

Security Noir: Critical Designs for Security & Privacy

Didem Demirag
UQAM
Montreal, QC, Canada

Leah Zhang-Kennedy
University of Waterloo
Stratford, ON, Canada

Jeremy Clark
j.clark@concordia.ca
Concordia University
Montreal, QC, Canada

Abstract

In this paper, we investigate an alternative method for communicating the somewhat elusive security risks and privacy invasions of technology, the web, social networks, and other aspects of modern life. Specifically, we draw from the *critical design* movement popularized by Dunne and Raby in the 90s. In this project, we have conceptualized 10 new technological designs and described their use through short vignettes. These designs are imaginary but plausible technologies that illustrate a security or privacy risk through absurdity, satire, or humour. We also consider the ability to co-design with generative artificial intelligence tools, finding the tools cannot (yet?) replicate our design process and produce satisfactory designs, but these tools are useful today in a ‘co-pilot’ role for generating high volumes of ideas for human designers to curate and refine.

ACM Reference Format:

Didem Demirag, Leah Zhang-Kennedy, and Jeremy Clark. 2026. Security Noir: Critical Designs for Security & Privacy. In . ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 Introductory Remarks

Critical design is commonly attributed to Dunne and Raby [5, 7] however it can be seen as a refinement of earlier movements in arts and design, beginning with the Italian radical design movement and extending through critical practice in HCI [18]. Consider the three examples provided in Table 1. Each describes the design of an object or technology. In contrast to industrial design, these technologies are not oriented around problem-solving, nor the creation of solutions that are functional, useful, and pleasurable, nor are they meant for mass production. Rather they offer a subtle criticism by illustrating something—electromagnetic radiation, the inevitability of death, the ubiquity and stealth of surveillance—that typically goes unquestioned. They are designed to invoke a dilemma for the user, raising questions, challenging values, and stimulating debate. Critical design considers these goals to be the function (or ‘para-function’ [5]) of the concept.

Critical design has significant overlap with conceptual art, utilizing common methods, and has been disseminated through exhibitions and the art literature. However critical design works by showcasing things that are plausible, believable, and easy for the

reader to situate, requiring design practice. Viewing critical design as art creates challenges. Consider the compass table (from Table 1) in the context of an art gallery exhibit—it is meant to illustrate electromagnetic radiation in every day life, in one’s home, in different positions and orientations within that home—all of this is lost when it sits statically in a gallery. Should critical design be displayed in galleries, a vignette or film or narrative will often accompany the object to describe its context of use.

1.1 Contributions

Security & Privacy. Security and privacy risks are often intangible and invisible, making concepts difficult to communicate to everyday users. We saw a parallel to electromagnetic radiation in this regard, and considered if critical design could similarly aid in illuminating risks.

Vignettes. Critical designs are commonly described through vignettes. Occasionally they are built and displayed. In the case of Dunne and Raby’s Compass Table (and the seven other objects in their *Placebo* project [7]), the table was built and given to individuals to use. This field study was followed by an exit survey on how the participants felt about the object. In this paper, we describe our designs through short vignettes that illustrate its design and use case. If they are a successful design, the reader should easily imagine its use in their life and question it critically, resulting in its ‘rhetorical use’ [18] rather than practical use. These designs are transgressive and intended to provoke the reader’s emotions; invoking humour, satire, ridicule, poetry, playfulness, lewdness, appropriation, irreverence, or deliberate overcompensation.

Methodology. In order to generate designs, we report on four methodologies. First, we (a) used traditional design methodology to generate 10 designs (Section 2). We then experimented with the generative AI tool ChatGPT version GPT-4g. Then we (b) fed our designs to ChatGPT and asked it to generate new designs (Section 3). It effortlessly generated designs that initially seemed sound, but under greater scrutiny, were lacking. We then (c) experimented with giving ChatGPT a sketch of what a possible design could involve and then dialogued with it to develop and refine the design, using it as a collaborative tool (Section 4). We found this approach produced good results and was less effort than method (a). Finally we (d) educated ChatGPT on the entire design process (Section 5), trying to involve it in the earliest stages of design (as opposed to (c) where we involved it after we had an initial sketch of an idea). We found this approach was less effective. In our experience, ChatGPT is capable of assisting in designs but requires human guidance and intervention. Generative AI’s strength is in generating large volumes of ideas, while humans can contribute by curating and refining the best ideas.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference’17, July 2017, Washington, DC, USA

