
P2 Individual data

VR context



Social data

## PERCEPTION MODULE



Inference Model

Collaboration constructs

Goals progress

Participant roles

Task distribution

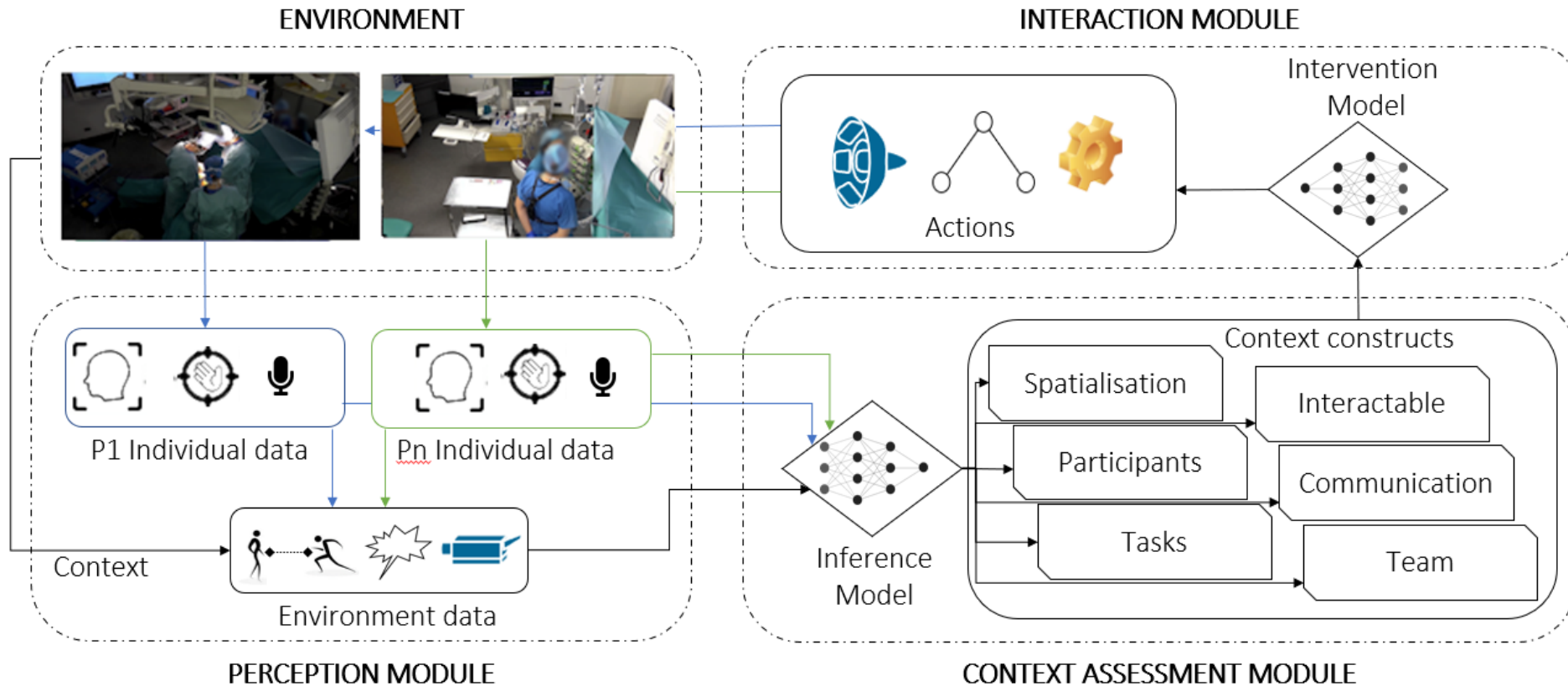
Collaboration archetypes

Communication efficacy

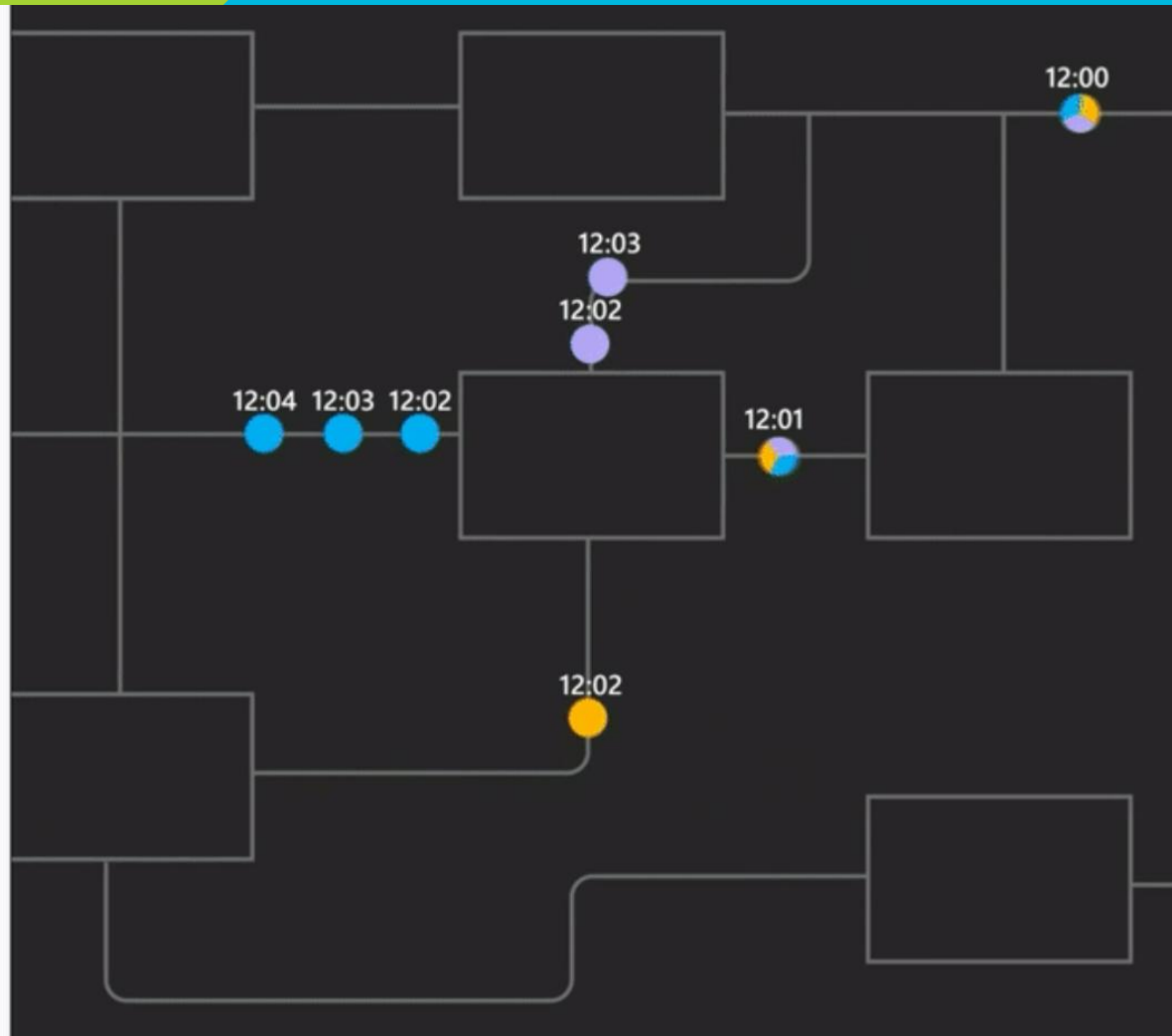
Team situation awareness

## COLLABORATION ASSESSMENT MODULE

## Architecture généralisable en HRI



# INFRASTRUCTURE FOR SYNCHRONIZATION



**IMT Atlantique**  
Bretagne-Pays de la Loire  
École Mines-Télécom

<https://camo.githubusercontent.com/1f2c3afc838e595a3d8738c25877b9cb108fd5576eef23e543e4a2985ab1ae11/68747470733a2f2f777772e6d6963726f736f66742e636f6d2f6556e2d75732f72657365617263682f75706c6f6164732f70726f642f323031382f30312f5073692d476966322d3936302d436f727265637465642e676966>

