Reductio ad data - Artificial Insanity

Project Proposal

Description

Theme

Mental health is one of the most significant contemporary global health concerns. According to the World Health Organization's latest report, approximately one in eight people worldwide experience atypical neuropsychological states.[1] The Centre for Addiction and Mental Health estimates that by age 40, half of the Canadians will have grappled with mental illness.[2] The space here prohibiting an extensive foray into the history of psychology and psychiatry, we will concentrate on elements of the recent past relevant to our present framework, specifically focusing on data-based technological approaches to mental health.

While computational technologies as a whole have an extensive history of iatric use – to varying extents of success – in applications such as neurofeedback,[3] VR therapy,[4] telehealth[5] or video game-based digital intervention,[6] the past decade has seen a more targeted surge of attempts to apply data-based methodologies to psychiatry, especially in aspects of diagnosis (e.g. NLP nosology[7]), surveillance (e.g. digital pills[8]) and therapy (e.g. AI chatbots[9]).

Through this project, I aim to question such approaches, which are inevitably anchored in broader issues pervading the current cyberinformational landscape including the data bias of algorithmic normativity, the depersonalization of parametrization / disincarnation of dematerialization and market pressures of the digital economy. As we steadfastly hurtle into a virtual(ized) era of which we do not yet fully grasp the implications, I find it necessary to consider the ramifications of these and related technopolitical decisions on vulnerable individuals as well as society at large.

Specifically seeking to problematize the phenomenon of data-based diagnosis, I hope to call attention to the computational and cultural pitfalls such an endeavour could fall prey to, between the illusory objectivity of algorithmic classification, the amplification of nosological demographic asymmetry and the further alienation of techno-determinist optimization.

Readings

In *The Point of Collection*, Mimi Onuoha lists a number of issues surrounding big data, including selection bias and the inherent subjectivity of data collection, technological determinism and confirmation bias, the versatility and exploitability of data, and the systemic dynamics of the data economy. Beyond serving as a concise ethical roadmap to data collection and analysis more broadly, Onuoha's concerns are directly applicable to the aforementioned pitfalls in algorithmic approaches to mental health, framing many of the issues I hope to address in this project. Indeed, while ongoing historical awareness of institutional failure and abuse in psychiatry have brought a certain measure of attention to hierarchical dynamics and information privacy, which is consequently at least explicitly addressed in data-based iatric tools (see, for example, Amazon Comprehend Medical's statements on Personal Health Information), other aspects of data collection mentioned in *The Point of Collection*, such as selection and confirmation biases, are strikingly underevaluated. This points to specific elements pertinent to explore in this project discussing AI-based diagnostic tools, including data provenance, socioeconomic discrimination and cybersecurity.

Hito Steyerl's *A Sea of Data* speaks more broadly to the sociological ramifications of the current data economy, between the dehumanization of "machinic modes of perception," the discrimination of signal processing, the neo-feudalism of the digital economy, and the lethal risks of "automated apophenia". While not explicitly discussing examples involving mental health in her 2016 article, Steyerl references instances of classificatory, filtering or inferential failures and their variably tragic outcomes. These stressed for me the importance of individual experience in grasping the nature and consequences of algorithmic bias, pushing me in a more provocative direction (e.g. soliciting/disclosing viewer's own personal information). *A Sea of Data*, as well as her piece *How Not to Be Seen: A Fucking Didactic Educational .MOV File*, also discuss the importance of (in)visibility, which have brought me to further consider the implications of being able to visualize data and algorithmic processes, or not. A corollary of this phenomenon – and a recurrent element in my work – concerns the notion of steganography, and the possibility of dissimulating information in plain sight. In this regard, Steyerl's exploration of "corporate animism" probes the probabilistic underbelly of algorithmic classification, an element I wish to explore the explicitation of and reckoning with.

Form

Reductio ad data - Artificial Insanity is currently being envisioned as a epistemically-flawed AI-based diagnostic application, transforming a glut of user input into baselessly overconfident psychiatric assessments. The viewer will be presented with an interactive installation hybridizing virtual and material data-gathering elements, deploying web-based processes via an online questionnaire form as well as tangible electronics through electrophysiological sensors.