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

Table 1: Examples of Critical Design

Work	Vignette
<i>Compass Table</i> , A. Dunne F. Raby	"This table reminds you that electronic objects extend beyond their visible limits. The 25 compasses set into its surface twitch and spin when objects like mobile phones or laptop computers are placed on it. The twitching needles can be interpreted as being either sinister or charming, depending on the viewer's state of mind. When we designed the compass table, we wondered if a neat-freak might try to make all the needles line up, ignoring the architectural space of the room in favour of the Earth's magnetic field." Quoted from [7].
<i>Life Counter</i> , I. Matsumoto	"With Life Counter (2001), you choose how many years you would like to or expect to live for and start the counter. Once activated, it counts down the selected time span at four different rates: the number of years, days, hours or seconds to go are shown on different faces. Depending on which face you choose to display, you may feel very relaxed as the years stretch out ahead or begin to panic as you see your life speed away before your eyes. The counter is designed to be visually unassuming and could easily fit into the slightly retro-futuristic style of the moment." Quoted from [7].
<i>Open Informant</i> , J. Ardern Y. Ushigome A. Jain	"Open Informant (2013) takes the form of a networked object including a phone app and e-ink badge. The app searches your communications for National Security Agency trigger words and then sends text fragments containing these words to the badge worn by the user for public display. Using the body as an instrument for protest, the badge becomes a means of rendering our own voice visible in an otherwise faceless technological panopticon." Quoted from [18].

1.2 Related Work

Critical Design for Security and Privacy. Much has been written on critical design and Malpass provides an excellent survey [18]. Since our focus is security and privacy, we only summarize the following concepts, situated within or close to critical design, since they deal with issues of privacy:

- **The Pillow.** Dunne and Gaver's design takes the form factor of a transparent pillow and plays radio signals it picks up from the technology of that time: e.g., phones, pagers, and baby monitors [6]. It challenges users to consider how confidential these signals are.
- **Open Informant.** Described in Table 1, this design exposes nation-state surveillance.
- **United Mirco-Kingdoms.** Dunne and Raby reimagine the UK as four 'supershires' where one is occupied by 'digitarians:' extreme technocrats ruled by algorithms in a state of total surveillance.
- **Alternet.** While not critical design, Gold's Alternet has the similar intention of provoking thought through design [10]. The internet is re-imagined with users retaining full control over their private data, touching on web tracking and our growing digital dossiers.

We are unaware of any attempt to build a systematic set of examples based on security or privacy, nor much attention to it from academic research papers, however critical design has been used in other areas of computer science. The uptake of critical design in HCI research is encouraged based on the anticipated value critical design will add to the field [1]. Speculative design is another related concept that HCI researchers rely on. An example is to analyze and question the usage of a certain technology and the consequences of it in the near future [14]. Ludic design is also

utilised to create designs that do not dictate a certain purpose or usage but rather leave the interpretation to users [9].

Methodologies for Design Creation and Evaluation. While design process can be based on well-established methods, that are relevant to us, like whiteboarding, using sticky notes, affinity diagrams; the integration of AI to the creation process has been proliferated with the abundance and the variety of AI-based tools. These tools can be used to generate the content itself or help the designer along the process of creation and ideation. There are also works that examine how users prompt generative AI ([23], [3]); utilize generative AI for co-ideation or co-creation [25],[4].

There are also efforts towards creation of computational tools to that would assist task that are manually done in general. One line of work is about creating tools to group design aspects to help ideation and exploration of design solutions[13]. Speculative design is also used to imagine the future of the tools based on generative AI [16]. Custom-made AI tools are proposed for specific design goals like integrating context awareness[8].

While we only created the designs and did not apply any techniques to evaluate them, there are several methodologies for this purpose. User studies like interviews and surveys are one of the most common techniques employed [23], [3]. They are deployed to evaluate different aspects of the design process like assessing the utility of the AI-tools used during the design process, understanding users' needs or deploying a participatory workshop. Methods to assess the creativity of AI-generated content are also proposed [2].

Alternative Pedagogy. Somewhat related to critical design, there are works on finding alternative ways (like design fiction [17], role playing game [19], science fiction and design fiction [26]) to explain tricky concepts not only in security but also in other fields like sensing technologies.

CS education literature has explored how to teach concepts related to computer science in general or cybersecurity to a diverse audience. Some of which are understanding how CS students perceive AI and cybersecurity [21], using visual cryptography for explaining concepts to K-12 students [22], how to initiate a security mindset [12]. Studying the metaphors used in CS education gives important insights about how students perceive certain concepts [11].

