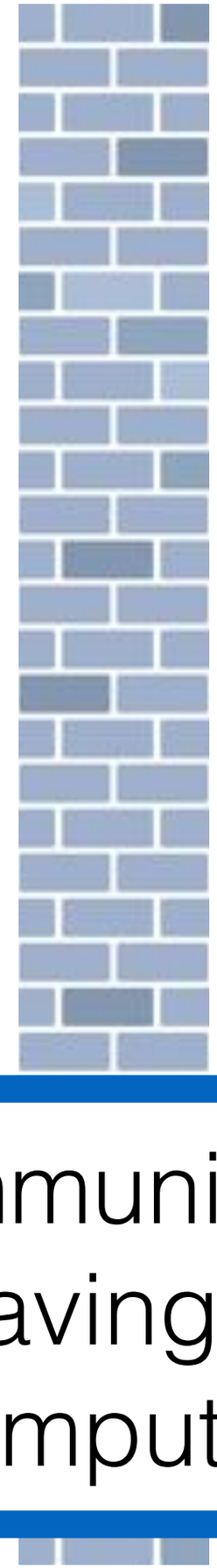
Synergie entre IHM & IA

Gilles Bailly

IHM

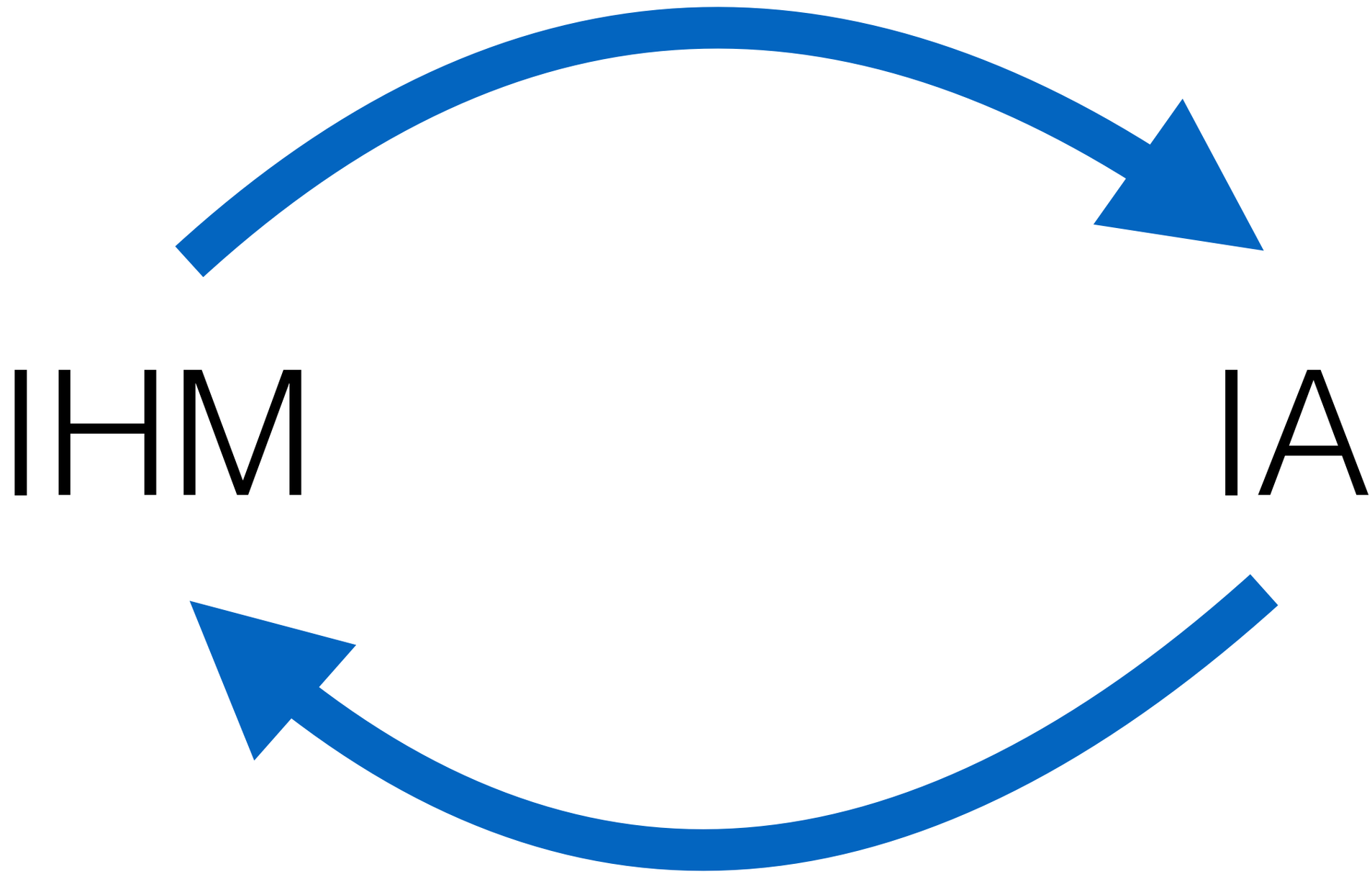


IA

The AI and HCI communities have often been characterised as having opposing views of how humans and computers should interact

IHM

IA


IHM
IA

Définitions

Intelligence Artificielle

Définition?

- Théories et outils pour réaliser des machines capable de simuler l’intelligence
- Traitement automatique de l'information

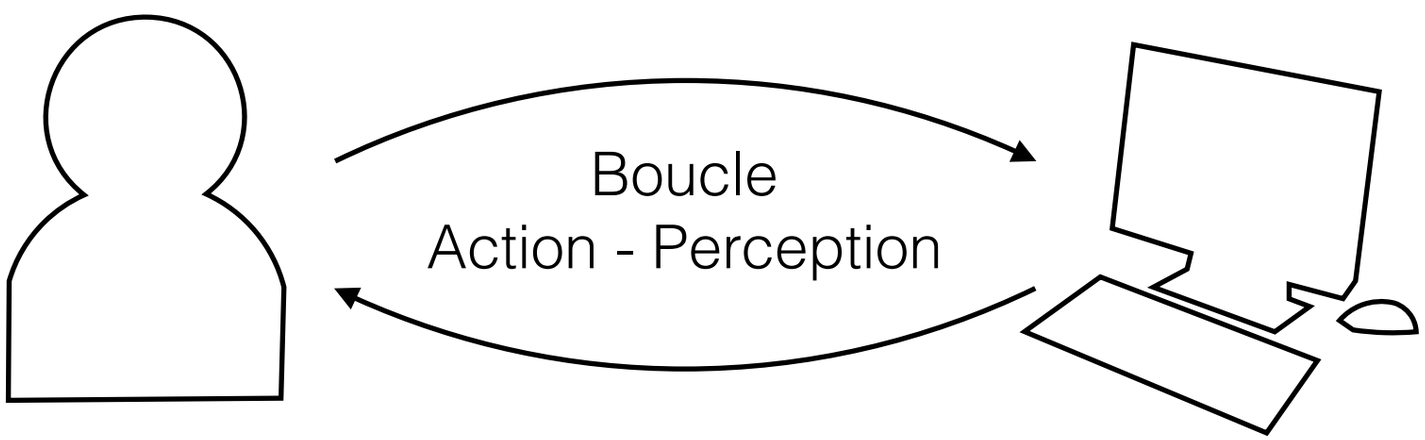
Au sens large

- Logiciel, e.g. système de recommendations
- Matériel (robotique),e.g. robot chirurgicaux

IHM

Interaction Homme-Machine

- La science qui étudie les phénomènes liés à la boucle action/perception entre un ou plusieurs utilisateurs et un ou plusieurs systèmes interactifs
- Domaine qui vise à concevoir, réaliser et évaluer les systèmes interactifs



www.AFIHM.org

IHM

IA

Systèmes intelligents

Problématique

- Divergence réelle ou perçues entre les fonctionnalités du systèmes et son interprétation sociale
- Confusion, absence de confiance, abandon de la technologie

Le cas d'IBM Watson

- Système de recommendation
  pour trouver un meilleur traitement pour le cancer
- Beaucoup d'efforts … faible adoption

IBM Watson

Baumer E. 2017 **Towards human-centred algorithm design.** Big Data & Society.

Harper R. 2019 **The role of HCI in the Age of AI.** IJHCS

Systèmes intelligents

p2. Impact sur les emploi
en 2025 - 2030

p3. Ca fait un peu peur,
Il y a des fantasmes

Problématique

- Divergence réelle ou perçues entre les fonctionnalités du systèmes et son interprétation sociale
- Confusion, absence de confiance, aban
technologie

Q1. Experts réticents
à utiliser l'IA

Le cas d'IBM Watson

- Système de recommendation
pour trouver un meilleur traitement pour le cancer
- Beaucoup d'efforts … faible adoption

IBM Watson

Baumer E. 2017 **Towards human-centred algorithm design.** Big Data & Society.

Harper R. 2019 **The role of HCI in the Age of AI.** IJHCS

Systèmes intelligents

Possible explication?

