

Inhabiting the Invisible

by

Paul Van Rijn

A thesis exhibition presented to OCAD University in partial fulfillment of the requirements for the degree
of Master of Fine Arts in Digital Futures

OCADU Waterfront Campus, 130 Queens Quay East,

March 26 to 30, 2026

Toronto, Ontario, Canada

Abstract

Since the mid 2010s social networking companies like Facebook and search engines like Google have implemented AI tools to improve and expand the capabilities of their ubiquitous platforms. This represents an important shift in our relationships to these technologies and poses questions for how we define our future interconnections (personal, social, political) with the digital realm. *Inhabiting the Invisible* is an interactive and immersive media installation that takes an aesthetic approach to defining our relationship with this realm - can we perceive these changes in both a visual and insightful sense? Have these changes informed how the platforms perceive us?

Inhabiting the Invisible explores the creative opportunities that these AI tools afford through many cycles of prototyping within a creation-as-research methodology.

Although the installation does not seek to provide any single answer to these questions, the intention is to provoke a response in the mind of the participant. This work, incorporating the AI tools used by the platforms, invites and even perhaps coerces participants to enter it and look into the digital realm.

Acknowledgements

Connie, my sister who has talked with me about art a lot over many years.

Simone Jones who has had to deal with some surprising things on my behalf as Program Director, and as a mentor and guide on my thesis project.

Parantap Bhatt and Bently Jarvis as my advisory team have been very perceptive and provided critical guidance when I needed it.

My cohort has been marvellous. Every day they steered me out of the ditches.

The Digital Futures Faculty have been universally fine teachers and have been tolerant of my profound ignorance of anything that has happened in the last few decades.

Simone Jones, Tom Regehr, Tien Van Rijn, Gerald Grison, and Jason Foley helped with the design, construction and cleanup of the final piece. Thank you.

Dedication

I dedicate this work to my late wife Olga Van Kranendonk, and my children Tien and Minna, who I love as life.

Table of Contents

Abstract.....	2
Acknowledgements.....	3
Dedication.....	3
Table of Contents.....	3
Introduction.....	9
Issue and Contribution.....	9
Scope and Limitations.....	10
The Interactive Installation.....	10
Relevant Artworks.....	11

Face Detection and Recognition Case Studies.....	11
Behaviour and Agency in Installation Works.....	12
Physical Environment.....	13
AI formed by the artistic process formed by AI.....	14
AI with Interaction.....	15
Germane Theory.....	16
Behaviour and Agency.....	16
Behaviour.....	16
Behaviour in Inhabiting the Invisible.....	17
Agency.....	17
Agency in Inhabiting the Invisible.....	18
Systems Approach.....	18
Physical Environment.....	19
The Digital Realm.....	20
Interaction.....	20
Methodology and Methods.....	21
Creation as Research.....	21
Methods.....	22
Iterative Prototyping.....	22
Initiation	22
Creative Cycle	22
Results	23
Prototypes	24
Gathering Concepts, Experiences, Influences.....	24
System Element: Portrait acquisition, processing and display.....	24
Emergence.....	24
Context.....	26

System Element: Interaction and Agency.....	27
Emergences.....	27
System Element: Presentation.....	28
Emergence.....	29
Context.....	29
Integration of Prototypes.....	30
Emergence.....	30
Context.....	31
SPARK Presentation.....	31
Emergences.....	34
Self-Presentation vs Self-Preservation.....	34
Interaction.....	35
Time	35
Context.....	35
Implementation and Experimentation.....	36
Gesture Recognition.....	36
Kiosk.....	36
Face Behaviour.....	37
Projection	38
Emergences.....	39
Software Design.....	40
Conclusion.....	40
Exhibition Commentary and Documentation.....	41
Work Cited.....	41
Appendix A: Software and AI References.....	45
Deep Neural Networks.....	45
Generative Adversarial Network (GAN).....	45

AI Tools Used in the Project.....	46
Deepface.....	46
Mediapipe.....	46
Kyutai Pocket TTS.....	46
Torch.....	46
Whisper.....	46
Other Software Tools.....	47
python.....	47
OpenCV.....	47
pytsx3.....	47
Unity Game Engine.....	47
Appendix B Final Installation.....	47
Physical Design.....	47
Location.....	47
Overview.....	48
The Screens.....	48
The Projectors.....	48
The Kiosk.....	48
Electrical and Other Wiring.....	51
Safety.....	52
Floor Plan.....	54
Images of the Exhibition.....	55

Table of Figures

- Figure 1: Richard Serra "My Curves are not Mad" 14
- Figure 2: Process and outputs diagram 23
- Figure 3: Face detection and information derived from an image by the deepFace AI showing dislocated facial points..... 25
- Figure 4: Incorrect age and emotional estimates derived from the image by deepFace..... 26
- Figure 5: Blue Gaze bird with programmed behaviours related to AI derived data..... 27
- Figure 6: Your Voice, (2025) a large-scale projection prototype project..... 28
- Figure 7: Uninteresting Display..... 30
- Figure 8: Overview of SPARK test presentation implementation..... 32
- Figure 9: User testing gesture-based image acquisition for portraits..... 33
- Figure 10: A participant’s face at a large scale. This also demonstrates the “mugging” behaviour..... 34
- Figure 11: Testing elements of the kiosk live with portrait projection..... 37
- Figure 12: Test to see face motions at scale..... 38
- Figure 13: Projection over distance tests to determine how large the screens might be based on the size of the exhibition space and the geometry of the projectors..... 39
- Figure 14: Entry kiosk with video camera and microphone..... 49
- Figure 15: Desk with computer and a back view of the kiosk made from movable panels..... 50
- Figure 16: Scale and placement of screen..... 50
- Figure 17: Floor plan of the exhibition space..... 54
- Figure 18: Screens illuminated by projectors, placed to make a corridor..... 55
- Figure 19: A viewer in the corridor between the screens..... 55
- Figure 20: Typical viewer reaction – standing against the walls..... 56
- Figure 21: Typical viewer reaction – identifying each other and themselves..... 56

Introduction

Issue and Contribution

In the mid 2010s software tools described as Artificial Intelligence were integrated as a key component of many very large digital platforms. This project will look at the technologies included in two of the largest platforms operating at that time and investigate how those AI technologies and others have been integrated into specific artworks.

Did our relationship with the digital realm change with this new way that they used our faces and our data that, in the words of Crawford and Paglen, create “the stark power asymmetry at the heart of these tools”? (Paglen and Crawford 2020) In their work *Training Humans*, Paglen and Crawford investigate how the datasets of images used to train computer vision and AI systems were created and how they are used. The power asymmetry is established when the subjects, those whose faces and data are incorporated into the databases, are represented in the final systems with all the motivations, history, politics, prejudice and ideology inherent in human endeavours, and are employed by two of the most pervasive and prosperous organizations in the world.

The concerns of invisible ubiquity, loss of agency and the general public’s lack of comprehension of the use of AI in these platforms will be expressed through the conceit that *Inhabiting the Invisible* is a window into the digital realm. The installation will not confer any agency to participants over how faces, data and daily activities are acquired and processed, but it might induce them to question their own relationships with the digital realm. Do they consider their role, or at least the role of their faces and data, in this power asymmetry? How might we question our perception of the digital realm and the roles of AI within it through a creative, interactive digital installation? What aesthetic affordances can be discovered through this creation-as-research process?

My creative process examines how those AI tools were made, and how I can use those tools to create an interactive installation that engages participants to provoke an examination of how they relate to these tools. My aesthetic concerns grew out of repeated cycles of creation and review - what “looks good”, what evokes in me the feeling that I hope to evoke in others. With every iteration, I have created processes and artifacts that might be of interest to others. I have endeavoured to break through a simple use of the tools to aim at a more nuanced emotional evocation of the wonder and unease that I feel about these platforms through scale, colours, behaviours, and control of the physical environment of the work.

The installation reveals through experiential perception the asymmetry between these ubiquitous AI driven systems and the people who use them.

Scope and Limitations

There are many important issues related to the widespread implementation of AI capabilities in the big technologies that increasingly govern our lives (Wu 2025). This project focuses on two specific platforms at a time when, in the opinion of observers of technology (Azizi 2015), these platforms implemented AI tools to solve problems and create opportunities for their businesses. This was the beginning of the epoch where AI tools went from experiments to widespread implementation and this directly affected billions of people (BBC 2015).

The Interactive Installation

My intention is to create an interactive installation that is engaging and will invoke attentiveness as described by Alois Reigl via Margaret Olin (1989). My experience is that people enjoy seeing their faces so to do this; the work uses faces selected from the audience using AI tools as an initial attractant.

Viewers passed through a constricted kiosk that contained lights, a camera and a microphone. This isolated them from any group they might have been in. The small scale of the kiosk will constrain and mildly coerce them to follow the process and perhaps experience some anxiety.

The system separated the faces of people in view of the camera from the rest of their bodies and their backgrounds in a form of digital vivisection. The portraits were digitally degraded then neutralized into shades of translucent white. This leaves them recognizable but “normalized” in the mathematical sense that the numbers that represent them are expressed in terms suited to the system rather than suited to a person for viewing.