2 Methodology A: Traditional Design

2.1 Methodology

Our research goal was to create critical design vignettes illustrating security and privacy issues. We will first describe our design methodology for human designers. We began by reviewing the critical design literature, and we settled on delivering vignettes, which would be short paragraphs describing a new technology and illustrating, with short narratives, how it functions and the feelings it might invoke. The technology is meant to be plausible in the near future (the reader can conceive why someone, somewhere might think it is a good idea), while also showing the darker side of convenience, risks that are normalized and overlooked, or ways a user might be betrayed by the technology. During this period, we were encouraged to notice ‘found examples’ of critical design in the wild, although perhaps not expressed that way.¹

We settled, after some trial and error, on a digital (Apple Freeform) ‘sticky notes’ methodology to generate affinity maps, where we clustered our notes in four groups: (1) security and privacy concerns (e.g., ‘if you don’t pay for the product, you are the product’), misconceptions (e.g., ‘nothing to hide’), or principles (e.g., ‘minimal disclosure’) that could be the basis of a critique; (2) domains where users might especially care about security or privacy (e.g., health records or social networks); (3) a form factor for the new technology (e.g., haptics or extended reality); and (4) a possible user response (e.g., amused or embarrassed). Figure 1 shows our board. Artefacts, including this board, are available on our GitHub repository.² We brainstormed around individual notes and drew random combinations to spur divergent thought. The design methodology provided structure but did not make designing simple. Each of our 10 designs was quite difficult to develop, each taking many hours of discussion and dozens of discarded ideas that we could not translate into designs.

New methodology

- We started by analyzing the different methods of design (divergent, convergent, mapping, prototyping...)
- We decided to follow a divergent method: the *How might we?* (HMW) technique from design thinking, starting from the general problem statement, each of us created *How might we?* cards. The cards reflected different aims we had for the design we will create like how might we provoke reflection of societal and ethical issues, make the narratives more relatable or create ideation cards that can eventually help with coming up with the vignettes. We also tried ChatGPT for creating

the card (**Any reflections on how it went?**). We then grouped the HMW cards around similar topics and applied dot voting to determine the most important questions.

- We also gathered images, news, concepts etc that reminded us of the project and inspired us in an inspiration board.
- We grouped the HMW cards we created (with and without help of ChatGPT) under themes. We came up with five suits (themes) for ideation cards (Critical, design, scenario, narrative, stickiness). Then we decided to keep only critical, which determines the main topic to be critiqued, and design, which determines the medium. Then redistributed the cards under these two themes. The rest of the themes serves more as a checklist for a quality check for the designs created (for evaluating the narrative style, whether the vignette is memorable and powerful).
- We divided the themes among ourselves so that each of us had cards from each suit (critical and design). Then we each created ideation cards based on the theme with the format: title and description.
- We then tested the cards. Each of us drew one card from each suit and came up with a design. We repeated this to see how we draw cards (e.g., keep drawing 2 cards until we find an idea, keep the one we like and continue drawing from the other suit, draw more than 2 cards from each category) and whether the combinations lead to ideas.
- We then refined the cards. Each of us looked at all of the cards to either present alternatives to the card content itself (changes regarding the title and/or the description) or we suggested to keep them as is. We also merged/unmerged cards, discussed whether the card is a good fit for its suit.
- After the initial refinement, we applied the MoSCoW (Must Have, Should Have, Could Have and Won’t Have this time) method to determine the most important, relevant cards to keep in the deck.
- After several iterations of refinement of the cards involving merging, deleting, moving the cards between the two categories and rewording; we finalized a deck of 37 critique and 30 design cards.
- We conducted an autoethnography for our design process (**See section XX**). Each of us created a design a day using the ideation cards. We created a template for our diary entries. We each created a first set of diary entries, and then we refined the fields in each entry to provide a comprehensive overview of our design process. (**Add the fields we filled each time, also add the total number of design, how long it took etc.**)
- We also tried co-designing and creating three designs each, using the same two cards (one from design, one from critical)
- We performed thematic analysis on our diary entries (**See section XX**).

2.2 Designs

¹For example, well-known cryptographer Matthew Green tweeted: *There’s an entire class of browsing many folks will feel uncomfortable doing if their browser has a little icon in the window that says “Hi <User> I know this is you!” This is Human Behavior 101.*