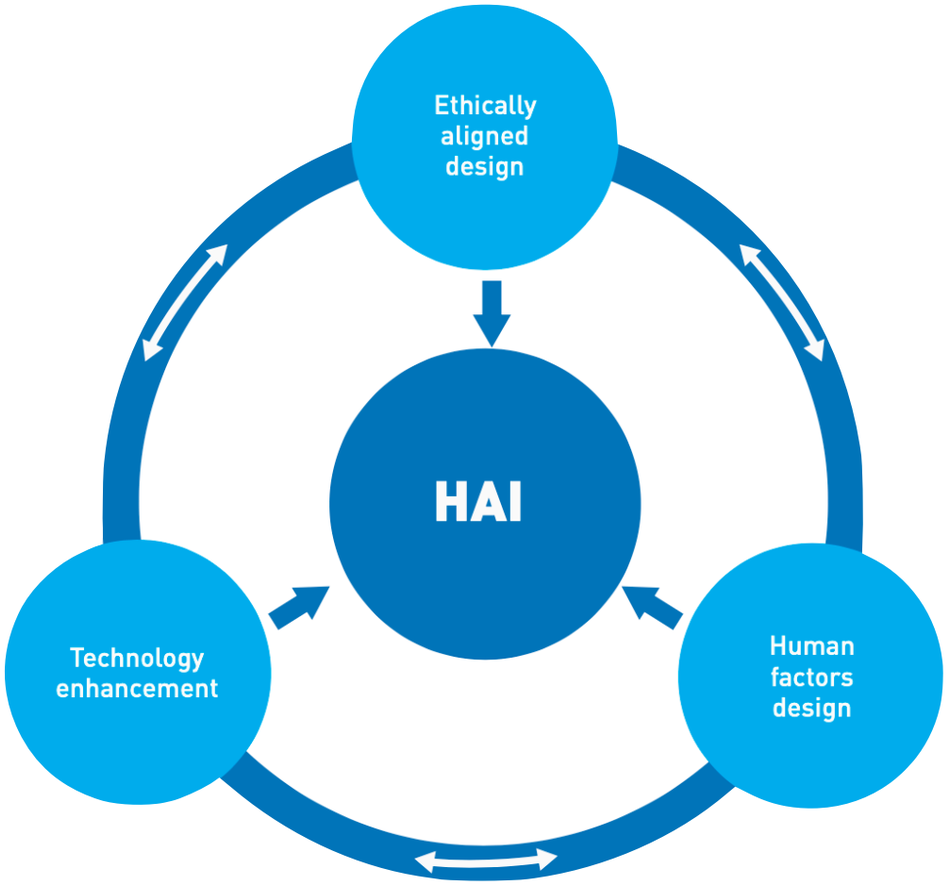
- L'histoire se répète…
- Équipes (d'experts) en IA
  conçoivent l'interaction (pour d'autres experts)

Solution?

- IA centrée sur l'humain
  (Human - Centred AI)
- Conception Centrée Utilisateur

Systèmes

- Éthiques, compréhensibles,
  utiles et utilisables



Xu W. 2019
Toward Human-Centered AI: A perspective from Human-Computer Interaction. Interactions.

Rôle de l'IHM

Principes

Guidelines

Études

etc.

Table 1. Heuristic biases that lead to decision (diagnostic) errors and strategies for mitigating them as reported in [64]. We consider how XAI can provide facilities to aid in these strategies.

Heuristic Bias	Description	Strategies to overcome systematic errors [64]	XAI Strategies for Medical Decisions
Represent-ativeness	Judging likelihood of an event 'A' belonging to a condition due to similarities between the two, but not judging whether A belongs to some other process that could be more similar.	**Compare** disease with **prototypes** of the condition; be suspicious when there is **no good match.**	- Identify **prototypes** of patient instances for each diagnosis - Show similarity between current patient and prototype(s) via **similarity distance**. - Highlight similarity and **contrast** differences in terms of data feature value or **attributions**.
Availability	Bias in perceiving that memorable, unusual or adverse events are more likely (frequent) than they truly are.	Seek **base rate** of a diagnosis.	- Show **prior probability** (equivalent to SHAP bias) of diagnoses (in dataset).
Anchoring	Skewed perception of a value due to a supplied numerical value (anchor).	Avoid confirmation and early closure; make use of **lab tests to "prove" other** leading diagnoses. "Crystal ball" exercise (**"premortem"** prospective hindsight [51]).	- Show input **attributions** for multiple outcomes to allow *contrastive reasoning.* - Facilitate **counterfactual** to test *How To* reduce the probability of primary diagnosis with Rules (e.g., aLIME, LORE). - Facilitate **sensitivity analysis** with *What If* explanations to test stability of primary hypothesis.
Confirmation	Collecting redundant information to confirm an existing hypothesis, instead of finding evidence of competing possibilities.	- Use **hypothetical-deductive** method to assess value and role of contemplated tests. - Try to **disprove your diagnosis**, consider conditions of **higher prevalence**.	- Show Findings (**input attribution**) first, instead of Hypotheses (**output posterior probability**). ***Insight:*** *this is opposite to typical Machine Learning apps to show output uncertainty first.* - Show **prior probability** (equivalent to SHAP bias) of diagnoses (in dataset).

system [70], the system should be ***transparent*** to show What and Inputs explanations and show its classification ***certainty***. Once the user detects erroneous reasoning in the system, *scrutability* features can be ***contrasted*** to allow for model debugging and correction (by a technical expert).

4.4.6 Summary. We have described how we can use various XAI components of our framework and identify pathways to support good reasoning behaviors in users. This seeks to mitigate cognitive biases that may arise due to heuristic reasoning.

4.5 Summary of Framework

Our framework describes how people reason rationally (Section 4.1) and heuristically but subject to cognitive biases (4.3), how XAI facilities do support specific rational reasoning processes (4.2), and can be designed to target decision errors (4.4). The framework identifies pathways between human reasoning and XAI facilities that can help organize explanations and identify gaps to develop new explanations given an unmet reasoning need. XAI application developers can use our framework as follows: etc. Next, identify which explanations help reasoning goals (4.2) or reduce cognitive biases (4.4) using pathways in the framework (Figure 2 red arrows). Finally, integrate these XAI facilities to create explainable UIs. Furthermore, XAI researchers can extend our framework by 1) examining new XAI facilities (4.2) to understand how they were inspired by, dependent on, or built from reasoning theories (4.1) and 2) identifying common biases and reasoning errors (4.3) related to reasoning theories (4.1) and then identifying appropriate mitigation strategies (4.2) to select specific XAI facilities. For example, Informal Logic [99] could be integrated into reasoning theories (4.1) and informal fallacies [35] into reasoning errors (4.3).

5 APPLICATION: EXPLAINING MEDICAL DIAGNOSIS

We have developed our framework using a theory-driven approach and we next apply it by implementing an explainable AI-based early decision aid to diagnose patients in an Intensive Care Unit (ICU). Given the critical importance of ensuring correct medical diagnostic

Rôle de l'IHM

Pr

Gu

Étu

etc

When wrong	G7	**Support efficient invocation.** Make it easy to invoke or request the AI system's services when needed.	[Voice Assistants, Product #1] "I can say [wake command] to initiate."
	G8	**Support efficient dismissal.** Make it easy to dismiss or ignore undesired AI system services.	[E-commerce, Product #2] "Feature is unobtrusive, below the fold, and easy to scroll past...Easy to ignore."
	G9	**Support efficient correction.** Make it easy to edit, refine, or recover when the AI system is wrong.	[Voice Assistants, Product #2] "Once my request for a reminder was processed I saw the ability to edit my reminder in the UI that was displayed. Small text underneath stated 'Tap to Edit' with a chevron indicating something would happen if I selected this text."
	G10	**Scope services when in doubt.** Engage in disambiguation or gracefully degrade the AI system's services when uncertain about a user's goals.	[Autocomplete, Product #1] "It usually provides 3-4 suggestions instead of directly auto completing it for you"
	G11	**Make clear why the system did what it did.** Enable the user to access an explanation of why the AI system behaved as it did.	[Navigation, Product #2] "The route chosen by the app was made based on the Fastest Route, which is shown in the subtext."
Over time	G12	**Remember recent interactions.** Maintain short term memory and allow the user to make efficient references to that memory.	[Web Search, Product #1] "[The search engine] remembers the context of certain queries, with certain phrasing, so that it can continue the thread of the search (e.g., 'who is he married to' after a search that surfaces Benjamin Bratt)"
	G13	**Learn from user behavior.** Personalize the user's experience by learning from their actions over time.	[Music Recommenders, Product #2] "I think this is applied because every action to add a song to the list triggers new recommendations."
	G14	**Update and adapt cautiously.** Limit disruptive changes when updating and adapting the AI system's behaviors.	[Music Recommenders, Product #2] "Once we select a song they update the immediate song list below but keeps the above one constant."
	G15	**Encourage granular feedback.** Enable the user to provide feedback indicating their preferences during regular interaction with the AI system.	[Email, Product #1] "The user can directly mark something as important, when the AI hadn't marked it as that previously."
	G16	**Convey the consequences of user actions.** Immediately update or convey how user actions will impact future behaviors of the AI system.	[Social Networks, Product #2] "[The product] communicates that hiding an Ad will adjust the relevance of future ads."
	G17	**Provide global controls.** Allow the user to globally customize what the AI system monitors and how it behaves.	[Photo Organizers, Product #1] "[The product] allows users to turn on your location history so the AI can group photos by where you have been."
	G18	**Notify users about changes.** Inform the user when the AI system adds or updates its capabilities.	[Navigation, Product #2] "[The product] does provide small in-app teaching callouts for important new features. New features that require my explicit attention are pop-ups."