Two rear projection screens were placed in the center of the exhibition space. The screens are meant to be large scale like the platforms in the metaphor of seeing the digital realm. This is intended to be both attractive in the sense of wonder, as well as repulsive in the sense of being overwhelming to the viewers. As the faces eventually got projected on the screens the viewers might have been surprised to see themselves or their friends display on a large scale.

The screens were placed close to each other leaving a corridor that some viewers choose to investigate. The gap between is narrow and the screens converge slightly overhead. The intention is for the experience to be fully immersive and overwhelming and perhaps even a little exhilarating.

The backgrounds of the screens are rendered in deep blue and the portraits in translucent shades of white. These colours are the official marks of IBM, Facebook and Google. The colours are associated with trust and comfort. (su 2019)

The faces move in the display with an attractive floating, pulsing motion. This is meant to be soothing despite the anomalies like colour, scale and other distortions. The installation demonstrated that participants did spend some time to contemplate why they are being treated this way.

Technical and physical elements of the work like supports, camera, microphone, projectors, screens and even a computer will be visible to the participants. The intention is to reveal what is being done to their images and data in a large system, as opposed to the slick interfaces that hide the true intentions of the digital realm. It is also a bit of fun as an imaginary peek behind the curtain.

Relevant Artworks

Face Detection and Recognition Case Studies

Paglen and Crawford (Paglen and Crawford 2020) comment directly about the classification of humans by AI and “the stark power asymmetry at the heart of these tools” in their work *Training Humans*. The work engages with the imagery used as the dataset and therefore the foundation of facial detection and recognition as used in my work.

The rock band Massive Attack made a commentary about the use of face detection and recognition at a recent concert in London (Bandara 2025) by taking live videos of people at their concert and projecting their images on a screen behind the band during the performance. The band also showed details like workplace, age, gender etc. of the person being displayed. The band’s apparently negative commentary might be undermined by the delighted reactions on most people’s faces when they see themselves projected on the screen behind the band. The audience

appears to be more interested in “self-presentation” as theorized by Erving Goffman (Bahar 2024) over “self-preservation” from a mendacious technology.

I encountered the same response in almost every prototype and iteration in the development of the display elements of this project. The result is that I have taken care to separate the interaction and depictions of faces. First, by giving the portraits their own behaviour unrelated to the actions of the viewer and second, in time by delaying the appearance of the viewer’s face in the display from the time it was acquired and processed.

Behaviour and Agency in Installation Works

Interactive environments like *Rain Room* (Scott 2013) where participants seemingly control the weather by moving around, self-motivated artworks like Norman White’s *Helpless Robot* (Transmediale 1986) where participants engage in a physical interaction with a demanding robot, demonstrate behaviours that engage us through interactive interplay. Those behaviours might be intelligent or might be just a few sensors and “a cleverly coded piece of software” (Markus 2014). These works do not integrate any technology that might be considered AI today but demonstrate how a work can contain behaviours and demonstrate autonomy.

Helpless Robot in particular demonstrates behaviours that change with interactions over time. White himself described this work “the possibility of simulation” which is how I describe *Inhabiting the Invisible* - it is an attempt at simulating and expressing a vast complex system in a small simulation.

In *Recurrent Kafka*, Rafael Lozano-Hemmer creates a digital “mirror” with unexpected non-mirror responses (Lozano-Hemmer 2025). He has imbued the reflected face, through the use of face detection and a generative re-creation of the viewer’s face, with behaviours that are simple but unsettling - the mirror does not look back at the viewer as we expect in a mirror or even in a face to face interaction with another person, but looks only at the text that appears on the screen. It is a comment on our expectations of reflection and interpersonal interactions and of digital systems that “represents us, poorly, poetically, and without consent”. This piece evokes the feeling of disquiet and disconnection while still maintaining a connection with the viewer through the mirroring aspects like head pose, but subverting that through the generative re-creation and redirection of the eyes.

Inhabiting the Invisible implements behaviours in acquiring, cropping and modifying images, and applying some motions and interactions between the resulting portraits. These behaviours create a world that contains the faces of the participants with the technical, in terms of identification, and other aspects that we might associate with faces, but is disconnected from the owners of those faces in that they have no agency over them once they have been assimilated. The intention is to express the disconnection of the digital realm from its users.

Physical Environment

In *Changing Light Corridor with Rooms* (1971) Bruce Nauman explores “a manifestation of the artist’s fascination with the affects that physical environments can have on people, especially the feeling of unease that comes from being in a space that is too compressed or too large” (Gryczkowska 2011). Nauman reconstructs the gallery space into an environment that controls and directs participants using walls, lights and open spaces.

In his large and immersive sculptures Richard Serra’s interest includes the subjective bodily and visual experiences his work evokes. His installation *My Curves Are Not Mad* is made of a pair of 14-foot tall and 44-foot-long curved sections of steel placed so that they create a corridor that controls the viewer’s motion within the piece. The narrowness of the corridor is both intimate in its narrowness and exhilarating in the large scale of the sides.



Figure 1: Richard Serra "My Curves are not Mad"

These works play an important role in the design of the physical environment of *Inhabiting the Invisible* because they purposefully evoke the feeling of unease and subtle coercion that define some of the affective goals of this project. *My Curves Are Not Mad* also evokes a feeling of exhilaration because of the large scale of the work and intimate proximity of the viewer which is intended to echo the scale of the digital realm.

AI formed by the artistic process formed by AI

Anna Ridler incorporates a Generative Adversarial Network (GAN) into her work *Mosaic Virus* (Ridler 2019). The video component of the piece was generated from Ridler's meticulous and beautiful collection of 10,000 photographs of tulips that she created for the work, modulated by the price of cryptocurrency via the GAN.(Ridler 2020) Her perspective on the "real world" data derived from the photographs she took and classified and the questionable quality, ownership and social relevance of widely available data sets.

She discusses the process she used in the creation of elements of *Mosaic Virus* and concludes that the physical image preparation, the tying together of many layers of abstraction and creation of the video would not have been possible without the machine learning tools in the GAN.

She has complete control over the source of her data (her photographs, crypto market values), the annotations of the images (where human judgment is applied), and the “tuning” of the GAN, where she admits with some trepidation, that she does not have any control over what the AI produced from these inputs.

This work clearly describes the self modifying nature of AI tools, and how much control the “creator” of the system has with initial conditions and system results (data, testing, post processing etc.) implying that they are created with intention, a definite goal for their operation, and an aesthetic result that the artist wants despite the “intelligent” elements of the final result as created by the generative AI.

This work describes how sophisticated AI technologies can be integrated into an expressive process. *Inhabiting the Invisible* uses a stream of images and AI processing as Ridler did in *Mosaic Virus* but does not use a generative tool to create the final imagery because those techniques did not become generally available until after the era that my work is exploring. Those generative tools do present excellent opportunities for expression and commentary.

AI with Interaction

Crawford and Paglen (Paglen and Crawford 2020) in their exhibition of photographs and other visual media *Training Humans* examines how images of people have been used as “datasets” for training computers to recognize human faces. They use an “out of the box” code for face detection and motion/emotion tracking to demonstrate to their audience the results of the recognition effort. The viewers see themselves in a video feed with a real time presentation of the location of their faces and information that the system has derived from their images. Seen in context with the faces of people the display is both fascinating and a little disturbing because it becomes clear that the training has worked. The interaction here is direct - as the audience appears and moves around, the video display clinically shows what it has derived.

I interpret this use as a polemic or information exchange, like a documentary, that directly addresses the viewer.

Rafael Lozano-Hemmer has created many fascinating works that explicitly include AI as part of their construction and meaning. As described above in *Recurrent Kafka*, Lozano-Hemmer uses face detection and reconstruction as interactive tools.(Lozano-Hemmer 2025) The interactivity is a key aspect of the meaning in that some of the viewers' actions, e.g. head pose, are interactively reflected but other expected actions are not, e.g. eye motion.

In *Inhabiting the Invisible* the interaction is delayed, as such interaction is delayed inside the digital realm. Both Facebook and Google have interfaces that respond in real time to provide search and social connection services, but from the business perspective, value is derived from the slower cycle of acquiring, processing and selling information about the user to generate advertising revenue. This process is invisible to users except that perhaps the advertisements in the platform or in subsequent sites seem to be aligned to the users' interests.

The generative elements of these pieces are fascinating, but they are outside the scope of *Inhabiting the Invisible* because they represent a later phase of AI development. This is a rich territory for future work.

Germane Theory

Behaviour and Agency

Behaviour

Behaviour is a difficult phenomenon to define (Uher 2016). I will take my definition from Uher with the understanding that the definition is anthropomorphic and is defined in terms of our perceptual and conceptual abilities. Behavior in an entity should have the following properties as described by Uher.

These particular properties are considered because they determine a given phenomenon's perceptibility by humans in everyday life—and thus also by researchers.

These properties are 1) spatial location in relation to the intact body of the individual under study (e.g., internal, external), 2) temporal extension (e.g., momentary, temporally extended) and 3) spatial extension, conceived in terms of the two complementary properties physicality (i.e., spatially extended) versus “non-physicality” (i.e., without spatial properties).

The question that concerns me is where in the system does the behaviour occur? In this project the location of the behaviour can include the audience, the physical elements of the system and software/compute components. The temporal extension will be in terms of human perception. Since I am including computers running software I will consider that “non-physicality” will include phenomena that are happening in software; that software might be also activating some physical elements of the system such as a projector.