The questionnaire component will attempt to collect basic and health-related personal information, using the engaging form of magazine personality quizzes and their contemporary online counterparts to coerce the user into disclosure. The user will be prompted to answer broad censal questions, rate specific psychological traits and provide photographs and voice recordings. The contrast between the ludic appearance of the questionnaire and the sensitive information gathered intends to speak to the gamification of data collection and the self-commodification rampant in social media prosumerism.

The physiological component will take the form of a pulse sensor and an EEG headset, the data of which will be deliberately albeit surreptitiously misused to bolster and/or illustrate semi-arbitrary assessments derived from the questionnaire. The purpose here is to problematize notions of epistemic authority and neuroenchantment in psychiatric contexts.[10]

As the end-goal of his piece is informative satire, it is geared towards a broad audience, with no specific viewer in mind other than one open to question sociocultural discourses surrounding mental health. Pedagogical elements will be included with the intention of bringing the focus back to and raising awareness about what remains a serious societal issue.

Medium

The main interface will be a web application presenting the user-facing questionnaire form and piping the data from both the form and the electronic interfaces to a database. This will be created using a combination of front and back-end development. The structure of the web application will rely on the Node.js runtime and Express.js library, and will communicate with a MongoDB database and AWS S3 storage using the appropriate Node modules. HTML, CSS and Javascript will be used to create a functional but also visually appealing and engaging UI.

The tangible electronics component will employ both commercial as well as ad-hoc instruments. Electrophysiological data will be collected using an open-source EEG headset (OpenBCI's Ultracortex) as well as a pulse sensor (DFRobot's Heart Rate Monitor Sensor). The headset is paired with a proprietary microcontroller board (OpenBCI's CytonDaisy) connected to the associated GUI via BLE. The pulse sensor will be connected to a microcontroller communicating with the Node.js web application via an appropriate module.

If time remains, output medium will also be considered. One avenue would involve the use of a thermal printer as a tongue-in-cheek foil to the resolutely ultra-contemporary themes addressed and tools employed. This would require another microcontroller to serve as an output interface.

Data

The data will be sourced from spectators interacting with the piece, as a combination of information provided by the viewer in a web questionnaire and physiological data extracted from electronic interfaces.

Questionnaire data will more precisely include fields for censal (e.g. name, age/DOB, gender, nationality, telephone number, address, professional status, marital status, religion / belief system, etc..), health (e.g. eating sleeping habits, physical activity, mood states, medication use, etc..) and additional (e.g hobbies/interests, program of study, film/music interests, etc..) as well as a photograph and voice recording.

Physiological data, while essentially misleading, will nonetheless be collected to bolster the impression of scientific rigour. Pulse and EEG data will be stored and processed to generate visual representations of physiological activity.

Algorithms

The questionnaire data will be collected from HTML forms using jQuery and/or Multer, then piped to a MongoDB database using mongoose. The photo and audio recording will be stored in an AWS S3 bucket using multer-s3, with an access link referenced in the relevant MongoDB entry. The extent and variety of data analysis remains to be determined, but an array of interpretative processes are under consideration.