²GitHub: <https://github.com/MadibaGroup/2017-Critical-Design>

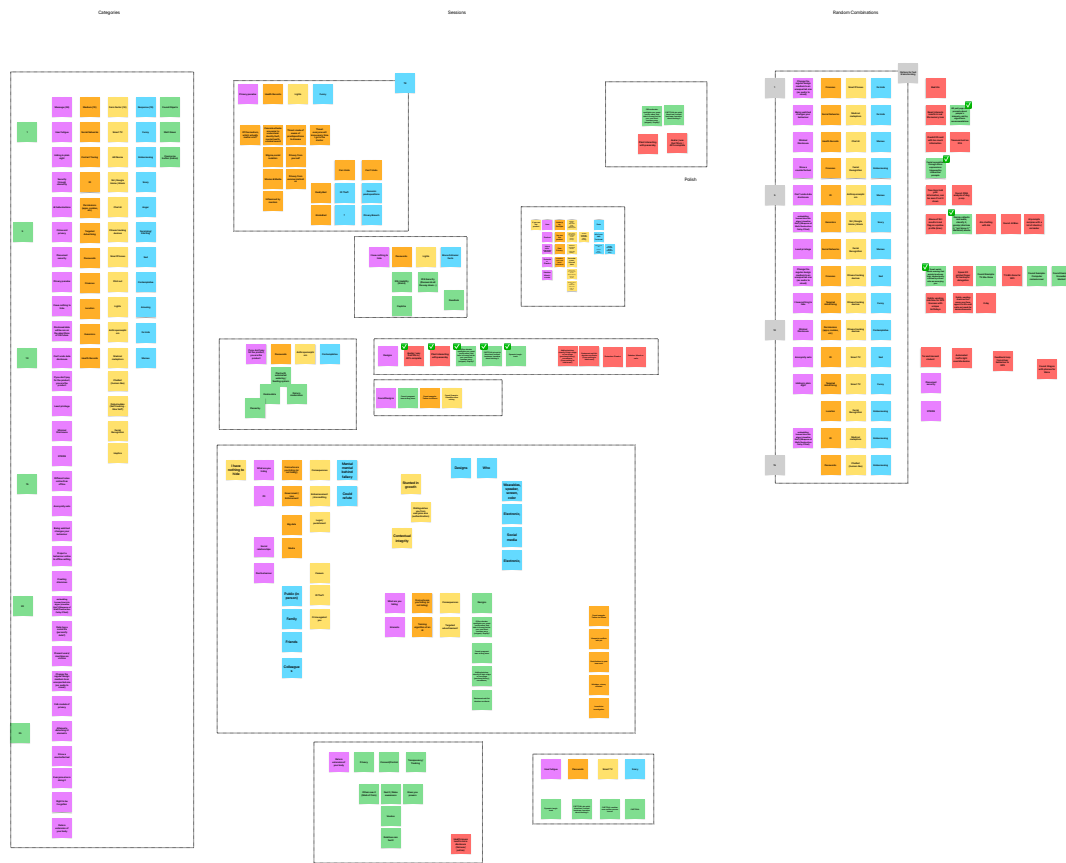


Figure 1: Final affinity board after brainstorming. Not intended to be legible; intended to communicate the scale of the board.

Crystal Avatars. Crystal avatars are created instantly but only gets better over time. At first, they are not recognizable and cannot be differentiated from other users, but as users browse the site, small details crystallize. Alice is happy to save time registering for the site and have a personal avatar, which automatically added a cat accessory after she spent a lunch hour reading about hairball treatments. However she grows uneasy about the "Woman, Life, Freedom" badge, a cause she believes in but only vocalizes discriminately.

Calm Watch. Calm Watch is a smart watch that tracks anxiety levels based on several biomarkers. When the anxiety level of the wearer crosses a threshold, a haptic vibration pulse prompts the user to look at the watch. From there, they can select a variety of calming techniques, including breathing exercises and guided meditation. Bob is nervous to talk to Alice and becomes flustered when his watch starts vibrating mid-conversation, just loud enough for Alice to hear it and notice his fidgeting with the watch.

Chatty. The latest software update for Alice's voice-activated home assistant Chatty adds machine learning to better infer Alice's commands, such as adding items to her shopping list and suggesting music she might like. It does not always know whether Alice is talking to it or not, so over time it picks up on pieces of Alice's life and has trouble unlearning them. Alice is startled but pleased when Chatty chimes in with the name of 'that actor who was in that thing' she was telling her friend about. She is bemused to find a specialized vinyl cleaning solution on her shopping list after she said Bob's records smells fishy.

Dynamic Laughtrack. Digital television content, such as sitcoms, encode laughtracks as a series of cues rather than an actual recording. Smart TVs listen and classify the viewer's level of laughter on a 10 point scale (with fine grain training over time). The TV adds laughter to the content in a sensitive and considerate manner, where the laughtrack is only marginally higher on the scale than the viewer's own laughter. This gently nudges the viewer to greater enjoyment of the program without bombarding her. It also mixes in actual past recordings of the viewer's own laughter to capitalize on emotional mirroring.