Behaviour in Inhabiting the Invisible

Inhabiting the Invisible incorporates behaviours in the usual way we expect software systems to work by operating the camera and projectors, moving and manipulating data, and continually assessing to ensure the correct operation of the system. It also gives behaviours to the portraits it has created as though they are individual, separate entities. The realization/visualization of these behaviours is done via some basic rules and randomizations in the Unity Game Engine.

Agency

Agency is apparently an even more difficult term to define. From the Stanford Encyclopedia of Philosophy “Usually, though, the term ‘agency’ is used in a much narrower sense to denote the performance of intentional actions.” (Schlosser 2019). This concept of agency has been the standard model from Aristotle, Hume to the mid 20th century, when thinking began to focus on the intentionality of action. This has resulted in the claim that there are different kinds of agency, including an entire realm of the metaphysics of agency.

Given that one concern of this project is to know where the agency lies in a system, and that this project is a physically implemented work that involves humans and complex software, this document will constrain the definition of agency for the implementation of the project to the initiation of a behaviour consisting in causation by the intentions of the agent. This concept is based on the concept, and perhaps theory, where agency is defined as initiation by the agent.

For example, if a person looks at something or waves at it, there is a question about those phenomena as behaviours initiated by that person, therefore they are the agent. It can be argued that if the person is

waving back at someone else there is a casual chain involved. This chain is the concern of this project's questioning of agency inside AI systems and even in the physical implementation of an art installation. Where in the chain do we, or can we, have agency?

Agency in Inhabiting the Invisible

The choice of creating an interactive, immersive installation, and the visibility of all of the components, confers some agency on the viewer that they do not typically have. It offers the possibility of examining the invisible digital realm which is one of the main ideas of the piece.

However, I have designed into the physical components of the work elements that have been shown to exhibit strong affect on viewers. This is a reversal of the promise offered by the title and the form of the piece, but the stated goal is to evoke some questioning of our relationship with the digital realm.

The portraits have been given some level of agency in that they are free to wander around in the space offered by the Unity Game Engine. The observer, whose face may make up the portrait being observed, is left out - they have no agency here - they cannot control or alter the data related to them. The agency remains with the system.

This is another main theme of the piece - we have no agency once our data have been assimilated.

Systems Approach

The technical and conceptual elements of this project are conceptualized, designed and integrated via the systems approach described by Jack Burnham (Burnham 1968).

But for our time the emerging major paradigm in art is neither an ism nor a collection of styles. Rather than a novel way of rearranging surfaces and spaces, it is fundamentally concerned with the implementation of the art impulse in an advanced technological society.

Burnham's observations and conclusions resonate with me. These constructed environments are, in a sense, maquettes of the real world, distorted and inverted in some cases, but always surrounding us with sights and sounds that invoke our attention and despite being large in some instances, are within our artistic control. Despite its military and industrial genesis, and perhaps even because of it, Burnham's thesis is as relevant today as it was in 1968. Caroline A. Jones (Jones 2023) writes in ARTFORUM magazine:

The systems virus is still here—but it has mutated from Burnham’s engineering and industrial setups into the social and electronic protocols that govern our world.

If we examine the “materials” description of a modern AI driven work, *Unsupervised - MOMA* by Anadol for example, we see software like VVVV and GANs, hardware in terms of dozens of visual and audio channels, credits for specialists including software developers, cloud data management analysts, mathematicians and so on. These works are significant systems.

Inhabiting the Invisible is composed of multiple software frameworks, computer languages and sophisticated AI models for detecting faces and spoken languages. This is implemented on a computer with a powerful graphics engine, a camera, microphone and two projectors.

These choices, as described by Burnham and Jones, are made for social, technical, and aesthetic reasons. The AI models are made from human faces and voices. I chose these elements to contribute to the ideas of AI and agency as defined by state of the art in 2015. The AI models conform to technical standards and protocols (e.g. CUDA) in order to be used by programmers on typical computers. The Unity Game Engine provides a framework for visualizing objects with the lifelike characteristics expected of modern electronic games. These elements have been customized through code to present an installation based on an aesthetic derived from all of these components.

Physical Environment

In a typical art gallery experience, an image or some other central focus frames or centers the visitor’s viewpoint, and often the image is made to have its own geometric focus that places the viewer's head to a very specific place and angle. (Bishop 2005)

This is similar to our experience in the digital realm. Both the physical requirement of sitting in front of a computer screen, and the experience of visiting the highly curated digital spaces like Facebook and Google are constrained and controlled.

In contrast, an immersive installation requires that the visitor physically move and look from many perspectives to experience the work. It gives the viewer some ability to control the experience and perhaps discover more about the work. This will offer viewers of *Inhabiting the Invisible* some, perhaps only illusory, agency in their experience of the digital realm.

This also invites the artist to include elements in an immersive installation that have control over how the audience can participate. Nauman in his *Hallway* installations went so far as to reconfigure the walls and lighting of the exhibition space to control the experience of the participants and to evoke feelings in them that he wished to convey.(Gryczkowska 2011)

In this work I use some elements of the physical installation to signal to the viewer that even if we see the technology, i.e. the laptop, projectors, microphone and camera, we don't necessarily understand it. The digital realm tends to hide its inner workings behind slick interfaces; here we are trying to expose what we can.

The Digital Realm

For the purposes of this document the digital realm is defined to be the two largest internet platforms, Google and Facebook. Today both are used regularly by about half of the population of the earth. (Gajic 2025) This is important because while “the digital realm” sounds ambiguous, even for this narrow definition, it is a reality for most people.

The term “Digital Realm” is used in this project as a metaphor for these two platforms. This document will focus on these two immense platforms in order to make concrete connections between the metaphor of inhabiting, and the much more large and diffuse “real world” digital realm.

Interaction

Pioneering interactive artist David Rokeby is interested in the nature of interaction, how interaction transforms the relationship between the interactor and the piece, and perhaps the sense of reality experienced by the interactor.(Rokeby 1998)

In “The construction of experience: interface as content” he proposes that “Interactive systems inherently involve feedback.”. Facebook and Google provide feedback on a certain level, but the main autonomous elements, that is the collection and processing of information about the user (Statista 2026), do not feed back to their public facing interfaces. Rokeby references a video disc-based game where the player might believe that they are in control, flying at a tremendous speed when actually the control software is tricking the player and not really reflecting their control inputs on the screen. There is also an interactive interface layer of illusion, search or socialization, placed over the reality of the digital realm's platforms' main purposes which are related to revenue (Statista 2026).

Inhabiting the Invisible has adopted this approach by acquiring the participant's image and utterances for processing and later presentation but does not offer the participants any ability to alter them. This is an echo of the digital realm's model of acquisition, processing and sale of a user's data so the duration of this interaction loop is very long.

Rokeby's classification of interactive systems (Rokeby 1995) is relevant here. The classification does not explicitly discuss agency, but it is an analysis of where agency might lie by exploring what role the work might take in the interactive interplay. He discusses maps/wayfinding, mirrors, automata all as classifications of interaction. These interactive models describe a direct, reciprocal engagement between the participant and the work. *Inhabiting the Invisible* implements a slow, insidious form of interaction that appears to be an automata in Rokeby's classifications. All of Rokeby's interaction models describe a direct and often immediate connection between the participant and the expression of almost any sound or image by the work. This implies that the work directly addresses and involves the participant. *Inhabiting the Invisible* examines the elements of the digital realm that do not interact in an immediate way with the participant but slowly ingest and digest and alter themselves with the participant's imagery and data.

Methodology and Methods

Creation as Research

This thesis uses the creation-as-research process described by Owen Chapman and Kim Sawchuk (Chapman and Sawchuck 2012) pg15:

It is a form of directed exploration through creative processes that includes experimentation, but also analysis, critique, and a profound engagement with theory and questions of method.

This is a qualitative, subjective approach to the project's research questions.

Methods

Iterative Prototyping

Initiation

The process began with a first guess at the interfaces and actions of the work. I mapped this to an achievable technical design based on proof of concept and prototypes that I have built or have had experience with. These initial technical elements were refined to some acceptable level then the creation/investigation/critical review cycle began.

An application to the OCADU Research Ethics Board (REB) was made to establish the levels of permission and informed consent required to use people's faces as part of the work. The result was that the REB mandate does not include creative works so no requirements were established.

The project advisors and I decided that we would restrict the project to ensure that the ethical problems we could identify were minimized. No permanent record of participants' images or derived data were kept. No connection to AI services were used to avoid exposing data or imagery to these services - all processing is done on the project computer. No images or data are shared.

Creative Cycle

Each pass through the creative cycle exposed problems in the technical framework that were addressed in later cycles, or the deficiency informed the scope of the project.

The questions of understanding and agency will be visited in each cycle. How has the technical framework of the work afforded an opportunity to express what I feel about the digital realm?

The creative aspect of designing the interactions, the portrait displays, interactions and the soundscape, were documented and critically reviewed as creation research results.

Where possible, the emergences were examined in context with related theory and artworks.

The assessments became input for the subsequent cycles. Unanticipated and interesting opportunities for creative expression of the themes were documented.

Results

Each of the parallelograms in the diagram below represents outputs of the process. The questions of method are the main focus: how do we use AI tools to express the goals of the project?