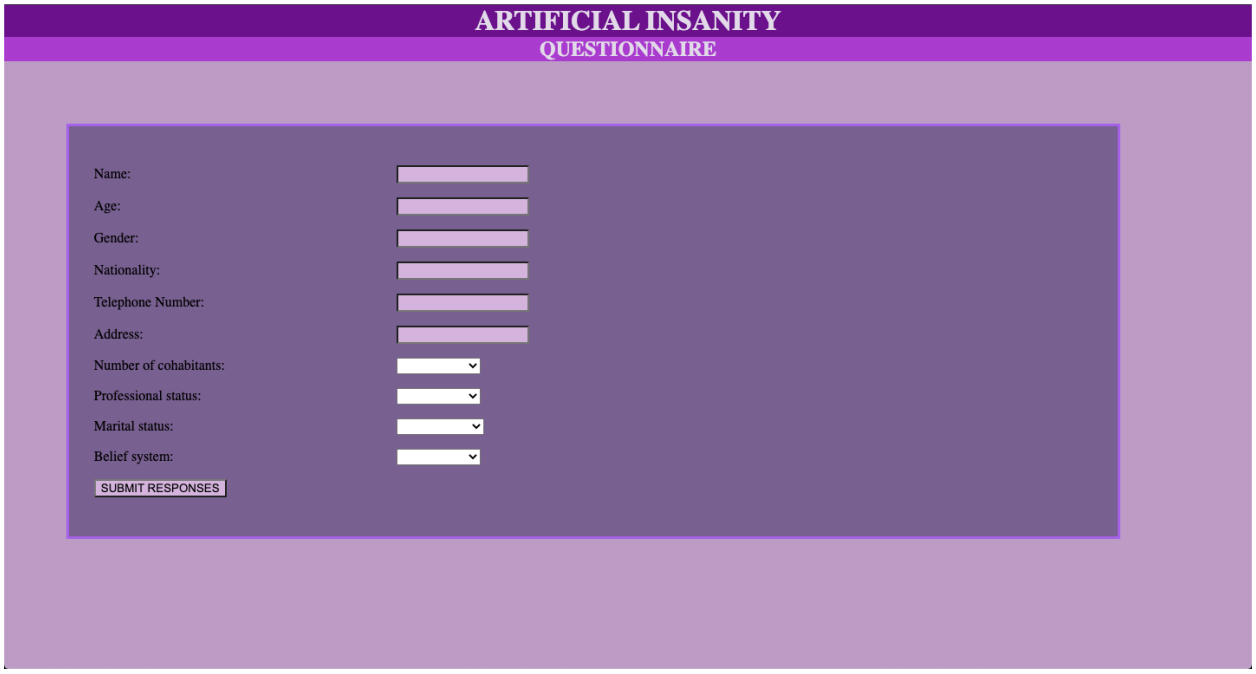
The main avenue I wish to explore is natural language processing (NLP), specifically named-entity recognition (NER) as a quick-and-dirty method to extract semantic information from text. Here, a commercial (AWS Amazon Comprehend Medical) or ad-hoc model could be used to extract medical semantic entities from user data, linking extracted entities to the CDC's International Classification of Diseases, 10th revision (ICD-10), principally diagnostic codes.

Given the imperfect nature of the probabilistic results NLP produces, other elements of personal information in the form as well as physiological data (e.g. pulse) will be used to either provide alternative or supplementary analytical data points. One example would be to use the uploaded photograph to derive fallacious assessments of health, harnessing for example the ml5.js/tensorflow.js library to extract body dimension, position and posture data to implement a contemporary AI-augmented version of phrenology.