Rôle de l'IHM

Pr

Gu

Étu

etc

When wrong	G7	**Support efficient invocation.** Make it easy to invoke or request the AI system's services when needed.	[Voice Assistants, Product #1] "I can say [wake command] to initiate."
	G8	**Support efficient dismissal.** Make it easy to dismiss or ignore undesired AI system services.	[E-commerce, Product #2] "Feature is unobtrusive, below the fold, and easy to scroll past...Easy to ignore."
	G9	**Support efficient correction.** Make it easy to edit, refine, or recover when the AI system is wrong.	[Voice Assistants, Product #2] "Once my request for a reminder was processed I saw the ability to edit my reminder in the UI that was displayed. Small text underneath stated 'Tap to Edit'
	G10	**Sc** En te	
	G11	**M** En sy	
Over time	G12	**R** M ef	
	G13	**Le** Pe ac	
	G14	**U** Li Al	
	G15	**En** En en	
	G16	**C** Im fu	
	G17	**Pr** Al m	
	G18	**N** In ca	

Capabilities and Limitations

- Anchor performance metrics to human-relatable benchmarks
- Define accuracy precisely given multiple interpretations (e.g. binary benign vs. cancer, multi-class accuracy, percent tumor, etc.)
- Gather known human pitfalls (e.g. well-known edge cases) and report AI performance on those sub-categories
- Describe the diversity (or lack thereof) of the training data to inform generalizability
- Relate volume of training data to what is considered reasonable scale for machine learning
- Describe theoretical limits of AI, given current knowledge

Functionality

- Enumerate the inputs / context accessible to the algorithm, particularly inputs that are not shown in the interface (e.g. patient history)
- Specify the main steps in the AI's analysis of its inputs, vis-a-vis steps taken in typical human analysis (e.g. multiple magnification levels, decomposition of input into sub-images)
- Compare and contrast AI schemas relative to known human decision-making schemas (e.g. the extent to which it has been explicitly trained on higher-level biological concepts)

Medical Point-of-View

- Show subjective thresholds of the model (e.g. examples of AI judgment on borderline cases)
- Include a human-AI calibration phase, for users to calibrate their own subjective thresholds to that of the AI, with an authoritative source provided as ground truth
- Specify where the algorithm received its medical source of ground truth (e.g. expertise level

Le cas de la robotique chirurgicale

Laparoscopie

Robot DaVinci

“Merveille" technologique

- Ultra précis
- Confortable

Inconvénients

- Cher
- Bénéfices cliniques?

Analyse du système

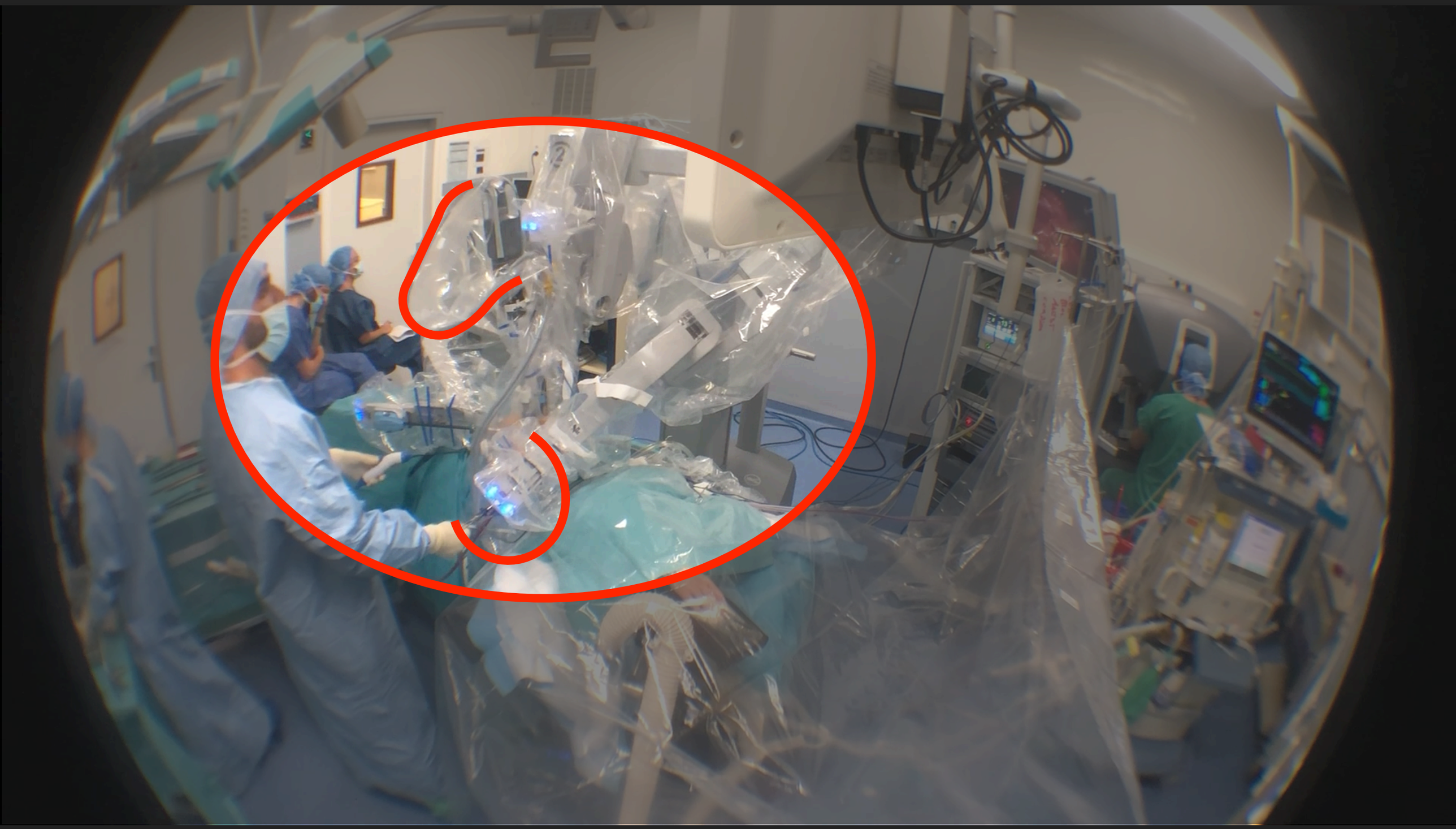
- Sur le terrain
- Observations, interviews, etc.

Un chirurgien autonome mais seul...

Conscience limitée de ce qui se passe autour

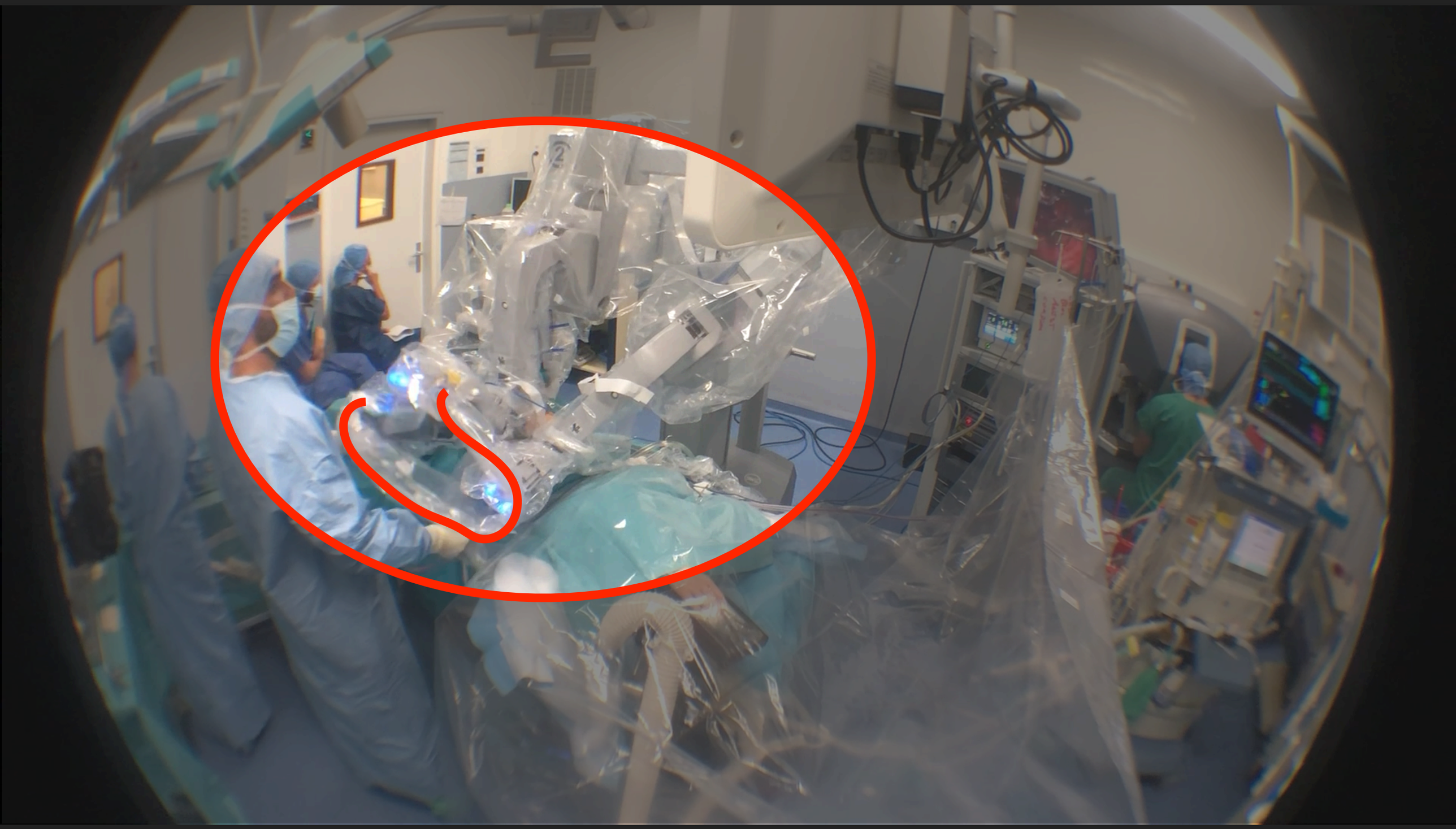
Un chirurgien autonome mais seul...