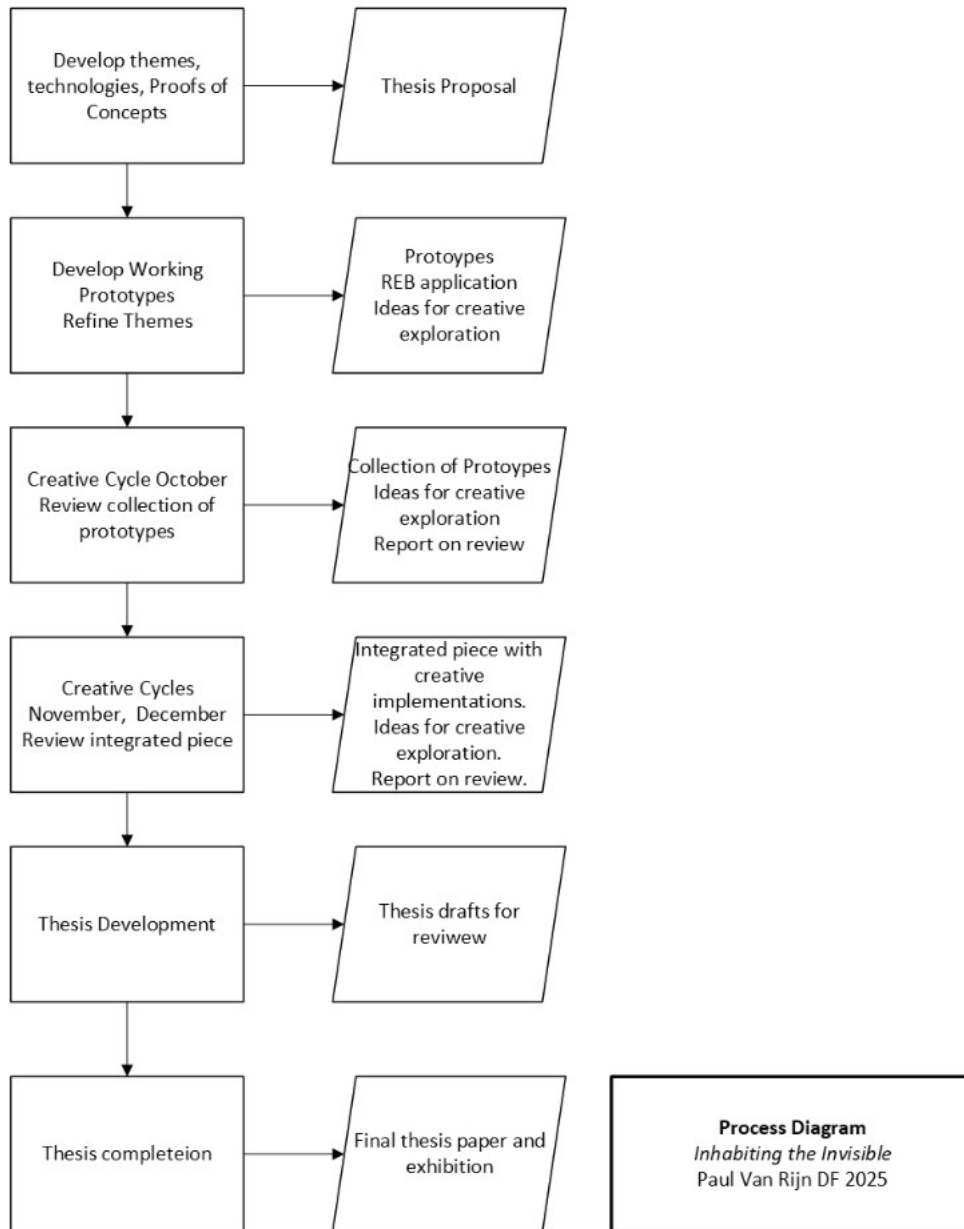


Figure 2: Process and outputs diagram

Prototypes

Gathering Concepts, Experiences, Influences

System Element: Portrait acquisition, processing and display

The intention of this prototype was to demonstrate some AI and visualization tools that might be important for the main project. Using a combination of AI (mediapipe, deepface) detect, isolate, and treat (using python, openCV) human faces to create "portraits".

The AI analysis derived the location and angle of the face, and other estimations like age, gender, emotion etc. I worked through these derivations thinking about how they might be important to the themes of the project. There were many expressive avenues presented by these derived values.

I created a script that used the location information to frame the faces and scale them to be a standard size.

The backgrounds were removed; the colours were balanced and then the AI derived values were used to assess if the portrait was "acceptable".

Emergence

It became clear that the AI components used in this project are prone to error. I used the AI drawing tools associated with the face detection tools to place the detection points (eyes, nose, ears) on the image that they were derived from, and the points were far from the related parts of my face. This might have been my fault in some way but in the second image we see a similar displacement in my age and emotional state. This will have an impact on how I can use these results in other code.

Many people who watched the demonstration commented on surveillance or invasion of privacy. While this is obvious, it is something that I wish to avoid because these are just two of the many issues concerning these technologies.

This exercise demonstrates camera based acquisition, analysis and interaction as the extension of a complex system (the neural networks of the AI tools that have been loaded with millions of faces and gestures) outside of itself into the physical realm of the audience.



Figure 3: Face detection and information derived from an image by the deepFace AI showing dislocated facial points.

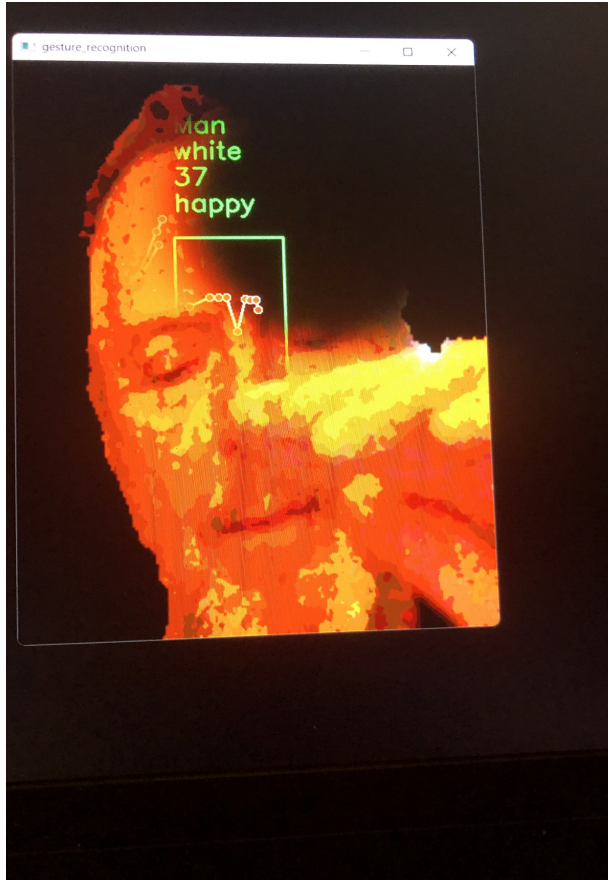


Figure 4: Incorrect age and emotional estimates derived from the image by deepFace.

Context

David Rokeby (Rokeby 1984) and Danny Rozin (Budow 1999) created interfaces using custom hardware, software, cameras and other specialized equipment as interactive interfaces to their works. These were groundbreaking for their time and represent a considerable investment in experimentation and development. This paradigm of camera to computation to expressive elements of a work is a key system design element of *Inhabiting the Invisible*.

It is also part of the design of Facebook, the organization that developed facial detection and recognition software in 2014 called deepFace using AI models that afford straightforward access to many properties of faces/people/gestures in the field of view. Kate Crawford and Trevor Paglen use what appears to be the same deepFace software in *Training Humans* as did Rafael Lozano-Hemmer in his work *Recorded Assembly*. *Inhabiting the Invisible* uses deepFace as a core component as part of the systems approach.

System Element: Interaction and Agency



Figure 5: Blue Gaze bird with programmed behaviours related to AI derived data

As part of a group project for OCADU (Pasian 2024) the other members and I implemented a bird that observed and reacted to people in its field of view. The motivation was to connect the external world to the internal world of the bird through a camera and software. We made a small behaviour model for the bird that took presence and proximity as detected by [P5.js](#) and reflected that in how the bird responded through “dancing”.

The system inside the bird is conferred agency through the scripts via the camera input and the AI analysis of the images from the camera.

The bird’s responses are not based on a real bird’s behaviour and are not necessarily an anthropomorphic representation, but rather a unique personality with set states in response to the AI estimates.

Emergences

Based on anecdotal comments and observations, the bird's physical appearance was very engaging to viewers. This was an important aspect of the work because it attracted viewers to the work and gave them some incentive to examine it further.

The bird's behavioural response was unexpected by the audience in that viewers would try to encourage the bird to dance by performing the bird dance themselves. We saw that it took a conscious effort for people to conform their behaviour to the bird's needs as described in the didactic posted next to it. Despite this effort, there was an interesting, if indirect, interaction between the viewers and the work where they would alter their behaviour in front of the bird and wait a few seconds for the response.