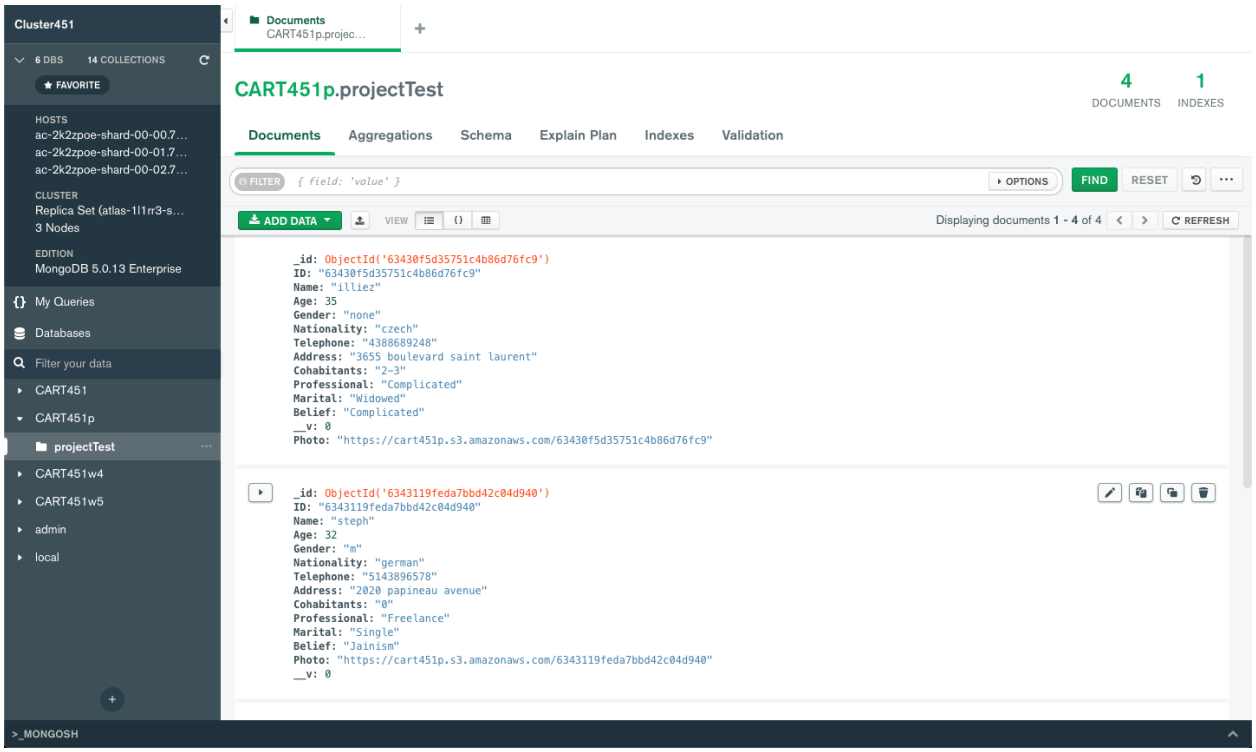
Remaining uncertainties:

- Debug preventDefault on form submission
- Amazon S3 for photo and sound clips
- Debug page routing (e.g. express res.redirect() + HTTP headers)
- How to get data from CytonDaisy to node app → OpenBCI Node.js library
- Piggybacking pulse sensor on CytonDaisy?

Storyboard



Form submission page ; aesthetics/CSS to be revised



MongoDB database w/ example entries (Photo field only filled after upload)