Conscience limitée de ce qui se passe autour



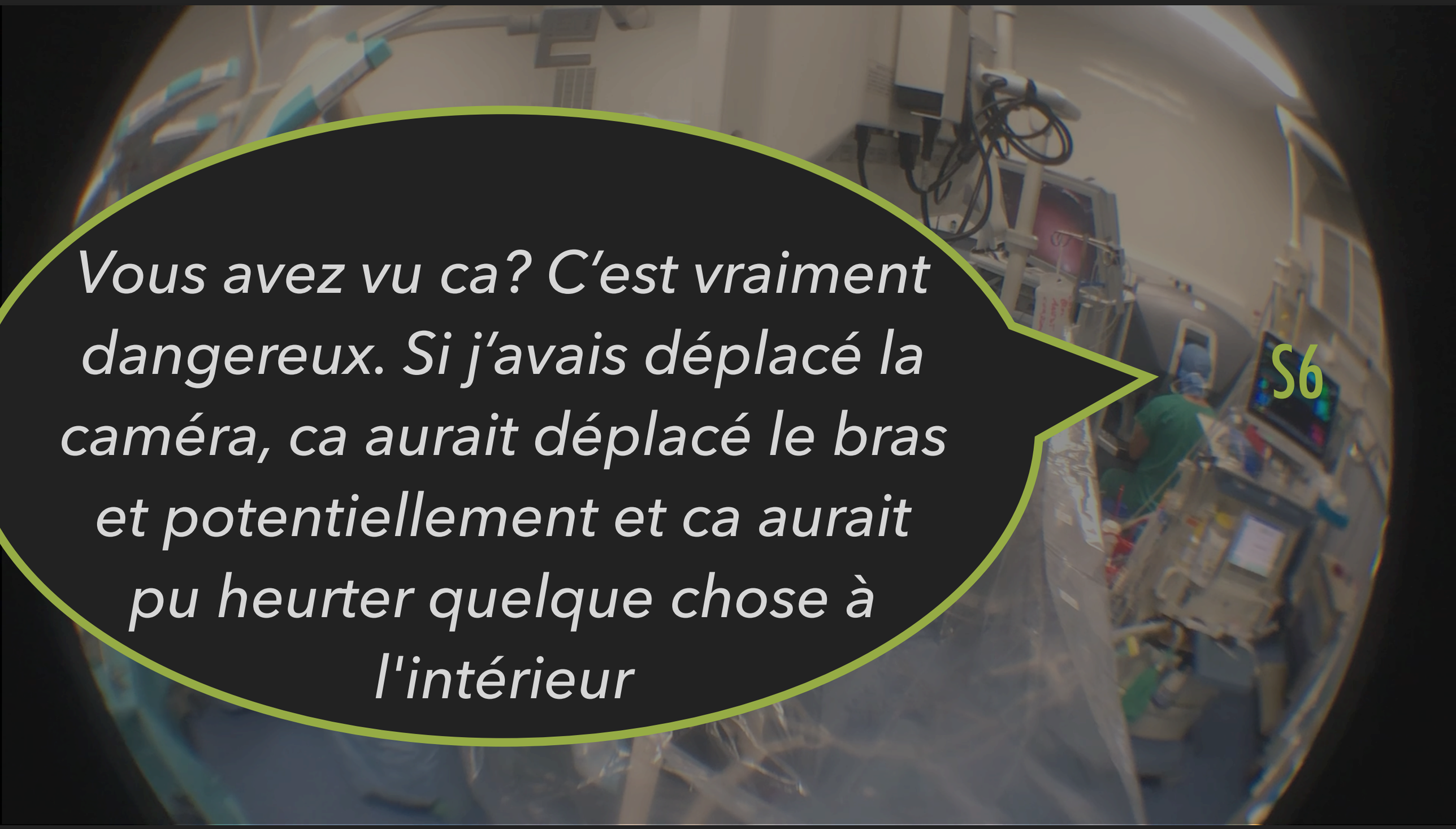
Un chirurgien autonome mais seul...

Conscience limitée de ce qui se passe autour



Un chirurgien autonome mais seul...

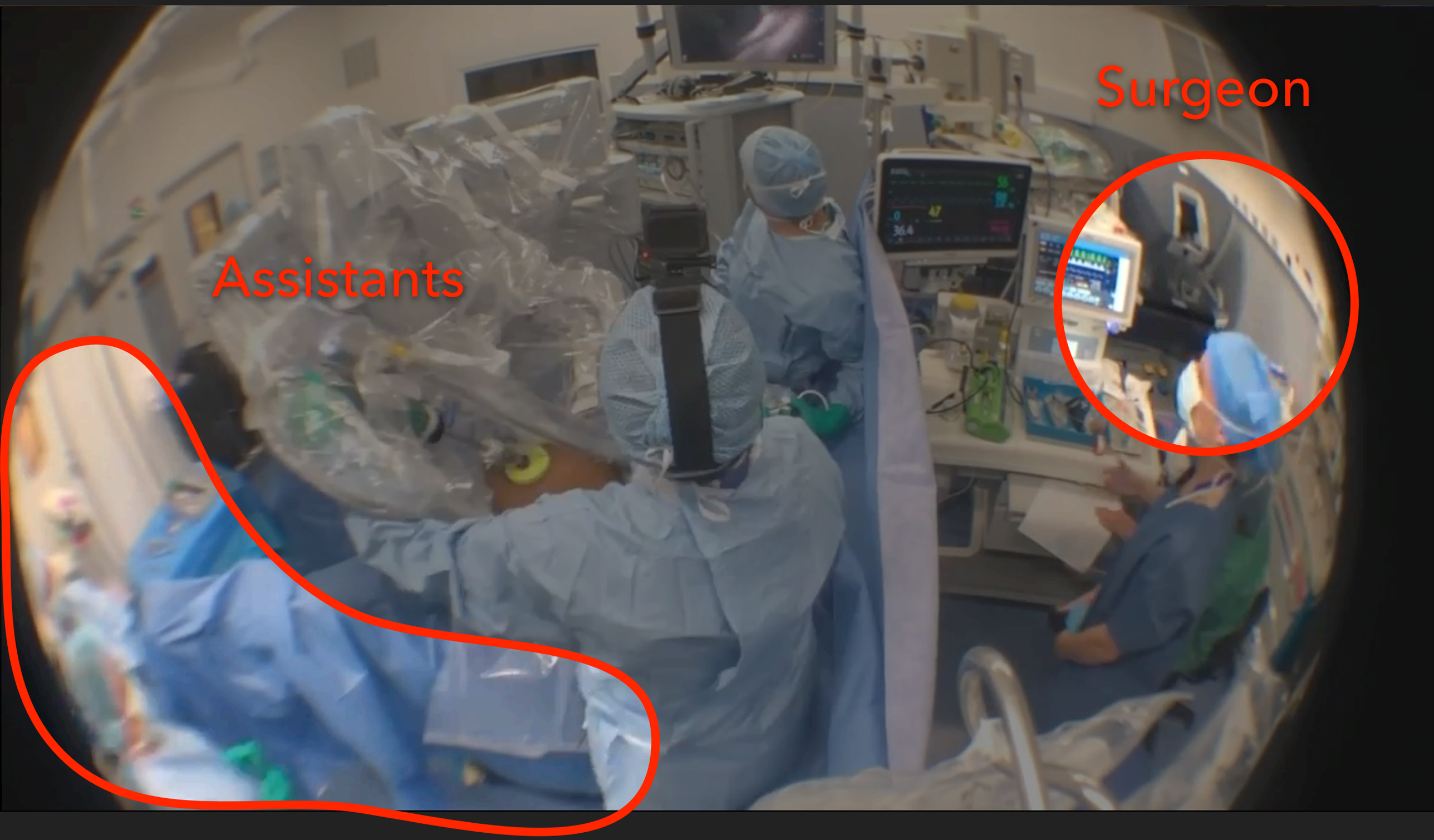
Conscience limitée de ce qui se passe autour



Vous avez vu ca? C'est vraiment dangereux. Si j'avais déplacé la caméra, ca aurait déplacé le bras et potentiellement et ca aurait pu heurter quelque chose à l'intérieur