System Element: Presentation

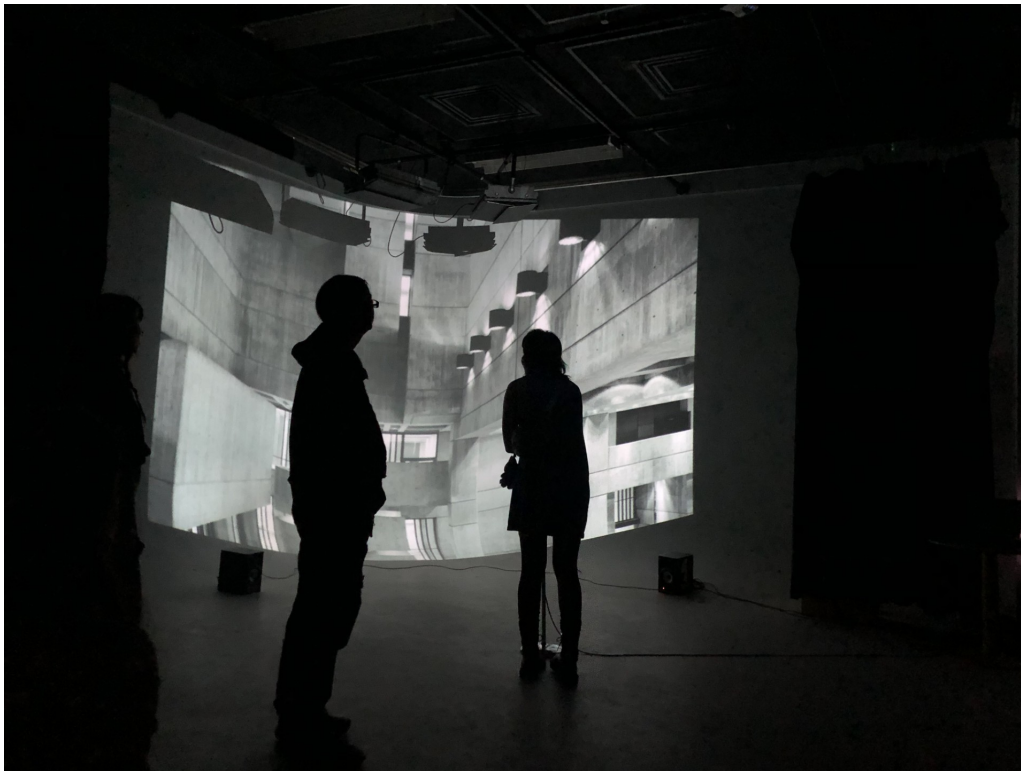


Figure 6: Your Voice, (2025) a large-scale projection prototype project

Your Voice was a project of mine in 2025. It is a series of black and white photos of the interior of Brutalist buildings projected onto a curved wall. A microphone was placed near the screen - the microphone invited the audience to speak into it, but when they did, their voices were scrambled by layers of reverberation and echo to the point of unintelligibility.

As the conceptual elements of inhabitation and ubiquity evolved through the iterative creative process, implementing the display for *Inhabiting the Invisible* at a similar large scale became appropriate to me. This scale, taller than a person, gave the sense of enclosure that encompassed the audience. The large scale of the faces echoed the large scale of the digital realm.

Emergence

General comments from the *Your Voice* project included a feeling of actually being in the buildings on the screen (inhabitation), awe at the scale of the images, and a sense of coldness at the stark interiors in the projected images. I felt that the scale and physically enclosing nature of the projection was as impactful as the subject and properties of the photographs themselves.

The invocation of these feelings of awe, and the transfer of the feelings of coldness and emptiness implied by subjects in the photos via the scale and the encompassing screen are important to the physical and visual designs of *Inhabiting the Invisible*.

For the purpose of this work, I have explored some definitions of emotions/affects as classified by Robert Plutchik. (Edine 2005) Milton Glaser, in his The Big Think video interview about art, discusses the importance of attention in art. This attention is one of the affects described by Pluchkin. Other affects that this project investigates, such as unease, questioning, surprise, and affinity are also described. At a deeper level these affects circle back to motivations, agency and behaviour in individuals but are outside the scope of this project because of the scale of that problem.

Context

Rafael Lozano-Hemmer uses large scale projection in many works as an element of his concept of “relational architecture” that he describes as “anti-monuments for public dissimulation”. This suggests that he is disguising the building in the projection. (Fernández 2007) These monumental projections often include images or other reflections of the audience, offering them agency in the visual experience of the cityscape. *Inhabiting the Invisible* inverts this monumental scale to the invisible yet “giant” infrastructure of the digital realm. (Hardy 2015) Instead of being projected outside on to “bricks and mortar” our faces are projected inwards to the immense, energy hungry data centres all over the world.

Integration of Prototypes

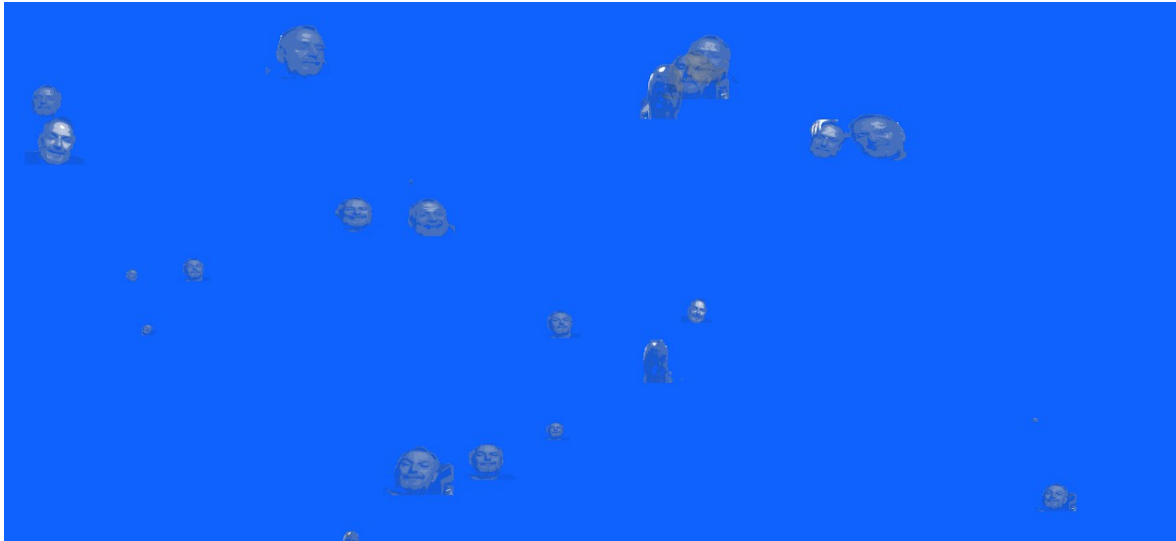


Figure 7: Uninteresting Display

I assembled the software components into a single python project that ran on a laptop. The AI components worked after the image/portrait generation was rebuilt.

I tried a few forms for the portrait projection/display and the degraded, translucent, grey scale images worked best as a first guess because they appeared to be diaphanous, mysterious yet recognizable.

Emergence

The result was not interesting or engaging. A strong signal arose - why did I immediately believe that the display was very unsatisfactory?

Subtleties of scale, motion, implied location in the space of the simulation are somehow a requirement of our perception and engagement. I would like to use the term “realistic” but nothing about the final visualization is realistic.

My conclusion was that I needed a better tool set to implement the simulation and visualization components so that I could take advantage of things like motion, energy transfer, distance and depth calculations and whatever else might be available. After canvassing my classmates, some of whom are accomplished game designers and builders, I decided to move the model/simulation elements of the work into the Unity Game Engine. There I discovered a broad range of tools that are used in games to offer me the ability to better control visual aspects of the display important to the affect that I was trying to achieve. The motion of the portraits could be controlled using energy inputs and reactions that appeared

to be more natural. The scale and location in the 3d space made it possible for me to create motion and interaction behaviours that I felt were more attractive.

Context

In his essay *The Uncanny Valley [From the Field]* Masahiro Mori focuses mainly on robots and the human form, describing how we feel revulsion rather than affinity to representations that are close to lifelike, but fail in some respect. (Mori 2012) On one side are cartoon or human-like shapes that we “like” and closer to human representations he contends that we are more profoundly repulsed unless the representations are accurate. We might think of movie zombies as an example. He writes that motion amplifies this effect, meaning that human-like motions are repulsive unless they are very accurately human-like.

Based on the revulsion/affinity ideas of Mori, I continually reworked the visualization and motion aspects of the portraits in the 3-dimensional space of the simulation until I felt that they were more attractive than repulsive. My revulsion was mild, more a dislike, but I had to concentrate on my feelings to find where the dislike turned around and I felt some affinity for the disembodied portraits.

SPARK Presentation

SPARK is an event where works in progress are semi-formally presented and discussed within the Digital Futures cohort.

This was a technical challenge in that the active elements of the project now needed to be distributed among several disparate technologies.

The portrait acquisition and image pre-processing was put into a separate python project that also included gesture detection and data management. These elements were set up in a kiosk model where people sat in a chair, saw what the system had captured as a portrait, then approved the images with a thumbs up gesture.

The model/simulation for the individual faces, the model for 3d rendering and the projection/visualization were moved to Unity.

Despite my intention to make this a strictly technical iteration, many conceptual elements presented themselves.

The distribution of components of the work among many technical processes reinforced the “systems” approach as described by Jack Burnham. It took a concerted systems design effort to get it to work.