Eleventh Finger. Biometric authentication based on fingerprints is generally user-friendly and fast. Eleventh finger is a 3D printed rubber finger, customized with the user's fingerprint. You can put it on a keychain and use it on cold winter days when you do not want to remove your gloves. Alice lets her friend Bob borrow it when he stays at her house. It can also serve as a backup if the worst happens.

GOTTCHA. GOTTCHA is a human-detection system to protect online accounts being accessed by bots. It invokes the device camera to analyze if a live human is using the device. To protect against video replays or machine-generated video, it unexpectedly prompts the user with a randomly selected image and carefully measures their reaction to it. Micro-expressions are involuntary facial displays of emotion that are too fast to mimic artificially. GOTTCHA's image bank can provoke disgust, anger, fear, sadness, happiness, surprise or contempt. Over time, Alice stops visiting websites that use GOTTCHA because of its capricious tendency to mix in disturbing images.

No(i)sy elevator. Elevator music has never been regarded as elevating your mood. This is why employees will be happier riding the noisy elevator. When it recognizes a user, it spins up a selection from their most-played songs on major music streaming platforms. Alice finds hearing her favourite Miles Davis song grounds her at the start of each work day, while Bob and Carol, coincidentally riding the elevator together, discover their shared love for 90s britpop. David is a bit more concerned about the profanity ridden banger that the nosy elevator plays for him and his boss.

Receiver plant. The receiver plant in the front yard of Alice records the information of the people that pass by. It is a greedy plant that needs to absorb enough bluetooth data to grow. Bob is singlehandedly responsible for its lustrous canopy from walking by Alice's house every day, mostly while checking his work emails on his phone. Perhaps the plant overstepped when it displayed, "you are 15 minutes late today, are you sure you can make it to work on time?"

ToS Fishing. ToS Fishing is a digital game where players earn points by hunting down excessive legalese on websites, including terms of service (ToS), privacy policies, and cookie policies. To play, users simply copy and paste the URL of the legalese. If the ToS has been seen before, the users are awarded points based on its word count. If it has not been seen before, it is reviewed by a human for validity and word count—and in this case, the user gets a finder's bonus. To mitigate people from launching custom websites, the game only accepts sites from the Alexis top million. Users display their aggregate score on a leaderboard with a profile showing their top catches (word count only, not the website), and can earn various badges for playing consistently.

Wifi Projector. Minimalist wifi routers have a single light indicator to show internet connectivity, such as a green light as opposed to red light. This understates the interesting and intriguing data flowing through the device as users browse the internet. By contrast, the wifi projector is a talkative router that projects visualizations onto the ceiling above where it is placed. Every website that pulls in scripts and cookies from other domains is displayed as a constellation of stars of various colours and sizes, which slowly fade away. While Alice's visit to dictionary.com is a beautiful galaxy, she is also perturbed by the number of ads and tracking companies surveilling her.

2.3 Discussion

We will provide some of our *stems*³ for one example: Eleventh Finger. We felt that many of our designs were privacy-focused and we wanted a design that was explicitly security-focused. STRIDE is a threat modelling framework that outlines six high level categories of threats: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege [24]. We considered possible designs for each and Eleventh Finger came from our ideation around Spoofing. We tried to avoid privacy implications, even though they are important as well (the ACLU and EPIC were raising concerns about Disney's use of biometric fingerprint scanning at its parks). Biometrics are often considered the holy grail of authentication by non-experts for their simplicity, however they have drawbacks in the areas of entropy, revocation after theft, and delegation. At the hacker conference Def Con, researchers showed that biometric scanners, like those found in modern smartphones, can be tricked using 3D-printed fingerprints [15]. We were open to using (dark) humour to avoid sounding too earnest in our designs. The reader can likely see how Eleventh Finger can emerge from these stems, however it is less apparent where the stems themselves came from (e.g., from our academic backgrounds and personal experiences), nor the large number of stems we discarded because we could not see connections between them.

³In this context, 'stems' refer to the core ideas, influences, or conceptual elements that contribute to the development of a design.

3 Methodology B: LLM Generated Designs

3.1 Methodology

The design concepts in the previous section were developed before or independent from generative AI tools like ChatGPT. However, once these tools became available, we felt it would be interesting to determine how they might assist in generating designs. In particular, since we are developing written vignettes, large language models (LLMs) are the appropriate machine learning technique suited to this purpose. We began with the simplest methodology: we fed ChatGPT version GPT-4g example designs from Section 2 and asked for the generation of new designs. All transcripts are available on our GitHub repository.⁴

3.2 Designs

The following are the first two designs generated by ChatGPT.

Transparent Wallet. Alice's digital wallet automatically categorizes her purchases and shares anonymized spending habits on a public leaderboard. She enjoys seeing her eco-friendly purchases earning her top marks in the green consumer category. However, Alice is troubled when she notices ads for luxury items appearing more frequently after her high spending months are tracked and shared by the wallet [ChatGPT Transcript].