# Platform for Situating Intelligence Overview

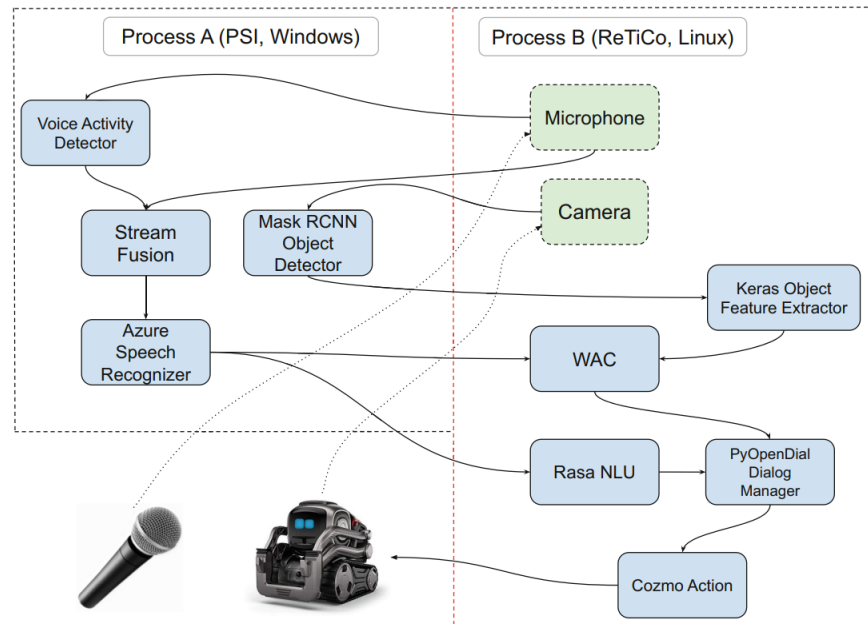
- ▶ Plateforme de synchronisation de flux de données.
- ▶ Framework open source basé sur des composants/pipelines (licence MIT).
- ▶ Possibilité d'utilisation en réseau / à distance (ex. : Hololens, application Sigma).
- ▶ Capacité d'enregistrement des flux de données.
- ▶ Possibilité de relecture en temps réel (ou de simulation à partir des enregistrements).
  - ▶ Utilisation des mêmes composants entre relecture et temps réel.
- ▶ Orienté apprentissage automatique (Machine Learning).
- ▶ Application de visualisation modulaire : PsiStudio.

# Who is Using

Platform for Situated Intelligence has been and is currently used in several industry and academic research labs, including (but not limited to):

- the [Situating Interaction](#) project, as well as other research projects at Microsoft Research.
- the [Interaction Research Lab](#) at Cornell Tech.
- the [PACCE team](#) at IMT Atlantique / L2SN.
- the [Interactive Robotics Group](#) at MIT.
- the [MultiComp Lab](#) at Carnegie Mellon University.
- the [Speech Language and Interactive Machines](#) research group at Boise State University.
- the [Qualitative Reasoning Group](#) at Northwestern University.
- the [Intelligent Human Perception Lab](#) at USC Institute for Creative Technologies.
- the [Tedia research group](#) at Carnegie Mellon University.
- the [F&M Computational, Affective, Robotic, and Ethical Sciences \(F&M CARES\) lab](#) at Franklin and Marshall College.
- the [Transportation, Bots, & Disability Lab](#) at Carnegie Mellon University.

## ► Human Robot Interaction (HRI) :



M. Spitale, C. Birmingham, R. M. Swan and M. J. Matarić, "Composing HARMONI: An Open-source Tool for Human and Robot Modular Open Interaction," *2021 IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China, 2021, pp. 3322-3329, doi: 10.1109/ICRA48506.2021.9560992.

## ► Human Robot Interaction (HRI) :



D. Kontogiorgos and J. Shah, "Questioning the Robot: Using Human Non-verbal Cues to Estimate the Need for Explanations," *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Melbourne, Australia, 2025, pp. 717-728, doi: 10.1109/HRI61500.2025.10974079.

Task Identification

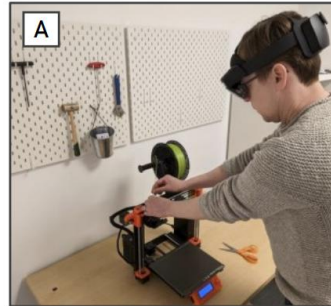
S: Hi, this is Sigma.  
S: What are we doing today?  
U: **Can you help me load the filament in a 3D Prusa printer?**  
S: Sure. I can help you load the filament in a 3D Prusa printer.  
S: [moves to the user's left-hand side]  
S: Before starting, I'll move more to the side here to keep out of the way, but you can always summon me anywhere by holding your palm open and saying 'Come here!'