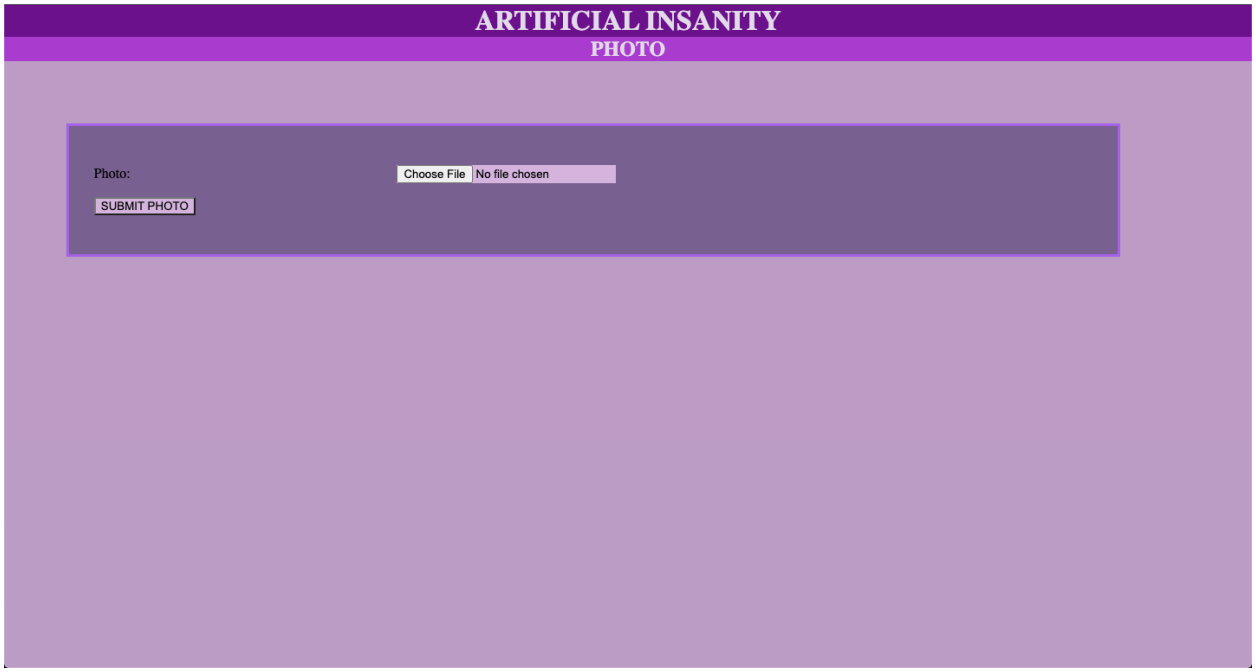
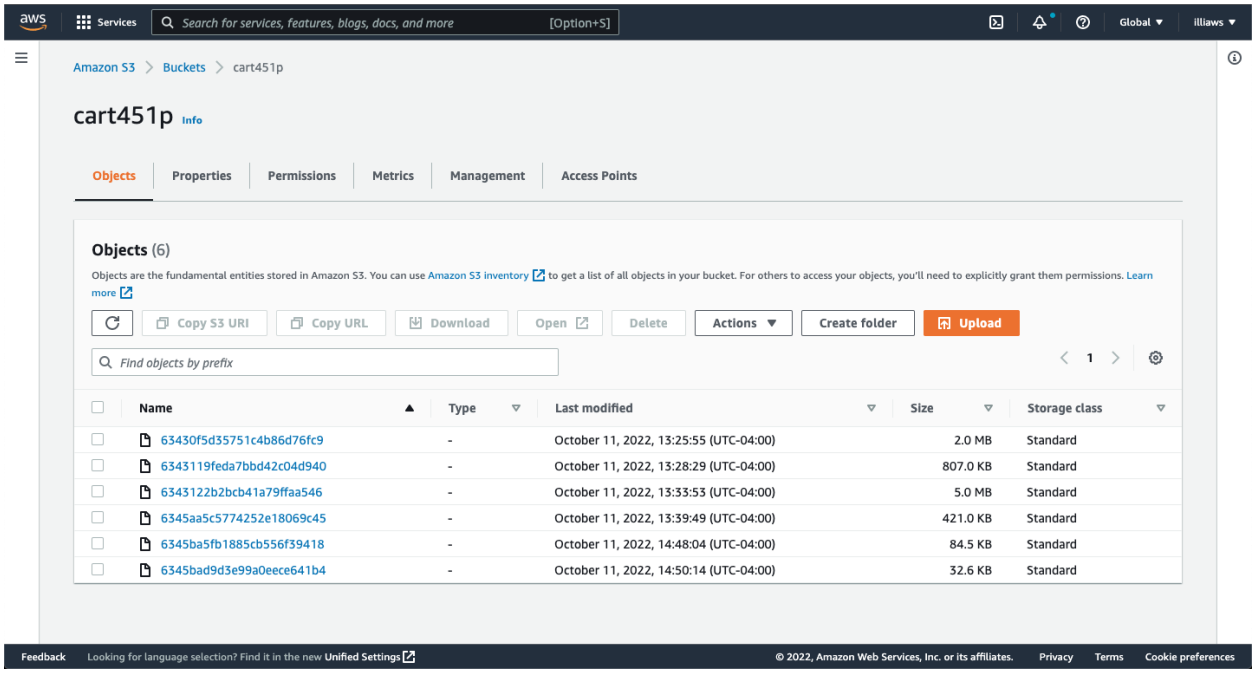