Echo Chamber Filter. A social media app uses an advanced algorithm to filter content to match users' perceived political views and biases. At first, Alice appreciates the streamlined news feed with content she agrees with. Over time, she realizes her interactions are limited to a narrow worldview, reinforcing her existing beliefs and making it difficult to encounter differing perspectives or engage in meaningful discussions [ChatGPT Transcript].

3.3 Discussion

Transparent Wallet can be considered an example of critical design but Echo Chamber Filter is not. The latter makes a valid critique but it does not describe a new (speculative) technology, personalized feeds have been utilized by social media platforms for years and have been thoroughly criticized for exactly this point (e.g., [20]). We asked for five designs and three of the designs, including Transparent Wallet, were based on the threat of data sharing being weaponized against the user (twice in the form of ads and once in the form of insurance premiums), demonstrating that ChatGPT can be highly repetitive.

Of all the five generated designs, Transparent Wallet was the closest to our own designs. It uses a similar structure: the design, plausible example of its intended use, and an unintended misuse. Gamification, data sharing, and algorithmic inference are all stems we ourselves considered (and even used) for our design. However we would work on improving what is being critiqued, which is pedestrian and straight-forward, strengthen the plausibility of Alice's willingness to share obviously sensitive personal data, and try to make the narrative clearer (we read it several times to 'get it').

This led us to consider a methodology where we co-designed with ChatGPT to iterate and refine designs.

4 Methodology C: LLM Collaborative Co-Designs

4.1 Methodology

We began each design session with a new window in ChatGPT and started each session with the same prompt: a few sentences about using critical design for security and privacy (assuming it could reference for itself what these terms mean) and our manually created designs from Method A as samples. We then input a set of stems we had brainstormed and asked for ideas about possible designs. As an LLM, we did not concern ourselves with making the stems coherent or even grammatically correct—just enough words to trigger data around the subject. An example set of stems (which was used for the design ElephantMind below) are:

Stems. Avoiding small talk; to know more about the person you meet; the app reminds you the last conversation you had with them; constantly reminding something trivial (Alice promised Bob to grab a coffee sometime. Next time Bob sees Alice the app reminds Bob what Alice said.); Summarize the last conversation you had with them; For a person you meet for the first time, device gives you some background info about them so that you choose the right topic of conversation.; Name for the app: something that is about never forgetting like having a memory like an elephant; Photographic memory; Or something related to recording everything

We then asked for examples along the lines of: technologies that could serve as the form factor for a critical design, potential names for the design, and examples of 'things that could go wrong' with the technology. In most cases, we would ask for 10 examples. We would consider the examples and apply redirection if the examples were not what we were looking for; if they were, we would curate 1 or 2 that were workable and ask for further examples. At certain points, we would ask it to write a draft vignette and then manually edit it.

4.2 Designs

BloodKey. BloodKey is a form of two-factor authentication that can protect online accounts from stolen passwords. With BloodKey, after providing their password, users authenticate with a pin-prick blood test using an inexpensive biomedical Bluetooth gadget. BloodKey was quickly adopted by leading websites because it was very hard to fake, despite its invasive nature. However, Alice grew tired of the daily inconvenience of pricking her finger, so she gradually stopped logging into her social media accounts altogether, while Bob's medical condition prevented him from using it at all. A media frenzy ensued when the most popular free-to-play online game sold users' genomic data to pharmaceutical companies, boosting their revenue streams [ChatGPT Transcript].

⁴GitHub: <https://github.com/MadibaGroup/2017-Critical-Design>

Civic Compass. Civic Compass is an add-on for popular GPS navigation apps that highlights the ethical considerations of various routes, such as avoiding surveillance hotspots or environmentally protected habitats. Public policy debates have emerged over the app's potential abuse by criminals. When Alice is in a rush, she ignores its suggestions. She is also disturbed to see that most of the surveillance hotspots are in impoverished, marginalized neighborhoods [ChatGPT Transcript].

ElephantMind. ElephantMind is an augmented reality (AR) social interaction app designed to enhance small talk by reminding users of past conversations. Alice cancels her coffee date with Bob when ElephantMind's service is down, fearing she won't remember their last chat. When Alice enthusiastically asks John about his trip to Paris, not knowing he was there for a difficult family matter, it leads to an awkward moment. Bob feels their friendship weakening as ElephantMind constantly pushes him to talk about trending political issues with Alice [ChatGPT Transcript].