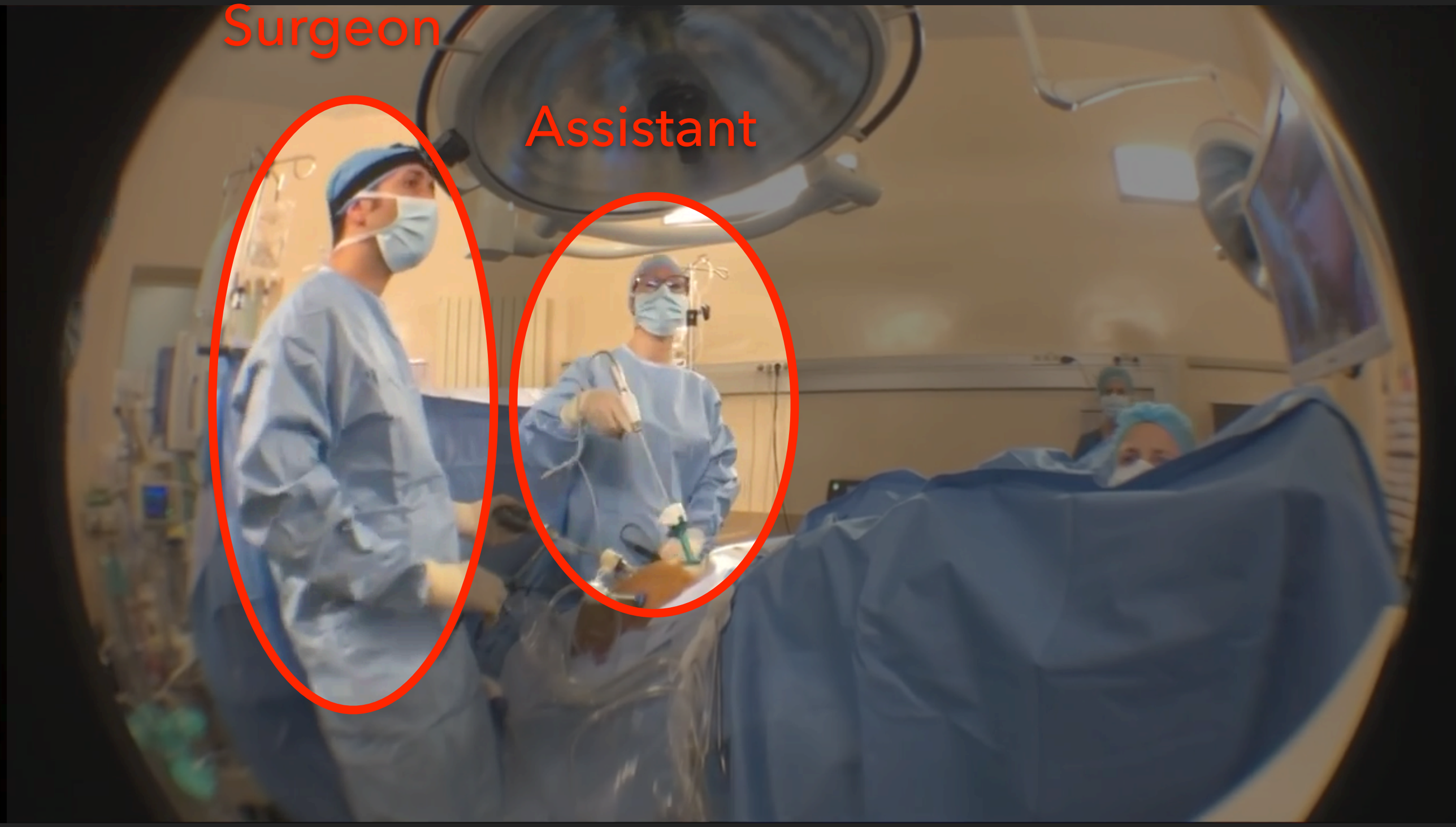
De l'assistant du chirurgien à l'assistant du robot

Les assistant font moins de gestes chirurgicaux



De l'assistant du chirurgien à l'assistant du robot

Les assistant font moins de gestes chirurgicaux



l'IHM pour l'IA:

le cas de la robotique chirurgical

De vrais avantages

- Confort du chirurgien

Avantages pour les patients

- À prouver…

Mais aussi des inconvénients (plus difficile à percevoir)

- Un chirurgien isolé
- Transformation du role des assistants
- Transfert de compétences plus difficile

Recommendations :

- Aller plus que loin que simplement la précision d'un système
- Penser à tous les acteurs
- Conception centrée utilisateurs
- Observations, interviews, questionnaires, étude sur le terrain, etc.

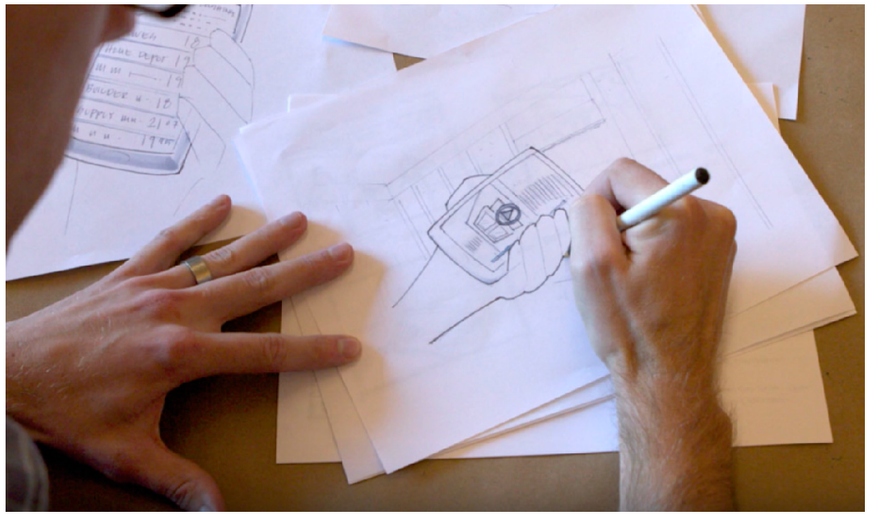
IHM

IA

Conception d’interfaces

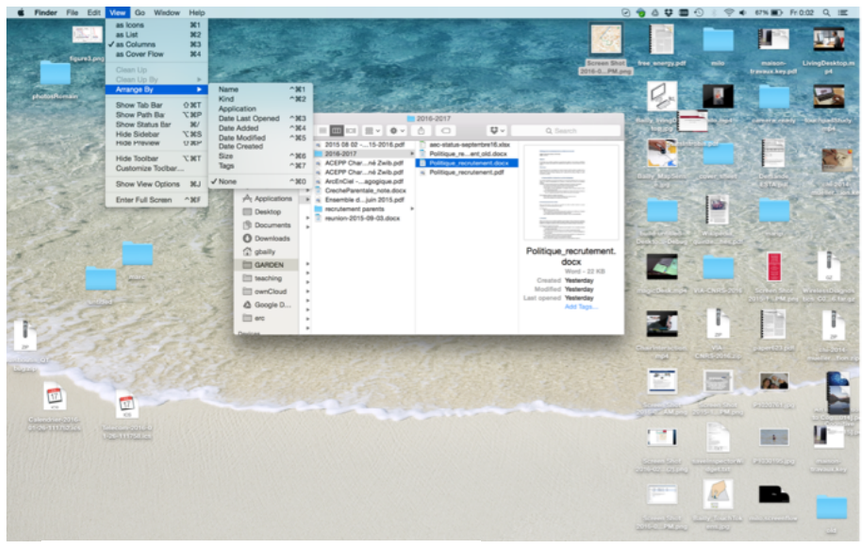
Aujourd’hui

- Conception manuelle
- S’appuie sur l’expérience
- Cher et long



Problématique

- Milliard de possibilités
- Quelle est l’interface optimale?