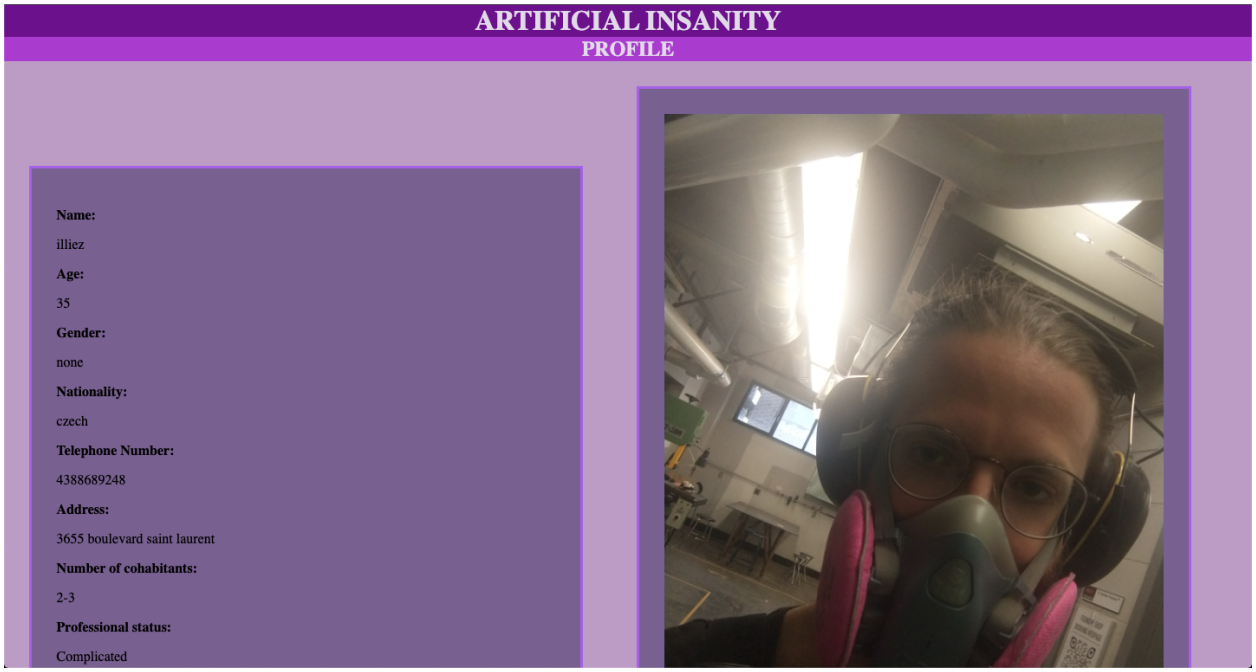