Profiled Plates. Profiled Plates is a technology for high-end restaurants that offers an extraordinarily personalized dining experience by using extensive data mining and online profiling. With just a reservation name, the restaurant gathers detailed information about guests and their friends from social media, public records, and online interactions to create a bespoke meal tailored to individual preferences and backgrounds. Alice books a table for her birthday and is thrilled to find her favorite flowers on the table and a menu that excludes cilantro, a dislike she once tweeted about. However the system led to a low tip from Bob when it served a meal highly customized for a different Bob Bradley [ChatGPT Transcript].

Square One. Square One is a smart deletion feature for social media accounts which resets interests, deletes content, and unfollows people based on a simple prompt from the user about what they want removed. Jack uses Square One to unremember a volunteer project he led, only to find that potential employers now see him as less community-minded and pass him over for jobs. Jack is dismayed as he watches these projects falter under less-qualified hires [ChatGPT Transcripts A and B].

4.3 Discussion

Our experience with this methodology was positive. ChatGPT excelled at providing countless examples and as long as it produced a few useful ones, we were able to work with it in refining them. We found this methodology required less effort than traditional design (A) and produced results at least as good. One issue we found is when we used new windows and settings to limit cross-contamination between sessions, we found ChatGPT seemed to remember past interactions. The primary drawback of this method was needing a set of stems before interacting with ChatGPT. For

this reason, we decided to attempt a final methodology where we involved ChatGPT from the beginning of the design process.

5 Methodology D: LLM Design Methodology

5.1 Methodology

To be completed again. We performed this methodology but will revisit it with more time and effort.

In the previous methodology, we provided ChatGPT with stems and found success in iterating these stems into designs using ChatGPT as a collaborator. This led us to question whether it would be useful to collaborate with ChatGPT from the very beginning of the design process, taking its input on setting up the design methodology, on finding additional sticky notes or different clusterings, and ideating through different combinations of sticky notes.

- Prompts about what methodology to use
- Prompts for help in clustering notes in affinity map
- Prompts for additional notes in each cluster
- Prompts for generating stems from combination of notes
- Prompts for expanding stems
- Once we have stems, go to Methodology C

5.2 Designs

To be completed

5.3 Discussion

To be completed

6 Concluding Remarks

To be completed

Acknowledgments

J. Clark acknowledges support for this research project from the National Sciences and Engineering Research Council of Canada (NSERC) through the NSERC, Raymond Chabot Grant Thornton, and Catalaxy Industrial Research Chair in Blockchain Technologies IRCPJ/545498-2018 (https://www.nserc-crsng.gc.ca/Chairholders-TitulairesDeChaire/Chairholder-Titulaire_eng.asp?pid=1045) and a Discovery Grant RGPIN/04019-2021

References

- [1] Jeffrey Bardzell and Shaowen Bardzell. 2013. What is "critical" about critical design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3297–3306. <https://doi.org/10.1145/2470654.2466451>
- [2] Tuhin Chakrabarty, Philippe Laban, Divyansh Agarwal, Smaranda Muresan, and Chien-Sheng Wu. 2024. Art or Artifice? Large Language Models and the False Promise of Creativity. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–34. <https://doi.org/10.1145/3613904.3642731>
- [3] Minsuk Chang, Stefania Druga, Alexander J. Fiannaca, Pedro Vergani, Chinmay Kulkarni, Carrie J Cai, and Michael Terry. 2023. The Prompt Artists. In *Proceedings of the 15th Conference on Creativity and Cognition (C&C '23)*. Association for Computing Machinery, New York, NY, USA, 75–87. <https://doi.org/10.1145/3591196.3593515>
- [4] Li-Yuan Chiou, Peng-Kai Hung, Rung-Huei Liang, and Chun-Teng Wang. 2023. Designing with AI: An Exploration of Co-Ideation with Image Generators. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1941–1954. <https://doi.org/10.1145/3563657.3596001>
- [5] Anthony Dunne. 2005. *Hertzian Tales*. MIT.