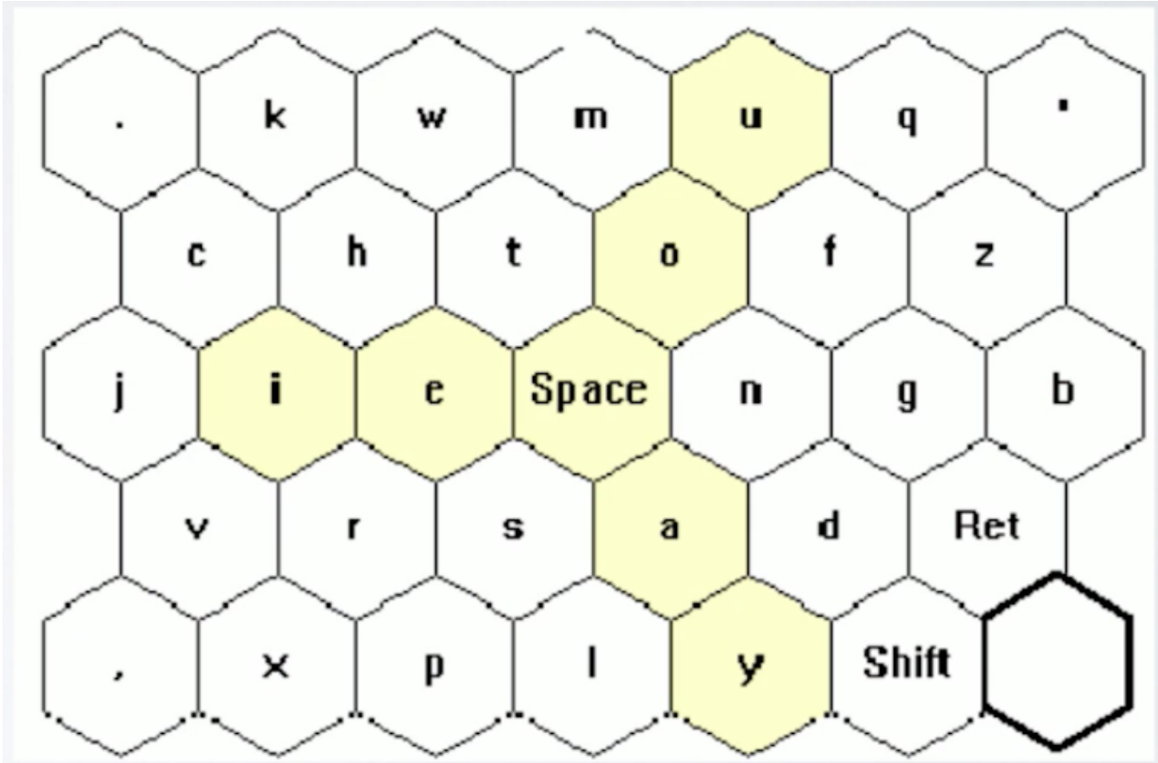
Défi

- Génération automatique d’interface

Conception d’interfaces

Clavier particulier

- PDA + Stylet



Optimisation interactive

Nouveau clavier AZERTY

Quelle est la clé

Pour générer automatiquement des interfaces?

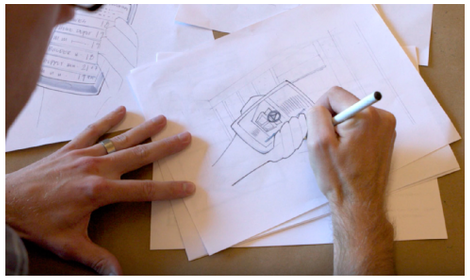
- clavier, menus, page web, etc.

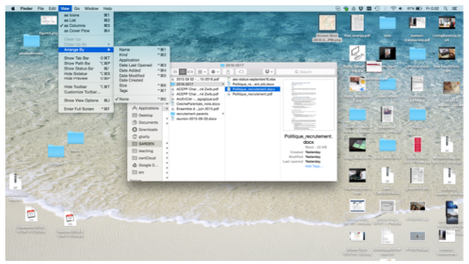
Modèle de performance

- Loi du comportement & Machine Learning

Conclusion

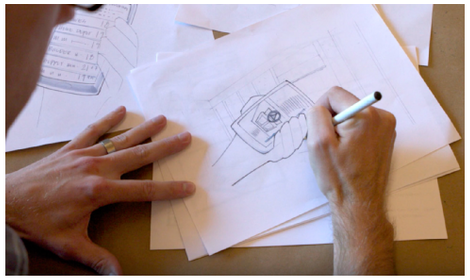
Concevoir

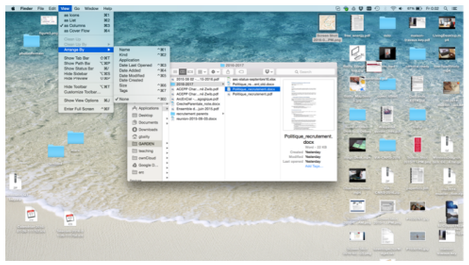




Conclusion

Concevoir → Utiliser

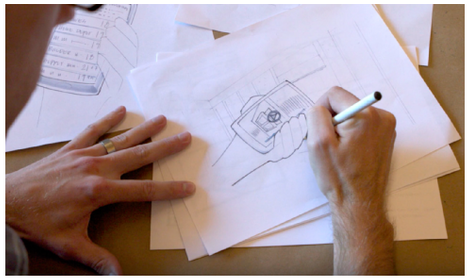


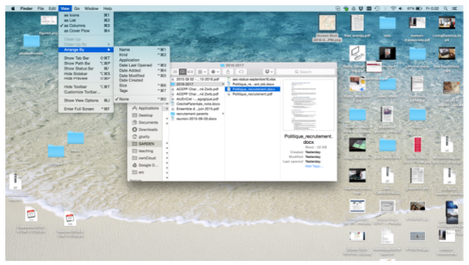


Conclusion

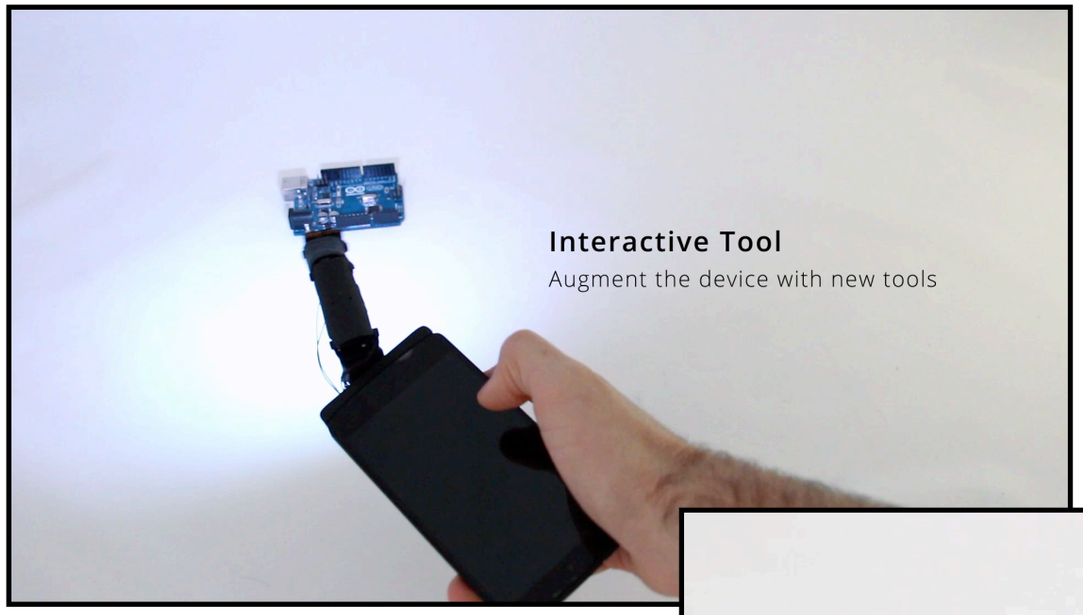
Que pourraient être nos dispositifs demain?