Photo upload page ; aesthetics/CSS to be revised

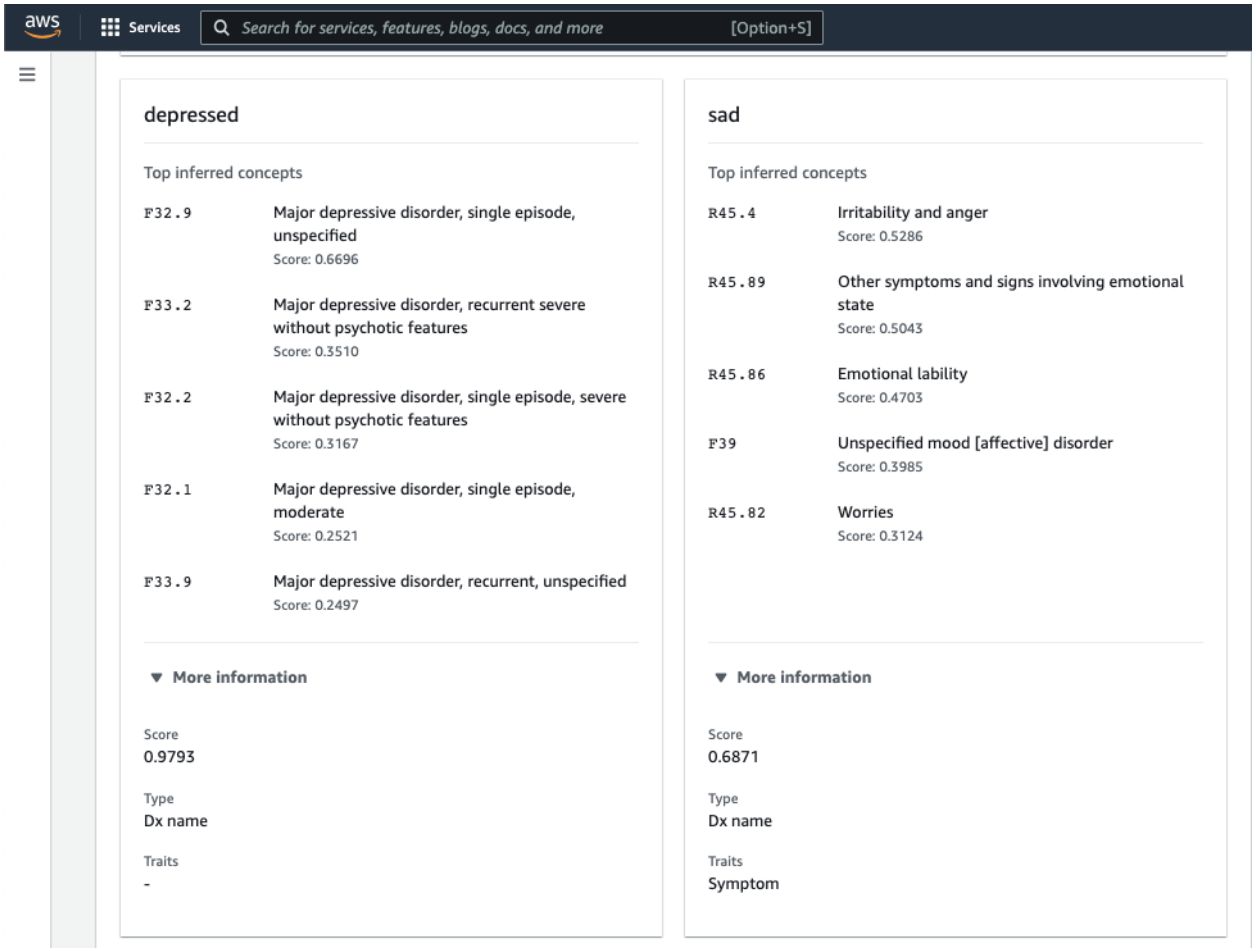


AWS S3 bucket w/ example uploaded photos, keyed to MongoDB ID



Example profile page, w/ questionnaire information and photo ; aesthetics/CSS to be revised

Example visualization of physiological data / EEG



Example AWS Amazon Comprehend Medical NLP/NER result

Inspiration

MANIC VR is a virtual reality piece attempting to creatively render an impression of an experience of bipolarity. Built off of Kalina's Bertin earlier autoethnographic documentary *Manic*, which delves into her family history, both biographical and medical, MANIC VR seeks to immerse the viewer further, using animated visuals and audio recordings to foster an impersonation of visual and acoustic phenomena described by her siblings during episodes of psychiatric distress. While the piece can at times seem slightly gimmicky in its use of computer-generated imagery to convey a profoundly subjective experience, it remains an effective and affective exploration of VR as a neurophenomenological tool (see also Conversations between self and self as Sigmund Freud for insight into the empathogenic value of embodiment), speaking to the psychological, interpersonal and social ramifications of heightened states of atypical cognitive experiences. Deeply moved while watching the initial documentary, I was apprehensive with respect to its transformation into a VR piece, but have ultimately come to admire Bertin's intention and attempt to use the media at her disposal to share her siblings' worldview as faithfully as possible. Having had the opportunity to speak with the artist, it became evident that the motivation was to honour her siblings' narrative and raise awareness about the discrimination, isolation and violence neurpsychological non-normativity can beget. This has joined a host of other influences, conversations and experiences in inspiring me to discuss mental health more openly, but also more earnestly, with less irony and sardonicism.