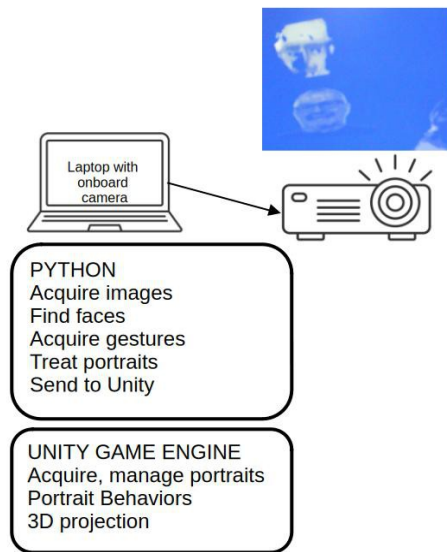


Figure 8: Overview of SPARK test presentation implementation

There was a set of instructions posted on the desk with the laptop/camera that described the system and indicated that participants could induce it to take a picture of them by making a “thumbs up” gesture. Based on my anecdotal observations, self-presentation, that is making faces, was more interesting to viewers than any risk demonstrated by an AI based system capturing, manipulating and displaying images of them.

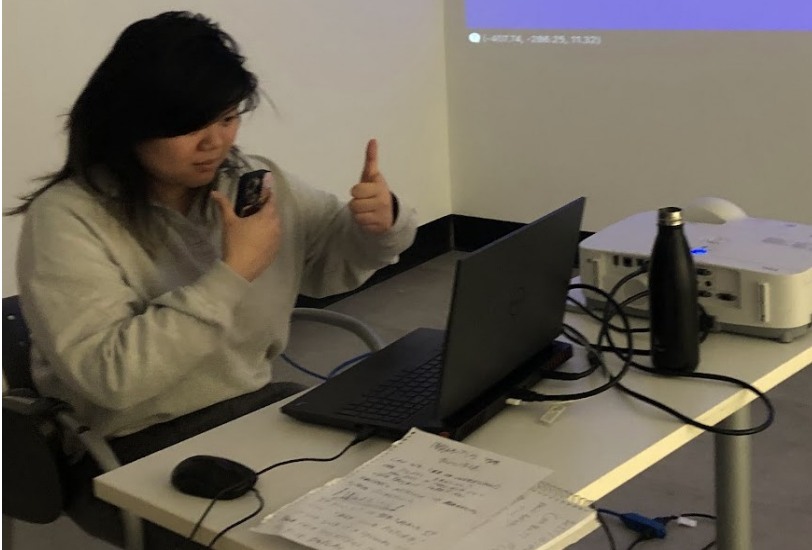


Figure 9: User testing gesture-based image acquisition for portraits



Figure 10: A participant's face at a large scale. This also demonstrates the "mugging" behaviour.

Emergences

Self-Presentation vs Self-Preservation

I observed in the demonstration that people enjoyed the process of having their portrait taken. They enjoyed seeing their own faces and faces of people that they know in the projection. This is nothing new, but it was remarkably powerful.

Interaction

My intention was to gauge the technical robustness of the system and to test if it was capturing and processing imagery correctly. My informal impression was that after the initial engagement with the image through recognizing themselves or their friends, there was little or no contemplation of the display of the disembodied faces. The capture process was a significant distraction from my intention that the viewers examine what the system had done to their images.

Time

Time emerged as an important factor in all elements of this work.

Thematically, faces move through the digital realm more slowly than typical immediate interaction is "supposed" to be. Images tended to accumulate faster than the system cleared them, so the displayed images were sometimes crowded with faces and sometimes quite sparse.

My intention was that there be a "surprise" factor, a resolution of sorts, when people finally see themselves represented there. That is a valuable "reward" for viewers who remain to contemplate the work. I didn't test this hypothesis.

The persistence of a face in the gallery is also important. I decided to start the installation without any stored faces and allow them to accumulate to a threshold that makes an interesting display and then delete old portraits as new ones arrive.

Context

Rokeby strove to make timing invisible in *The Very Nervous System*. In the article "The Construction of Experience: Interface as Content" in the "tricks of the trade" section he points out that the apparent interface, with implications of freedom and agency, overlays the reality that the options for system response are limited by design and technology. In the digital realm the interfaces that we associate with search or social interaction represent a skin that overlays the actual motivation, revenue per user, of the entire system. This lower level does not represent itself to the users. It is interactive only in a cycle that is unrelated to the perception on the "user interface" level.

Inhabiting the Invisible generates an imaginary world at that unseen level with a slow interaction cycle and a specifically engineered disassociation between the participant and the information that represents them.

Implementation and Experimentation

Over the months of October and November 2025 I implemented several experiments to determine if they would contribute to the outcomes described above.

Gesture Recognition

I added functionality to recognize some hand gestures that invoked a reaction where the faces scattered in response to an outheld palm. This interaction was unsatisfying because it felt to me as though I was influencing the interior behaviour of the simulated virtual realm. This was anathema to the theme of lost agency.

Kiosk

The gesture recognition element was replaced with a process where the participants enter the exhibit through a small kiosk. I tested a few different configurations for the kiosk.

The camera is on a tripod in plain sight. The kiosk is lit with a small spotlight reflecting off a wall with a poster on it. There is also a microphone on a stand. Eventually it was enclosed in movable partitions to isolate the viewers as they pass through into the display.

The spotlight adds to the feeling of control and process. It also serves the practical purpose of controlling the lighting so that the image acquisition software can make consistent decisions about which images are acceptable.

The system is continuously observing the video and audio via the kiosk. Based on timing and the AI estimation of the emotional state of the viewer, an image is selected, turned into a portrait and injected into the pool of faces.



Figure 11: Testing elements of the kiosk live with portrait projection.

The voice system continuously converts input from the microphone to text and back into audio. These tests showed problems. More tests showed that the voices played back in different timbres was, like the interactive gestures above, a distraction. The voices appeared humorous to me, but not evocative in any other way. The sound system evokes a disquiet related to the unexpected repetition of words and phrases that had been uttered moments ago.

Face Behaviour

The processed face images move around in a virtual 3-dimensional space on slowly changing random vectors. They “bounce” off the edges of this space outside the virtual camera range and eventually return to view.

I implemented a flocking behaviour to provide some form of relationship between the faces. The faces huddled together and became less interesting, so I removed the behaviour.

After operating for a few minutes, the faces became dispersed and the presentation became relatively empty. I re-implemented the “collision avoidance” element of the flocking algorithm and after a bit of tuning the faces appeared to behave better. However, this became predictable and therefore not engaging for viewing over more than a few minutes. I implemented a randomly timed mixing of the faces that helped make the display more engaging over a longer period.



Figure 12: Test to see face motions at scale.

Projection

Ideally the projections would be impressively large to give some feeling for the scale of the digital realm. The location of the exhibition and the geometry of the projectors determine the maximum size that the projected images can be made.

Screens made of fabric looked very good in both reflection (side facing the projector) and transmission (side away from the projector) but the structure required to support them and keep them taut was distracting. I tested a screen made from photographic background paper and, because it required much less support and infrastructure, it seemed to disappear under the light of the projectors. This approach was abandoned because of limitations imposed by the final exhibition site.

The projectors were placed in the corners and above head height. The screens were accessible all around and viewers could walk between them. The intention is for the displays to be overwhelming and, at about a meter apart, they felt uncomfortably large and close.

All the electronics, supports, wiring and even the computer running the software will be visible to the participants.

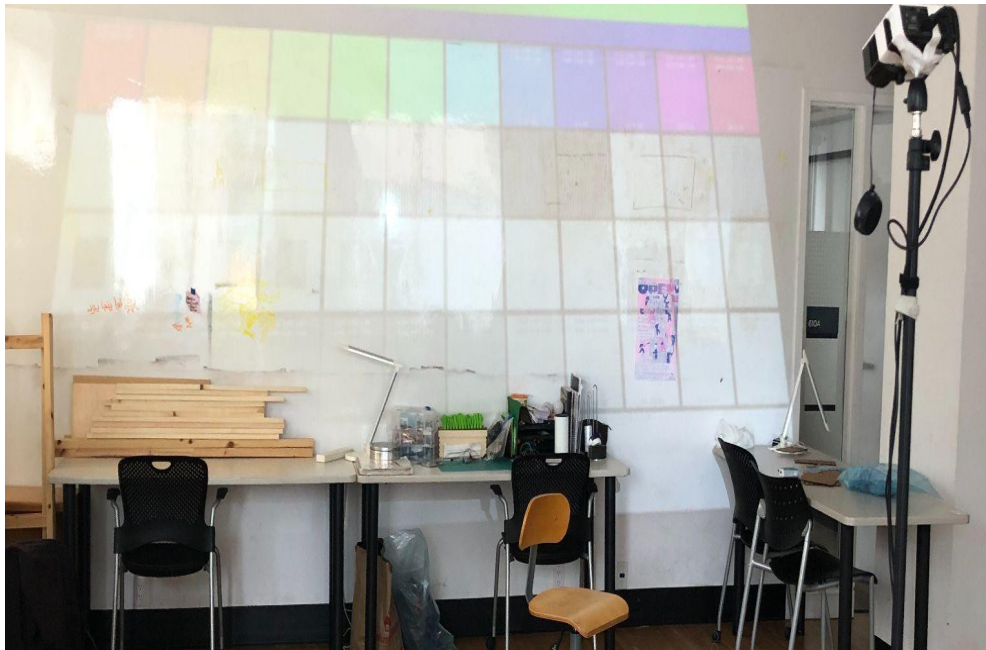


Figure 13: Projection over distance tests to determine how large the screens might be based on the size of the exhibition space and the geometry of the projectors.