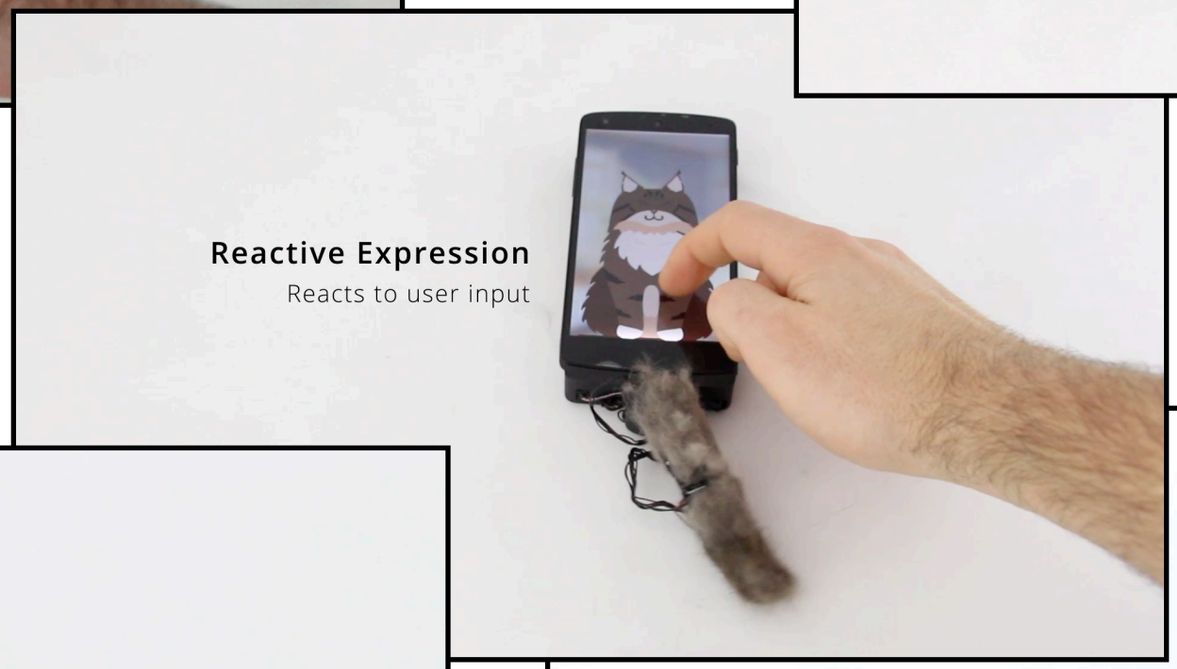
Imaginer → Concevoir → Utiliser

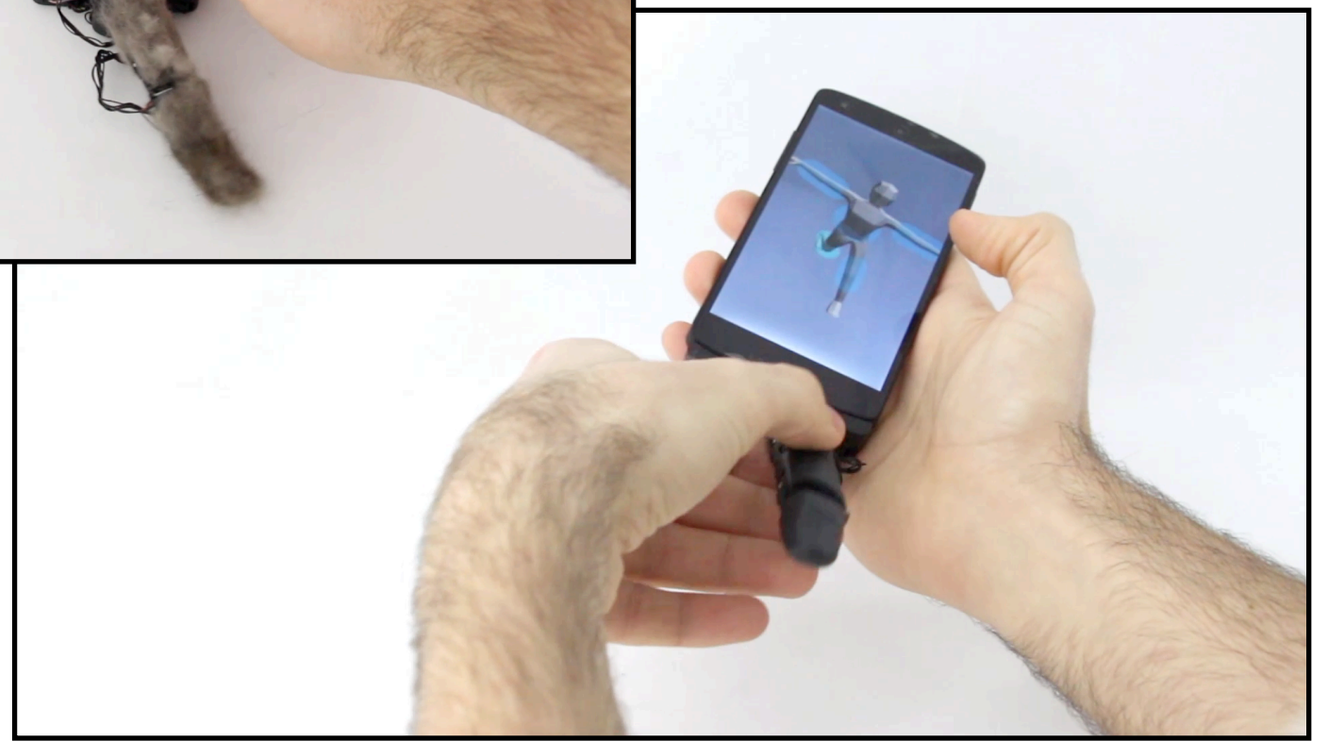




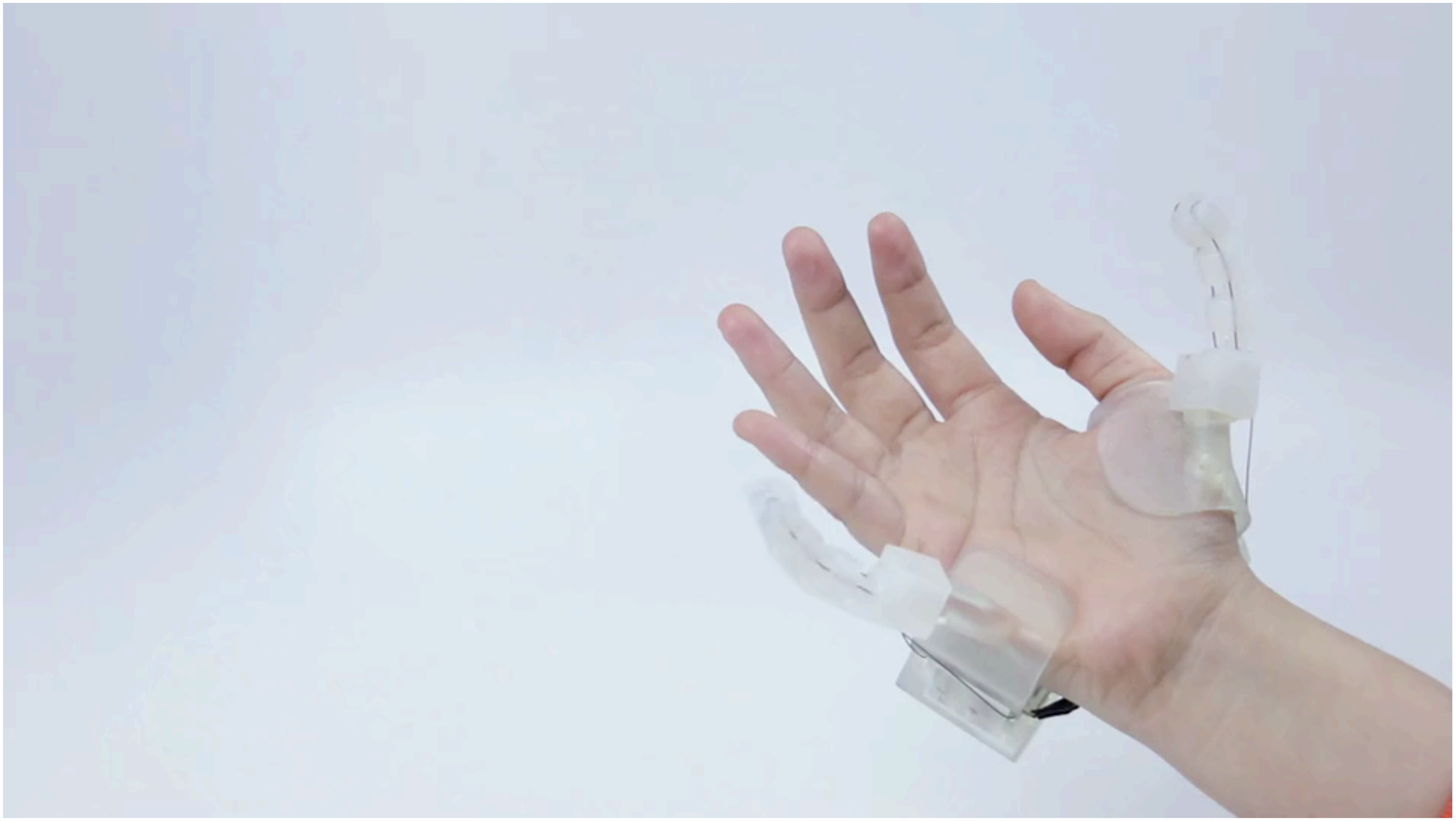
Telephone Intelligent





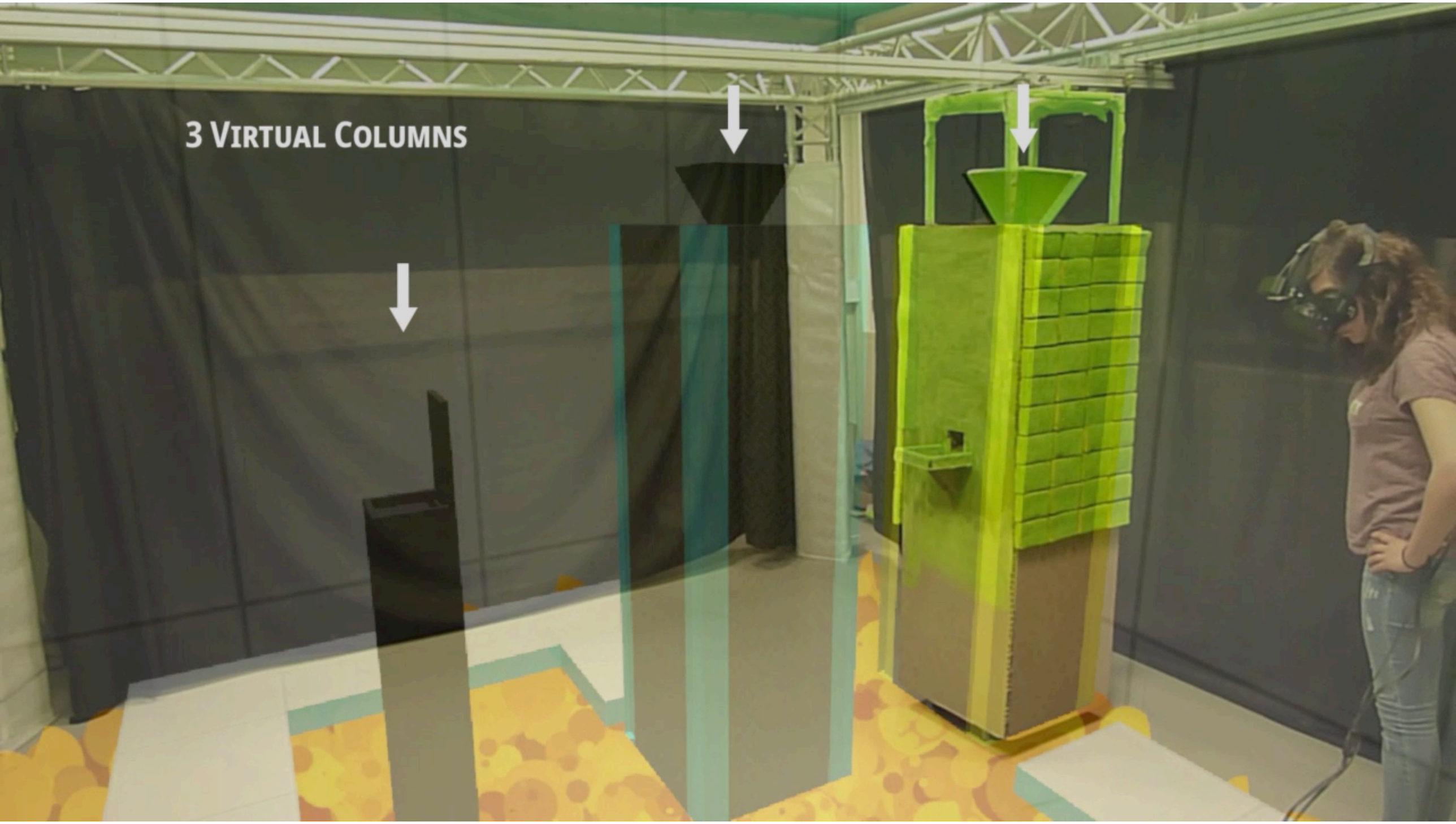


L’humain augmenté



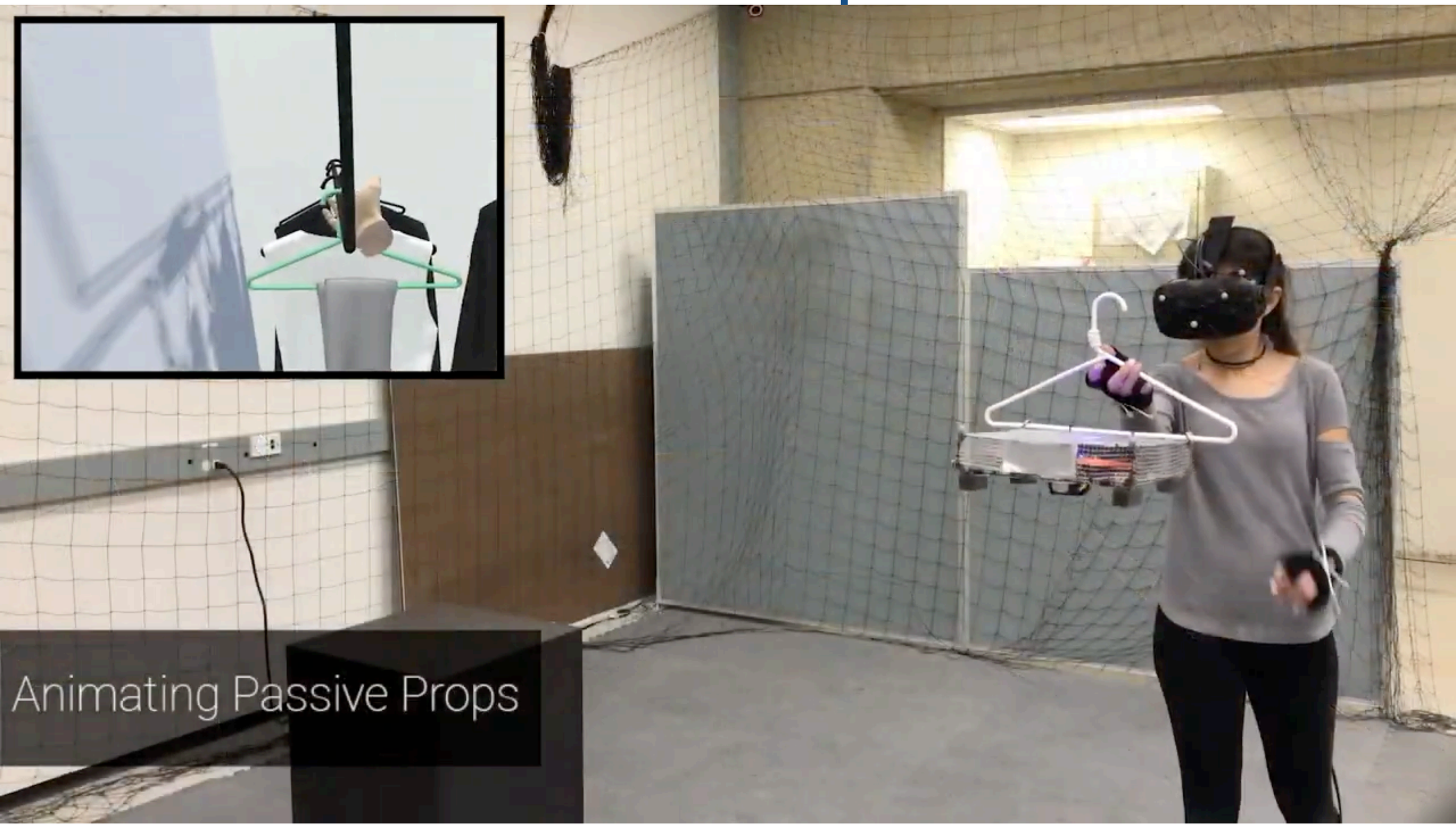
Y. Hu, S. Leigh, P. Maes (ACM **CHI’17**)

Hand Development Kit: Soft Robotic Fingers as Prosthetic Augmentation of the Hand


3 Virtual Columns


MARSHALL

Quadcopter

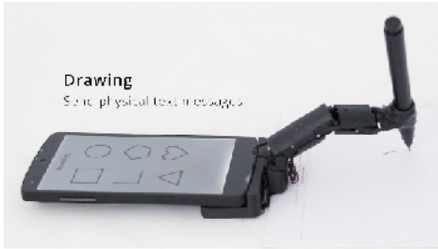