An insightful project in terms of data art more specifically is *White Collar Crime Risk Zones*, a satirical web piece proposing a predictive policing model for white collar crime. The website takes the form of an interactive map of New York City's Manhattan Island, distributed in location-based geohashes. Each geohash presents statistics for the most likely type of white collar crime in the area as well as nearby financial firms. By inverting the statistical skew of conventional criminality data mining, usually focusing on blue-collar crime and consequently overreliant on racist and classist priors which get computed into the resulting models, the artists expose this methodological limitations and ideological bias. Their white paper demonstrates a discursive performance reminiscent of cultural jamming duo Yes Men, and serving as an inspiration for the use of brusque satire. A particularly compelling element in *White Collar Crime Risk Zones* is a section on the website titled "Most Likely Suspect", an algorithmically-generated face based on LinkedIn-sourced and statistically averaged faces of corporate executives in the geohash. This raises questions about the deindividualization of crime prediction models, caught between the Scylla of inadequate scope where, according to the artists "all entities within high risk zones are treated as uniformly suspicious" and the Charybdis of averaged stereotype, playing into a pot-pourri of biases (source misattribution, own-group recognition, etc..). This automated depersonalization is also readily apparent in This Person Does Not Exist, a website presenting fictitious digital portraits generated by a GAN model trained using a large corpus of photos.

Formally, this piece might be inspired by works tackling notions of cyberbureaucracy. A first example is Alexei Shulgin's 1997 *Form Art*, an early instance of net art exploring the interplay between form and function through HTML forms, where the content (network art, digital aesthetics, online collaboration), becomes subsumed in the parafunctional context (or vice-versa). For all practical intents and purposes, *Form Art* is a nonteleological ludic series of interactions with UI elements such as checkboxes, radio buttons, text fields, drop-down menus and so on, but becomes an almost cyber-architectural paragon, a quintessential ur-form of the Internet in its primeval self-awareness. These themes have more recently revisited by Pippin Barr under the guise of triple post – -Internet, -ironic and -capitalist – aesthetics in *It is as if you were doing work*. In this piece, the user engages in a simulacrum of cyberbureaucracy, accomplishing required tasks with variable agency and latitude of action. Both of these examples point to the byzantine nature of administrative formalization (whether Sisyphean or Kafkaesque), amplified by the routed nature of online navigation inherent to the hypertextual structure undergirding the Internet. This speaks to David Graeber's notion of "utopian thinking" addressed in Hito Steyerl's *A Sea of Data*, an element of built-in UX failure, or at least frustration, I intend to include in this piece.

Bibliography

1. World Health Organization. World mental health report: transforming mental health for all. 2022.
2. Smetanin, Paul, Carla Briante, Minhal Khan, David Stiff, and Sheeba Ahmad. "The life and economic impact of major mental illnesses in Canada." (2015).
3. Marzbani, Hengameh, Hamid Reza Marateb, and Marjan Mansourian. "Neurofeedback: a comprehensive review on system design, methodology and clinical applications." Basic and clinical neuroscience 7, no. 2 (2016): 143.
4. Liu, Zhen, Lingfeng Ren, Chang Xiao, Ke Zhang, and Peter Demian. "Virtual reality aided therapy towards health 4.0: a two-decade bibliometric analysis." International Journal of Environmental Research and Public Health 19, no. 3 (2022): 1525.
5. Bulkes, Nyssa Z., Kaley Davis, Brian Kay, and Bradley C. Riemann. "Comparing efficacy of telehealth to in-person mental health care in intensive-treatment-seeking adults." Journal of Psychiatric Research 145 (2022): 347-352.
6. Li, Jinhui, Yin-Leng Theng, and Schubert Foo. "Game-based digital interventions for depression therapy: a systematic review and meta-analysis." Cyberpsychology, Behavior, and Social Networking 17, no. 8 (2014): 519-527.
7. Guzman, Benedict, Isabel Metzger, Yindalon Aphinyanaphongs, and Himanshu Grover. "Assessment of amazon comprehend medical: Medication information extraction." arXiv preprint arXiv:2002.00481 (2020).
8. de Miguel Beriain, Iñigo, and Marina Morla González. "'Digital pills' for mental diseases: an ethical and social analysis of the issues behind the concept." Journal of Law and the Biosciences 7, no. 1 (2020): lsaa040.
9. Nadarzynski, Tom, Oliver Miles, Aimee Cowie, and Damien Ridge. "Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study." Digital health 5 (2019): 2055207619871808.
10. Ali, Sabrina S., Michael Lifshitz, and Amir Raz. "Empirical neuroenchantment: from reading minds to thinking critically." Frontiers in human neuroscience 8 (2014): 357.