Gather Step

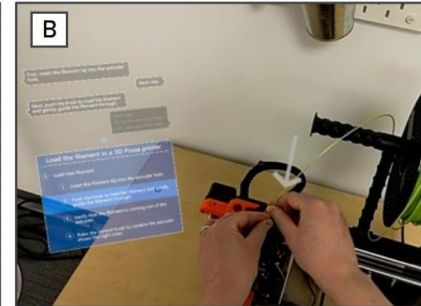
S: Now let's get started.  
S: The first step is to gather the objects listed below.  
S: [gem moves to highlight position of scissors in the world]  
S: I see that you already have the scissors.  
S: [moves to highlight position of filament spool in the world]  
S: I see that you also have the filament spool.  
S: Now that you have everything you need let's move on to the next step.  
...

Complex Step

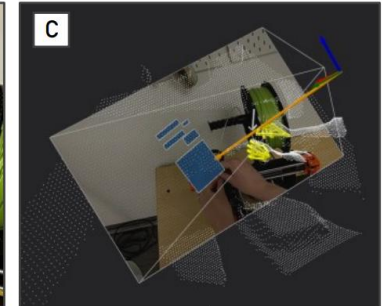
U: **OK I've done that what's next?**  
S: Great. The next step is to prepare the new filament.  
S: [displays hologram indicating placement on spool holder]  
S: First place the filament spool on the spool holder.  
S:  
U: **[places the spool on the holder] Okay next.**  
S: Next unwind a length of filament and straighten as much as possible.  
U: **How much should I unwind?**  
S: You should unwind enough filament to reach from the spool to the extruder, with a little extra to ensure smooth loading, typically around 30-40 centimeters.  
U: **Okay I'm done what's next?**  
S: Next, use the scissors to trim the end at an angle to create a pointed tip.  
U: **Uh how much should I trim?**  
S: You should trim about 1 inch (or 2.5cm) off the end of the filament at a 45-degree angle.



[A] SIGMA is a mixed-reality assistant system that can provide guidance for performing procedural tasks on site and in real time.

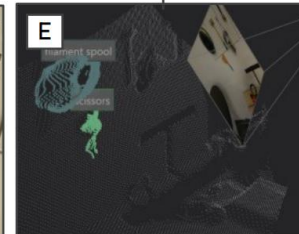


[B] The system displays a virtual panel containing the history of the interaction, and the list of steps to be performed. It can also show predefined holograms relevant for the execution of each step.