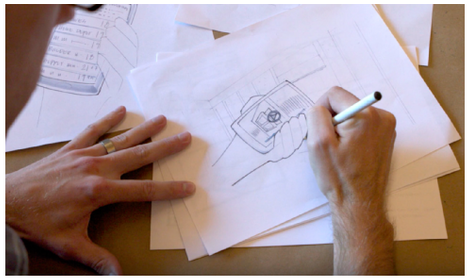
P. Abtahi, B. Landry, J. Yang, M. Pavone, S. Follmer, J. Landay (ACM **CHI'19**)

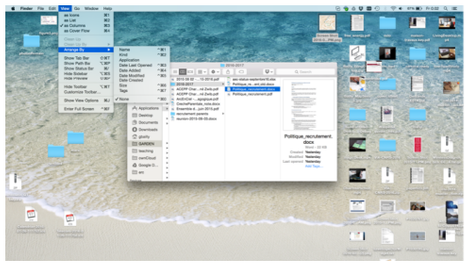
Beyond The Force: Using Quadcopters to Appropriate Objects and the Environment for Haptics in Virtual Reality

Conclusion

Imaginer → Concevoir → Utiliser



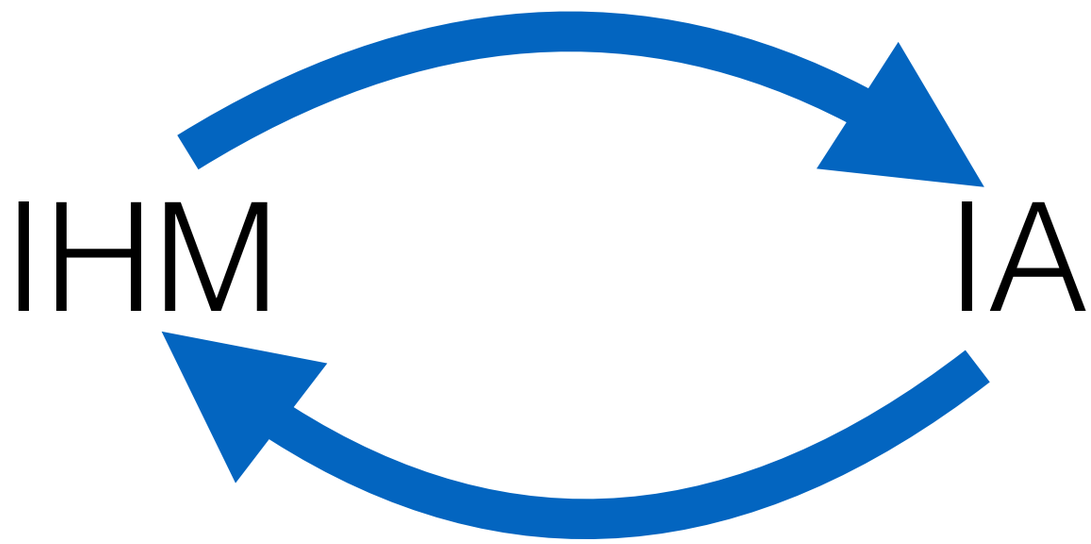




Synergie entre IHM & IA

Gilles Bailly
www.gillesbailly.fr

Conception centrée humain



Imaginer, concevoir, utiliser

SORBONNE UNIVERSITÉ
CRÉATEURS DE FUTURS DEPUIS 1257