- [6] Anthony Dunne and William W. Gaver. 1997. The Pillow: Artist-designers in the Digital Age. In *CHI*.
- [7] Anthony Dunne and Fiona Raby. 2000. *Design Noir: The Secret Life of Electronic Objects*. Birkhauser.
- [8] Xianzhe Fan, Zihan Wu, Chun Yu, Fenggui Rao, Weinan Shi, and Teng Tu. 2024. ContextCam: Bridging Context Awareness with Creative Human-AI Image Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–17. <https://doi.org/10.1145/3613904.3642129>
- [9] William W. Gaver, John Bowers, Andrew Boucher, Hans Gellerson, Sarah Pennington, Albrecht Schmidt, Anthony Steed, Nicholas Villars, and Brendan Walker. 2004. The drift table: designing for ludic engagement. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 885–900. <https://doi.org/10.1145/985921.985947>
- [10] Sarah Gold. 2014. *Alternet Rules*. Master's thesis. University of the Arts London.
- [11] Colton Harper, Keith Tran, and Stephen Cooper. 2024. Conceptual Metaphor Theory in Action: Insights into Student Understanding of Computing Concepts. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2024)*. Association for Computing Machinery, New York, NY, USA, 463–469. <https://doi.org/10.1145/3626252.3630812>
- [12] Buffie Holley, Dan Garcia, and Julia Bernd. 2023. Teaching Cybersecurity: Introducing the Security Mindset. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 2 (SIGCSE 2023)*. Association for Computing Machinery, New York, NY, USA, 1185. <https://doi.org/10.1145/3545947.3569637>
- [13] Ziheng Huang and Stephen MacNeil. 2023. DesignNet: a knowledge graph representation of the conceptual design space. In *Proceedings of the 15th Conference on Creativity and Cognition (C&C '23)*. Association for Computing Machinery, New York, NY, USA, 375–377. <https://doi.org/10.1145/3591196.3596614>
- [14] Shaun Lawson, Ben Kirman, Conor Linehan, Tom Feltwell, and Lisa Hopkins. 2015. Problematising Upstream Technology through Speculative Design: The Case of Quantified Cats and Dogs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2663–2672. <https://doi.org/10.1145/2702123.2702260>
- [15] Yamila Levalle. 2020. Researcher uses 3D-printed fingerprint to spoof biometric authentication. *Biometric Update* (2020). <https://www.biometricupdate.com> Presented at Def Con 2020.
- [16] Lauren Lin and Duri Long. 2023. Generative AI Futures: A Speculative Design Exploration. In *Proceedings of the 15th Conference on Creativity and Cognition (C&C '23)*. Association for Computing Machinery, New York, NY, USA, 380–383. <https://doi.org/10.1145/3591196.3596616>
- [17] Cecilia Loureiro-Koechlin, Jose-Rodrigo Cordoba-Pachon, Lynne Coventry, Soferis Demetriou, and Charles Weir. 2022. Vision: Design Fiction for Cybersecurity: Using Science Fiction to Help Software Developers Anticipate Problems. In *Proceedings of the 2022 European Symposium on Usable Security (EuroUSEC '22)*. Association for Computing Machinery, New York, NY, USA, 79–84. <https://doi.org/10.1145/3549015.3554295>
- [18] M Malpass. 2017. *Critical Design in Context: History, Theory, and Practices*. Bloomsbury.
- [19] Nick Merrill. 2020. Security Fictions: Bridging Speculative Design and Computer Security. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1727–1735. <https://doi.org/10.1145/3357236.3395451>
- [20] Arvind Narayanan and Sayash Kapoor. 2024. *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*. Princeton University Press.
- [21] Vidushi Ojha, Christopher Perdriau, Brent Lagesse, and Colleen M. Lewis. 2023. Computing Specializations: Perceptions of AI and Cybersecurity Among CS Students. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2023)*. Association for Computing Machinery, New York, NY, USA, 966–972. <https://doi.org/10.1145/3545945.3569782>
- [22] Pranathi Rayavaram, Ashwin Jagadeesha, Sashank Narain, and Claire Seungeun Lee. 2023. Designing a Visual Cryptography Curriculum for K-12 Education. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 2 (SIGCSE 2023)*. Association for Computing Machinery, New York, NY, USA, 1319. <https://doi.org/10.1145/3545947.3576266>
- [23] Téó Sanchez. 2023. Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs?. In *Proceedings of the 15th Conference on Creativity and Cognition (C&C '23)*. Association for Computing Machinery, New York, NY, USA, 43–61. <https://doi.org/10.1145/3591196.3593051>
- [24] Adam Shostack. 2014. *Threat Modeling: Designing for Security*. Wiley, Indianapolis, IN. <https://www.wiley.com/en-us/Threat+Modeling%3A+Designing+for+Security-p-9781118809990>
- [25] Jakob Tholander and Martin Jonsson. 2023. Design Ideation with AI - Sketching, Thinking and Talking with Generative Machine Learning Models. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1930–1940. <https://doi.org/10.1145/3563657.3596014>
- [26] Richmond Y. Wong, Ellen Van Wyk, and James Pierce. 2017. Real-Fictional Entanglements: Using Science Fiction and Design Fiction to Interrogate Sensing Technologies. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. Association for Computing Machinery, New York, NY, USA, 567–579. <https://doi.org/10.1145/3064663.3064682>