Author(s): Marilène Oliver, Alissa Rossi, Jonathan Cohn:

**Know Thyself as a Virtual Reality: Navigating the ethics of working creatively with personal data**

**Abstract:**

Research-creation has existed in an ethical gray-area since its introduction to the academy. In developing the *Know Thyself as Virtual Reality* project, we realized that the current standards for ethics review for university-based artists are not adequate for research-creation projects which tend to involve ethical concerns distinct from conventional research and art. This is particularly clear when a research creation project, like *KTVR* requires the use and manipulation of the personal data of others. Digital data can be useful to researchers and artists alike, but it also implies a wide variety of unique ethical concerns. While regulations and policies need to be updated for all researchers, the lack of ethical guidelines for artist-researchers compounds the risk that they face when working with personal data. In order to gain a better understanding of the implications of the growing proliferation of data, much of the focus of the KTVR project (and the content of the VR artworks) has turned to understanding emerging and evolving frameworks for the ethical use of human data in research-creation projects.

**Agenda:**

**Introduction** 3

**The Ethics of Creative Practice/Research-Creation** 5

**The Legal Framework of Research Involving Human Data Subjects** 5

**Artists, Research-Creation, and Institutional Ethics** 6

**Potential Risks and Harms** 7

**Working Towards a Solution** 8

**Conclusion** 9

Assistant Professor Marilène Oliver:

- University of Alberta, 3-98 Fine Arts Building, Edmonton, AB, Canada T6G 2C9
- marilene@ualberta.ca, www.marileneoliver.com

Alissa Rossi

- University of Alberta, 3-98 Fine Arts Building, Edmonton, AB, Canada T6G 2C9
- ahamilt1@ualberta.ca

Associate Professor Jonathan Cohn

- University of Alberta, 3-73 Humanities Centre, 11121 Saskatchewan Drive, Edmonton, AB, T6G 2H5
- cohn@ualberta.ca
Introduction

Figure 1 Screen capture of My Data Body VR app

My Data Body and Your Data Body are virtual reality (VR) artworks being created at the University of Alberta as part of the interdisciplinary research creation project Know Thyself as a Virtual Reality (KTVR). These projects bring together different forms of personal data such as medical scan, social media, biometric, banking and health data in an attempt to make visible and manipulable our many intersecting data corpuses so that in VR they can be held, inspected and dissected.

In My Data Body, the MR scanned body of artist Mariène Oliver floats prone within a cloud of textual data. Into the semi-transparent, virtual body are multiple other data corpuses downloaded from social media platforms plotted into cross sections of the body. In the horizontal plane, Oliver’s Mac terminal data is plotted into bone, Google data into muscle and Facebook data into fat. In the vertical plane are plotted data usage agreements and into the depth plane are texts from various data privacy charters. Passwords and logins flow back and forth through veins and arteries and hashtags pool in organs. To echo historical anatomical public dissections, there is an audience of facial recognition scans captured using an iPhone. The medically scanned, passive/obedient semi-transparent body becomes a data processing site that can be pulled apart and dis/organised.

Your Data Body is a partner project to My Data Body made using a combination of open source and donated datasets. This project focuses on issues of data privacy and ownership, playing on the etymology of the word data meaning ‘given’. In VR the user can pick up, move, resize, re-colour and duplicate the scanned body parts, or stack them to make Frankenstein-like figures. Audio files attached to each scan play when the user holds and manipulates it. Anonymized open source datasets are accompanied by an automated voice recounting the study data published alongside the dataset, whereas datasets ‘donated with explicit consent’ have a ‘personal story’ based on the original subject of the data.
Author(s): Oliver, M. Rossi, A. Cohn, J.

Know Thyself as a Virtual Reality: Navigating the ethics of working creatively with personal data

Figure 2 Detail of “dis-organised” My Data Body

Figure 3 Screen capture from Your Data Body VR app
The making of these artworks that consider the ethics and privacy implications of various new technologies has itself raised many practical and ethical questions about the use of personal data as artist material and subject matter; from access to data through to ownership of the data once it has been transformed into an artwork and to what extent data can be manipulated and re-presented in the name of affective, socio-political artistic research. These projects have highlighted the complexities of conducting creative research using personal data, especially with regards to the intersection of research ethics, data privacy, and rapidly emerging technology from an interdisciplinary perspective. How can artists critically use these new technologies, media, and materials in a way that respects and shows care for those who originally provided their data as well as others impacted by it? Here, we point out various ethical dilemmas that one may encounter when making art with these technologies and suggest how to improve these projects not by attempting to avoid this ethical quagmire, but rather by more fully engaging with it throughout the process.