Emergences

Every attempt at adding real time interaction with the portraits led to an unsatisfactory result because my attention skewed from engagement about the overall impression to a focus on the effects of the interaction.

Face behaviour required a balance between preserving the disembodied and flattened impression of the portraits with motion that is believable. It took many tries to define and tune a motion algorithm for the portraits that is engaging and still expresses the strangeness of the transformed portraits. Eventually it appeared that they became some other kind of creature. Cephalopods perhaps.

Large scale projections of the faces were more interesting than mid-sized faces. I moved the limits of the “bouncing” inside the virtual camera’s frustum so that faces moved closer to the camera and appeared to be larger.

Software Design

Portrait processing starts with python/openCV acquiring images from the video stream from the camera set up in the kiosk. These are passed to the Google mediapipe library to detect faces and the faces are cropped out. The resulting face images are assessed by the deepFace library for emotions and then are quality checked. If the emotions and other aspects like timing are acceptable the images are decimated, decolourized and made into portrait files.

The portraits are passed as files to the Unity Game Engine where they are turned into graphics objects and are put into the game space. There they acquire some behaviours coded in C# related to motion and reaction to each other and the edges of the 3d space. Two virtual cameras are set up to send their point of view to the projectors.

The sound processing starts with an implementation of the openAI *whisper* python library that monitors a microphone set up in the kiosk. The resulting series of text strings are quality checked and then made available for the text to voice and sound modification libraries. The texts are randomly associated with a voice/emphasis file and audio is sent out to the sound system.

The code is available on <https://github.com/paulvanrijn/inhabit>

Conclusion

This project demonstrates that contemporary AI systems can operate through forms of delayed, opaque interaction and subtle forms of coercion that limit user agency while maintaining the illusion of participation. The aesthetic/attractive appeal of the viewers’ faces moderated their perception of the power asymmetry between the viewer and the installation.

By translating these conditions into an immersive installation, this thesis offers a way to critically engage with the experiential and affective dimensions of the digital realm.

It demonstrated that design through scale, colours, behaviours, and control of the physical environment of the work played a key role in making these invisible systems perceptible and open to reflection.

Exhibition Commentary and Documentation

Please see Appendix B for details about the exhibition.

Work Cited

Anadol, Rafik. "Machine Hallucinations: Unsupervised." November 2021.

<https://refikanadol.com/works/unsupervised/>

Azizi, Othman. "The Rise of Deep Learning: Why 2015 Was a Pivotal Year." 2015.

https://www.researchgate.net/publication/385273558_The_Rise_of_Deep_Learning_Why_2015_Was_a_Pivotal_Year

Bahar, Varqa Shamsi. "Self-Presentation Theory: A Review." In TheoryHub, edited by S. Papagiannidis, 2024. <https://durham-repository.worktribe.com/output/3342035/self-presentation-theory-a-review>

Bandara, Pesala. "Band Uses Live Facial Recognition on Audience in Powerful Critique of Technology." PetaPixel, September 17, 2025. <https://petapixel.com/2025/09/17/band-massive-attack-uses-live-facial-recognition-on-audience-in-powerful-critique-of-technology/>

BBC News Technology. "Internet Used by 3.2 Billion People in 2015." BBC News, May 26, 2015.

<https://www.bbc.com/news/technology-32884867>

Big Think. "What Is Art and What Is Not Art." Video, April 23, 2012. <https://www.youtube.com/watch?v=LR AptNtN9-A>

Bishop, Claire. Installation Art. London: Tate Publishing, 2005.

Budow, Steve. "Wooden Expression." Wired, December 1, 1999.

<https://www.wired.com/1999/12/wooden-expression/>

Burnham, Jack. "Systems Esthetics." Artforum (1968).

https://monoskop.org/images/0/03/Burnham_Jack_1968_Systems_Esthetics_Artforum.pdf

Chapman, Owen, and Kim Sawchuk. "Research-Creation: Intervention, Analysis, and 'Family Resemblances.'" Concordia University, 2012.

<https://www.concordia.ca/content/dam/artsci/cissc/docs/Owen%20Chapman%20%20Kim%20Sawchuk%20Research-Creation-%20Intervention%20Analysis%20and%20'Family%20Resemblances'.pdf>

Chowdhry, Amit. “Match Faces with 97.25% Accuracy.” Forbes, March 18, 2014. <https://www.forbes.com/sites/amitchowdhry/2014/03/18/facebooks-deepface-software-can-match-faces-with-97-25-accuracy/>

Crawford, Kate, and Trevor Paglen. Training Humans. 2020. <https://www.fondazioneprada.org/project/training-humans/?lang=en>

Ednie, Kathryn J. “Emotions and Life: Perspectives for Psychology, Biology, and Evolution.” American Journal of Psychiatry 162, no. 2 (2005). <https://psychiatryonline.org/doi/10.1176/appi.ajp.162.2.409>

Fernández, Maria. “Illuminating Embodiment: Rafael Lozano-Hemmer’s Relational Architectures.” Architectural Design 77, no. 4 (July 2007). https://www.lozano-hemmer.com/texts/bibliography/articles_panorama/31_ArchitecturalDesign.pdf

Gajic, Ana. “25 Google Search Statistics and Trends.” 99firms, August 10, 2025. <https://99firms.com/research/google-search-statistics/>

Gryczkowska, Agnieszka. “Changing Light Corridor with Rooms 1971, Bruce Nauman.” National Galleries, 2011. <https://www.nationalgalleries.org/art-and-artists/89096>

Hardy, Quentin. “Facebook Takes Its Hardware Designs to a Giant Scale.” New York Times, February 11, 2015. <https://archive.nytimes.com/bits.blogs.nytimes.com/2015/02/11/facebook-takes-its-hardware-designs-to-a-giant-scale/>

Ji, Haru Hyunkyung, and Graham Wakefield. “We Are Entanglement.” In Proceedings of SIGGRAPH, August 2025. <https://dl.acm.org/doi/10.1145/3721249.3731635>

Jones, Caroline A. “Caroline A. Jones on Jack Burnham’s ‘Systems Esthetics.’” Artforum, 2023. <https://www.artforum.com/columns/caroline-a-jones-on-jack-burnhams-systems-esthetics-200869/>

Lozano-Hemmer, Rafael. “Recurrent Kafka.” 2025. https://www.lozano-hemmer.com/recurrent_kafka.php

Marcus, Gary. “What Comes After the Turing Test?” New Yorker, 2014. <https://www.newyorker.com/tech/annals-of-technology/what-comes-after-the-turing-test>

Metz, Rachel. “Facebook Is Shutting Down Its Facial Recognition Software.” CNN Business, November 8, 2021. <https://www.cnn.com/2021/11/02/tech/facebook-shuts-down-facial-recognition/index.html>

Mori, Masahiro, Karl F. MacDorman, and Norri Kageki. “The Uncanny Valley.” IEEE Robotics & Automation Magazine 19, no. 2 (June 2012): 98–100. <https://ieeexplore.ieee.org/document/6213238>

Olin, Margaret. "Forms of Respect: Alois Riegl's Concept of Attentiveness." *Art Bulletin* 71, no. 2 (1989): 285–299. <https://www.collegeart.org/pdf/artbulletin/Art%20Bulletin%20Vol%2071%20No%202%20Olin.pdf>

Olivia Pasian, Paul Van Rijn, Kasper Zhang. "Blue Gaze" OCADU Creation & Computation 2024 web site. <https://blogs.ocaduwebpace.ca/digf-6037-301-2024fa/2024/11/12/blue-gaze/>

Radford, Alec, et al. "Robust Speech Recognition via Large-Scale Weak Supervision." arXiv preprint, 2022. <https://cdn.openai.com/papers/whisper.pdf>

Ridler, Anna. Mosaic Virus. 2019. <https://annaridler.com/mosaic-virus>

Ridler, Anna. "Mosaic Virus." Video, 2020. <https://youtu.be/TGEDxqMyuGk>

Rokeby, David. "Body Language." Accessed March 18, 2026. <http://www.davidrokeby.com/body.html>

Rokeby, David. "The Construction of Experience." In *Digital Illusion: Entertaining the Future with High Technology*, 27–34. New York: ACM Press, 1998. <http://www.davidrokeby.com/experience.html>

Rokeby, David. "Transforming Mirrors: Subjectivity and Control in Interactive Media." In *Critical Issues in Electronic Media*, edited by Simon Penny. Albany: State University of New York Press, 1995. <http://www.davidrokeby.com/mirrors.html>

Schlosser, Markus. "Agency." *The Stanford Encyclopedia of Philosophy*. Winter 2019 edition. <https://plato.stanford.edu/archives/win2019/entries/agency/>

Scott, Felicity. "Limits of Control: Rain Room and Immersive Environments." *Artforum*, September 2013. <https://www.artforum.com/features/limits-of-control-rain-room-and-immersive-environments-217722/>

Statista. "Average Revenue per User (ARPU) of Meta Platforms from 2011 to 2025." 2026. <https://www.statista.com/statistics/234056/facebooks-average-advertising-revenue-per-user/>

Su, Lixu, Annie Cui, and Michael Walsh. "Trustworthy Blue or Untrustworthy Red: The Influence of Colors on Trust." *Journal of Marketing Theory and Practice* 27, no. 3 (2019): 269–281. https://www.researchgate.net/publication/334550253_Trustworthy_Blue_or_Untrustworthy_Red_The_Influence_of_Colors_on_Trust

Transmediale. "The Helpless Robot." 2006. <https://archive.transmediale.de/content/the-helpless-robot>

Uher, Jana. "What Is Behaviour? And (When) Is Language Behaviour?" *Journal for the Theory of Social Behaviour* 46, no. 4 (2016). <https://onlinelibrary.wiley.com/doi/10.1111/jtsb.12104>

Wu, Tim. "Has Britain Become an Economic Colony?" Guardian, November 23, 2025.