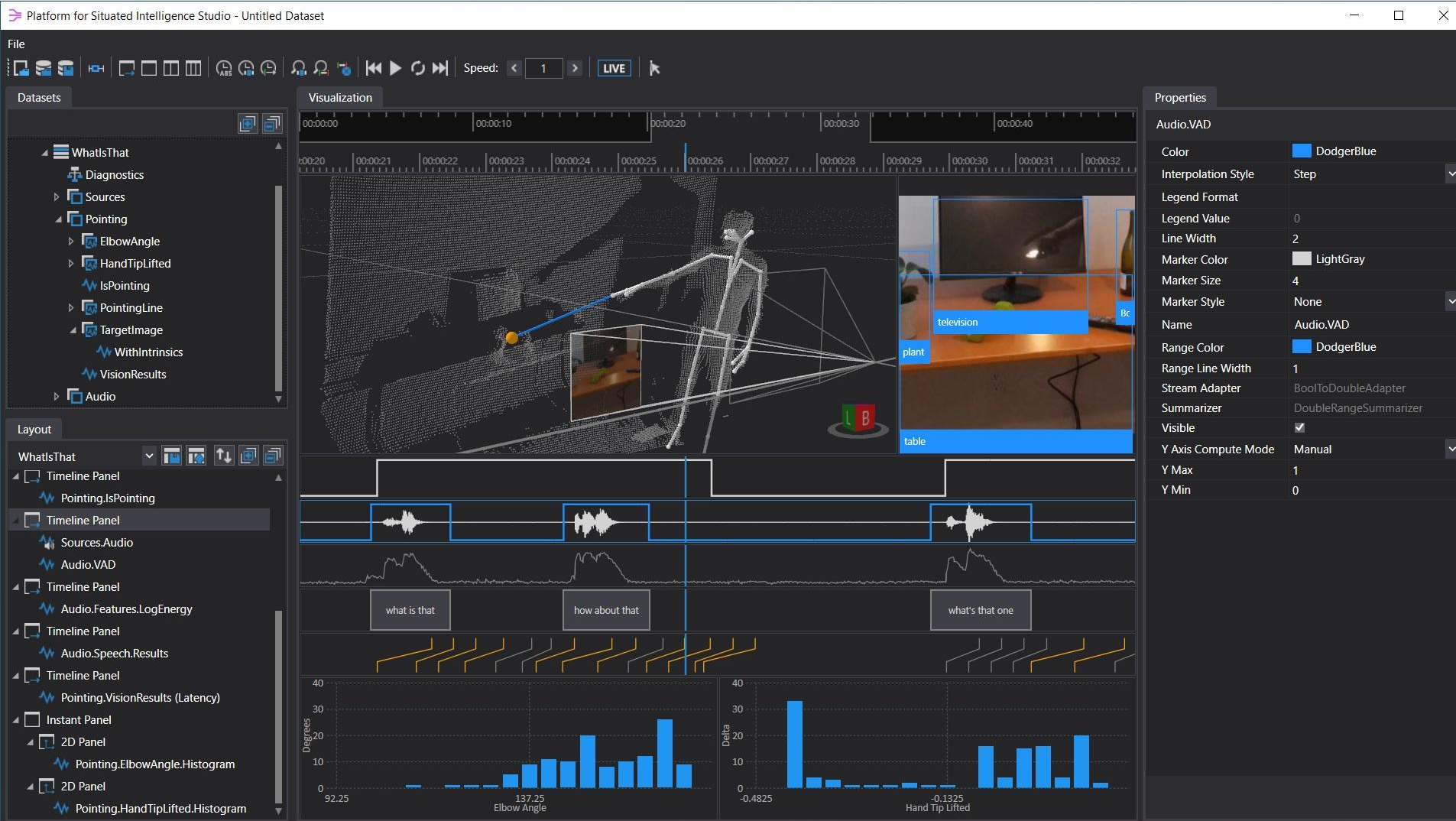


[C] SIGMA leverages and provides access to a variety of multimodal data streams including camera view, depth map, detected objects, gaze, hands, and head position.

During Gather steps, SIGMA uses computer vision models to detect objects in RGB space [D]. The object masks are back projected on the point cloud [E], and SIGMA informs the user about and points to the location of these objects [F].



SIGMA can respond to open-ended user queries during task performance, such as "How much should I unwind?" or "How much should I trim?" by leveraging large language models.



- ▶ Visualisation des données (en temps réel ou à partir d'enregistrements).
- ▶ Création de jeux de données à partir des enregistrements.
- ▶ Annotations !
- ▶ Analyse des données via API :
  - ▶ Exécution d'algorithmes spécifiques
  - ▶ Processus d'extraction pour l'apprentissage automatique
  - ▶ Statistiques sur les données
- ▶ Possibilité de développer des modules de visualisation pour l'application.

- Exemple d'un chatbot avec \psi:

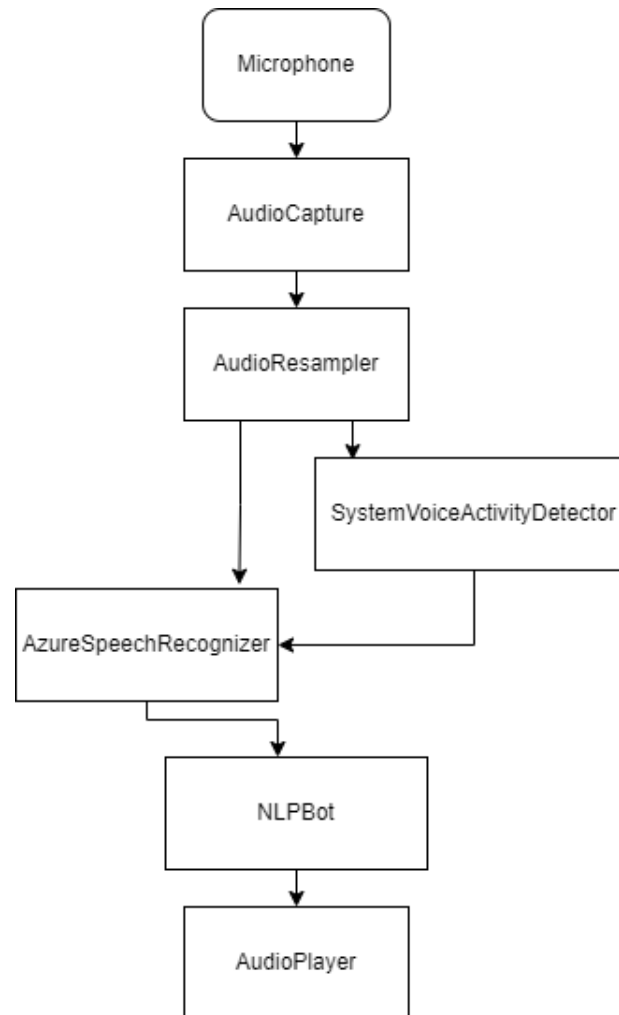
```
var mic = new AudioCapture(p);
var resampler = new AudioResampler(p);
// link : micro -> sampler
mic.PipeTo(resampler);

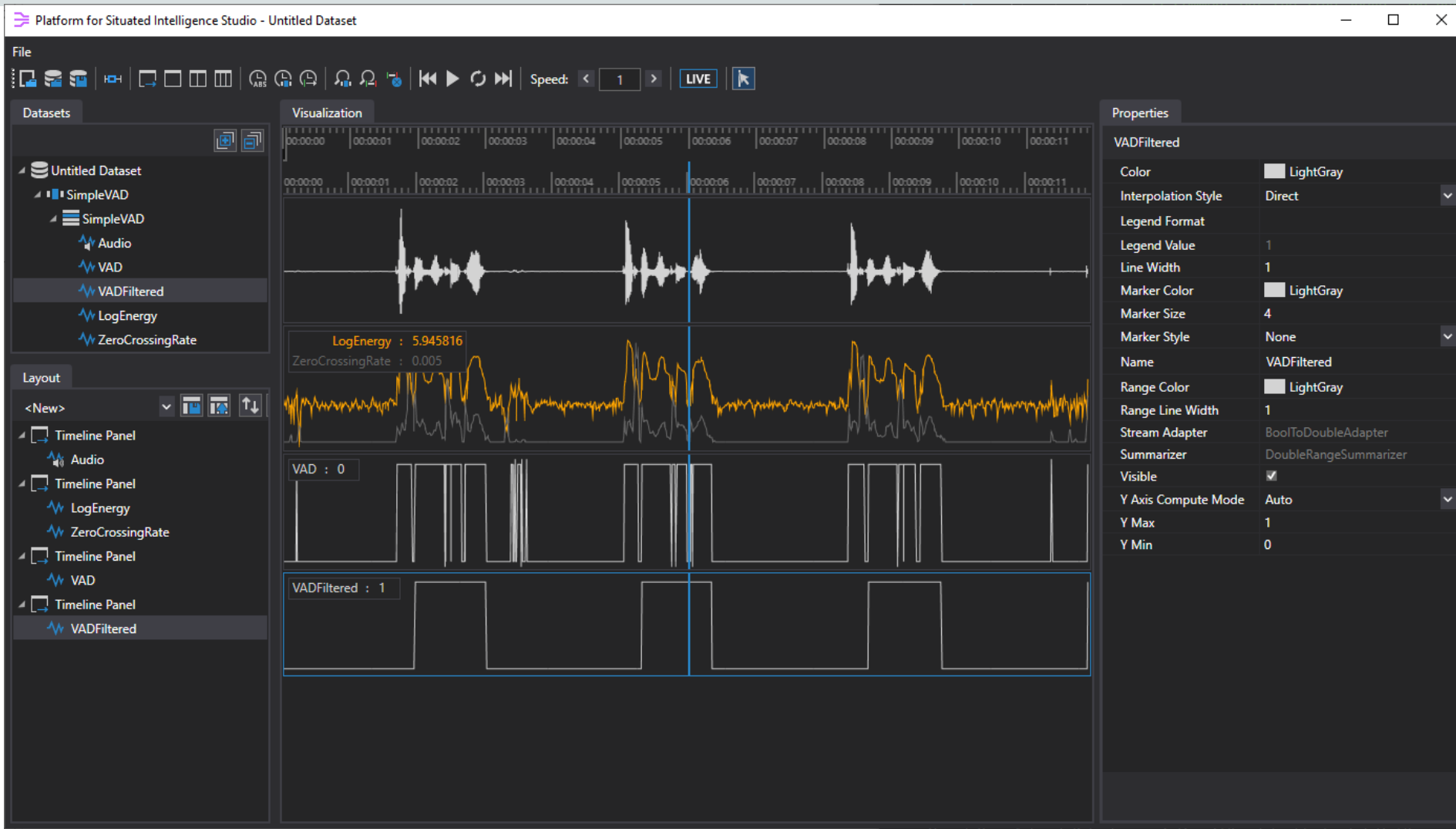
var voiceActivityDetector = new SystemVoiceActivityDetector(p);
// link : sampler -> voiceDetector
resampler.PipeTo(voiceActivityDetector);

var azureSpeechReco = new AzureSpeechRecognizer(p);
// link : sampler + voiceDetector -> SpeechToText Azure
resampler.Join(voiceActivityDetector).PipeTo(azureSpeechReco);

// Process : complete results only
var finalResults = azureSpeechReco.Out.Where(result => result.IsFinal);
NLPBot nlp = new NLPBot(p, botDescription);
// link : SpeechToText Azure -> NLP
finalResults.PipeTo(nlp);

var audioplayer = new AudioPlayer(p);
// link : NLP -> TextToSpeech
nlp.Voice.PipeTo(audioplayer);
```