<https://www.theguardian.com/books/2025/nov/23/has-britain-become-an-economic-colony>

Appendix A: Software and AI References

The following is a quick overview of some AI techniques and tools discussed in this document and used in the project.

Deep Neural Networks

A machine learning, or deep neural network model is used by Google to rank pages based on a client's search text. It consists of a dataset, billions of web pages in this case, and a neural network that has been trained to "give the right answer" based on the query input by platform users. Its role is to convert the often ambiguous and incomplete search requests into a ranked list of pages - it is a meaning detector rather than a word search engine. It is trained by a large contingent of paid search quality raters (about 16,000 at this time) (Google 2025) and the staff at Google.

Generative Adversarial Network (GAN)

A GAN is an AI model made of three main parts: the dataset, which can be immense as in Anadol (all of the art in the catalogs of MOMA), or modest in the case of Ridler (approximately 1000 images of tulips), the generator which creates things based on a neural network trained on the dataset and a discriminator which accepts or rejects the outputs of the generator. The discriminator is initially trained on real data, imagery in the dataset for example, and fake data, created using the generator on random noise for example.

In terms of agency, an artist might choose the dataset, the training protocol for both the generator and the discriminator and can actively guide both of them. (Ridler 2020)

This is an example of a technique that can generate a continuous stream of new and different outputs based on the data, training and constraints as guided by the creator of the system.

Anadol and his team add layers upon the emerging data points where further discrimination by the artist can be applied. For example, the rate of change (difference over time) of a particular model output can be converted to a visualized quantity rather than simply the values appearing in the output. (Anadol 2021).

AI Tools Used in the Project

The same basic deep neural network is used in the face detection, gesture recognition, and face recognition applications used in many of the artworks referred to in this document.

Deepface

Deepface, the face detection and recognition system created by Facebook, was trained on a database of faces and related data, drawn, without permission, from a million or so of their users' data starting in 2014 until 2021. (Chowdhry 2014), (Metz 2021) Deepface is the main AI tool used in this project because it is about, and is made from, our faces, and our faces are the main interactive artifact used in this project.

It is used in the face detection and emotion estimation for the creation of portraits.

Mediapipe

Mediapipe is a newer AI toolset created by Google AI as a uniform interface to video, audio and text AI and machine learning models. The interface works with multiple platforms and programming languages and therefore presents a standardized, though proprietary, way to implement many AI models. It is used in the image manipulation portion of this project for the creation of portraits.

Kyutai Pocket TTS

This library generates audio output from a string of text and a “donor” audio file that guides the AI to create the spoken text in the donor’s voice and emphasis. This was used in the testing phases but was found to be too distracting.

Torch

This is a generalized AI framework used to ensure that the standardized neural networks used by most of the AI tools in *Inhabiting the Invisible* run on the graphics accelerator rather than the CPU.

Whisper

Whisper is a speech to text system based on neural networks that were trained by openAI on 680,000 hours of audio and related transcripts found on the internet (Radford 2022).

Other Software Tools

python

python, mostly version 3.12 in virtual environments, was used to access the libraries above. The python libraries pygame, pedalboard and numpy were also used.

OpenCV

Since 2000 a consortium has managed OpenCV to build a large library of computer vision algorithms. OpenCV is used in this project for generic camera and image processing tasks like acquisition from the camera, image manipulation like scaling, rotation, colour manipulation etc.

pyttsx3

pyttsx3.py was used to generate voice from texts. It is an old library with a robotic sound that suits the old time feel of the project.

Unity Game Engine

The Unity Game Engine, initially created in 2005, was used to run the 3d simulation and generate a point of view for each of the two projectors. The behaviours of the portraits, e.g. motion, scale, interactions with each other and the virtual world, were coded in C# scripts for Unity.

Appendix B Final Installation

Physical Design

This is an overview of the physical design at the time of the exhibition.

Location

The presentation was in room WFA 435 at 130 Queens Quay E

The room is approximately 22x29 feet.

Overview

The installation was two screens in the centre of the room that formed a corridor. Two projectors were mounted on stands in the corners of the room driven by a laptop on a desk beside the kiosk.

A kiosk was built at the eastern side door where viewers could read the didactic, an advisory note and where they were photographed via the web cam and a microphone transcribed their voices from voice to text.

Viewers walked around the outside of the room or walked inside the corridor between the screens. All the equipment, camera, stands, projectors, laptop, etc. were visible to them.

The Screens

The screens were made of wood with scrim fabric stretched over them, mounted on floor stands made of the same wood as the screens. The projection surfaces were angled to give the impression of being an enclosure.

The Projectors

The projectors were short throw 1080p models that defined the geometry of the installation. They were placed on top of stands with the appropriate ball attachments so that they accurately aimed at the screens. They were 8 to 10 feet high to reduce glare in the eyes of viewers.

The Kiosk

When viewers entered the exhibit, they passed through a small enclosure made of movable wall panels. A camera and microphone were situated to record them as they look at a poster on the wall. A light illuminated the poster, and the reflected light ensured that the viewers are appropriately lit for the camera and AI software to generate a consistent result.



Figure 14: Entry kiosk with video camera and microphone.



Figure 15: Desk with computer and a back view of the kiosk made from movable panels.



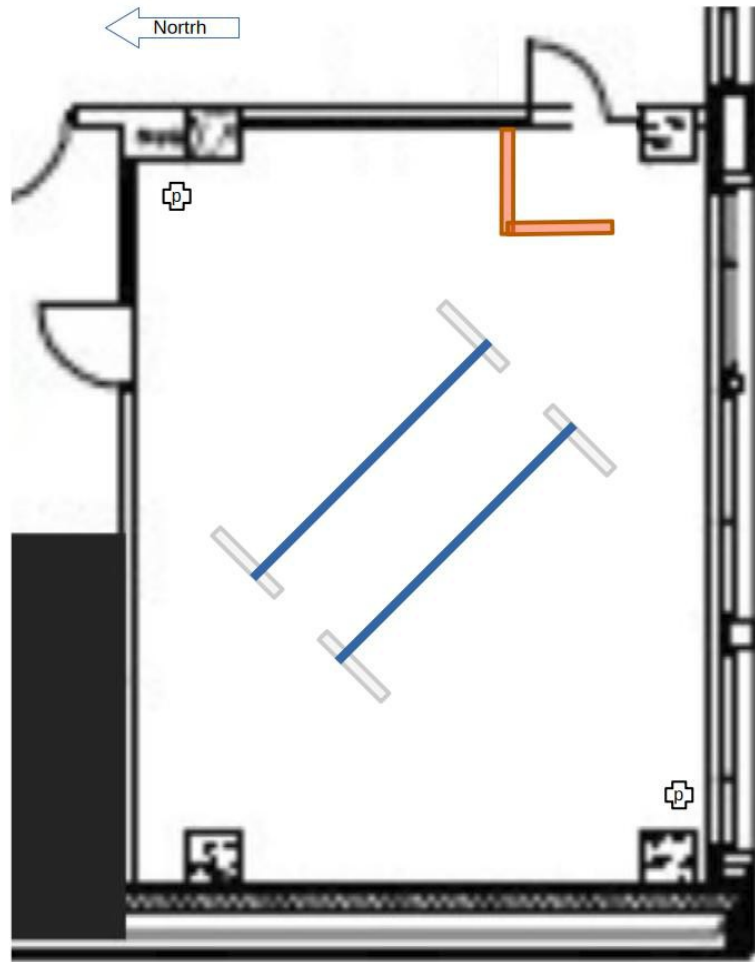
Figure 16: Scale and placement of screen

Electrical and Other Wiring

The room is well supplied with electrical outlets, but extension cords were required for the projectors because they were approximately 10 feet up. HDMI cables were run around the perimeter of the room and were taped down.

Safety

The stands for the screens were held down with sandbags. A path was marked on the floor with white tape so people walking around the perimeter were directed away from the stands. All wires on the floor were



taped down. *Figure 17: Floor plan of the exhibition space.*

Floor Plan

The + shapes marked p are projectors. The white shapes are stands supporting the screens, which are shaded blue.

A small table with the laptop was placed against the wall next to the entry kiosk marked in brown.

Images of the Exhibition



Figure 18: Screens illuminated by projectors, placed to make a corridor.



Figure 19: A viewer in the corridor between the screens.



Figure 20: Typical viewer reaction – standing against the walls.



Figure 21: Typical viewer reaction – identifying each other and themselves.