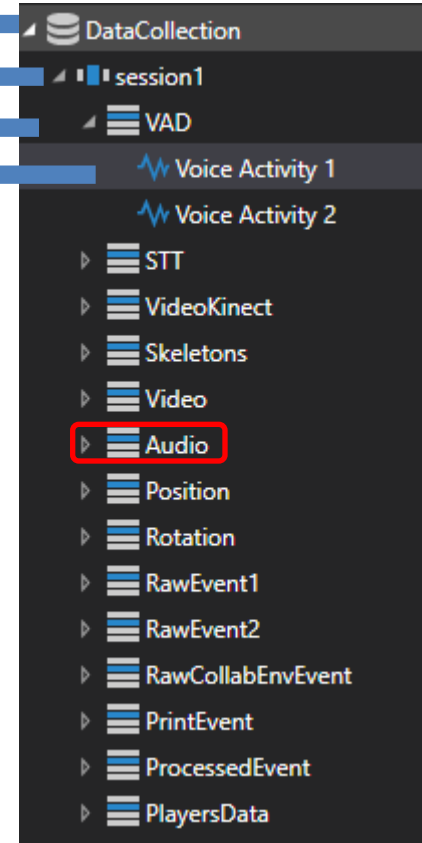


- ▶ Certains composants ont déjà été développés:
  - ▶ Azure services
  - ▶ Whipser
  - ▶ Ollama
  - ▶ OpenFace
  - ▶ OpenSmile
  - ▶ ...
- ▶ Dans le cas d'environnement spécifique :
  - ▶ <https://github.com/microsoft/psi/wiki/Bridging-to-Other-Ecosystems>

- ▶ Compatibilité avec Hololens 2.
- ▶ Ne fonctionne pas dans Unreal Engine (sauf via des requêtes HTTP).
- ▶ Fonctionne dans Unity :
  - ▶ Application de bureau
  - ▶ Quest (2, 3 & Pro)
  - ▶ Mais nécessite des modifications du code (fork du repo officiel)
- ▶ Ne permet pas de traitement d'image.
- ▶ Si vous voulez utiliser \psi dans Unity:
  - ▶ <https://github.com/SaacPSI/psi>
  - ▶ <https://github.com/SaacPSI/saac/tree/master/Components/Unity>

# DATASETS

- ▶ \psi utilise son propre format de donnée
  - Dataset
  - Session
  - Partition
  - Donnée
- ▶ Mais fournit les outils pour les convertir en CSV, JSON ou MessagePack:
  - ▶ <https://github.com/microsoft/psi/blob/master/Sources/Runtime/Microsoft.Psi.Interop.Format/Readme.md>
- ▶ Deux datasets disponibles:
  - ▶ <https://doi.org/10.1145/3710996>
  - ▶ <https://doi.org/10.1145/3706598.3713628>



# DISCUSSION

- ▶ Il existe d'autres frameworks de synchronisation de données :
  - ▶ Quels sont les outils de traitements et de visualisation utilisés ?
  - ▶ Est-il possible de converger vers un ensemble interopérable ?
- ▶ Il y a un intérêt de plus en plus grand à partager des datasets :
  - ▶ Sous quel(s) format(s) ?
  - ▶ Avoir un inventaire des datasets partagés avec leurs spécificités ?
- ▶ Développer des liens avec d'autre(s) PEPR ?

# Merci de votre attention !



**IMT Atlantique**  
Bretagne-Pays de la Loire  
École Mines-Télécom