The Problem

Research-creation exists in an ethical gray-area within university research. Except when directly involving human data subjects, research-creation and artistic work is exempt from ethical review (Tri-Council Policy Statement, Article 2.6). As KTVR involves human scan data, it benefitted from the ethics review process which prompted rigorous consideration of how consent, data privacy and potential harms would be addressed during the research project. While useful, the process highlighted significant gaps in the process. Because the Tri-Council Policy Statement (TCPS) exempts most research-creation from ethics review, most practitioners do not benefit from the process, and because it views ethics narrowly, the TCPS does not address the ethical issues particular to research-creation. Research-creation can be powerful in a way that conventional research cannot since it “has an ability to convey emotion and to engender empathy” (Foster, p.372), and that power opens it up to ethical issues that are not present in conventional research. Added to this, the TCPS is out-of-date in regards to emerging technology, preventing it from anticipating potential issues.

Ethical use of data has become more complicated with the changes in the use and development of novel technology. The rise of social media, smart devices, and the wealth of personal data they generate creates a rich source of data for researchers to mine, but also challenges individuals’ right to privacy (Baron, p.2; McKibbin, p.4; Rocher, p.2). While regulations and policies need to be updated for all researchers, the lack of ethical guidelines for artists compounds the risk that they face when working with personal data. In order to gain a better understanding of the implications of the growing proliferation of data, much of the focus of the KTVR project (and the content of the VR artworks) has turned to understanding emerging and evolving frameworks for the use of human data in research.

Legal Framework of Research Involving Human Data Subjects

Privacy and consent to use of personal information are regulated at the national level in most western countries. The European Union’s 2018 General Data Protection Regulation (GDPR) is the most stringent and laid the groundwork for regulations that followed. These regulations inform national policies on research involving human data subjects which dictates to institutions the framework for ethics boards that they must implement to grant ethical clearance to research applications for government funding.

Exemptions to privacy and consent laws are broadly consistent across policies. Research is exempt from ethical review if its purpose is journalistic, artistic, or literary, or if it is considered secondary use of anonymized data (Rossi, p.59). Whatever reason this exemption was originally added no longer makes sense as we need to

---

1 In Canada, research involving human data subjects must undergo an ethics review by the research institution’s Review Ethics Board (REB). The Tri-Council Policy Statement (TCPS) is a federal document that sets out guidelines for how REBs function, what research requires ethical review, and define research ethics and consent. Both the UK and Australia have similar structural approaches to ethical review.
recognize and account for how journalism and artistic efforts can also be ethically fraught. Given the harms that misusing personal data or making it public can cause, it is not appropriate for artists to produce work without considering its social and ethical implications. It never was. Researchers need guidance from ethics review boards. Likewise, smart devices and social media make a wealth of information publicly available (Cooper, p.162-3; Parks, p.6) and undermine the methods for protecting human subjects represented in anonymous data sets (Rocher; McKibbin). Regulations need to catch up to the social and technological developments of the twenty-first century and ethical guidelines need to be developed to help artists navigate the ethical problems they encounter while creating work.

Artists, Research-Creation, and Institutional Ethics

Artistic freedom has traditionally been the mechanism that allows artists to freely produce work in order to critically examine changes in society and stimulate dialogue with a broad audience (Bolt, p.187; Lowry p.44, 45; Noury, p.44). As academic and artist Barbara Bolt says “…the speculative and provocative nature of art...enables it to become a site of engagement for ethical debate and hence gives art the ability to illuminate the ethical issues of our times” (Bolt, p.193), but current policies use ethical guidelines that are meant for the social sciences and are ill-fitted or even antithetical to artistic practice. Ethical guidelines tailored specifically to research-creation cannot be prescriptive. The goal is not to discourage artists from engaging in important topics, nor from modifying their work in order to avoid ethical review (Bolt, p.195, 196; Lowry, p.43). Instead, guidelines would prompt artists to reflect throughout the design, research, and practice stages of their research and provide review boards with a resource to consult when assessing a project.

When research-creation entered academic institutions, the system for ethics review was not adapted to accommodate this new form of research. In Canada, a committee identified a critical need for adaptation of the TCPS to the particular needs of research-creation and provided options for ways to adapt the TCPS, including adding a new chapter devoted to research (Blackstone, (2007); Blackstone (2008); van den Hoonaard). Ultimately, the suggested chapter was not incorporated. The updated TCPS2 simply refers to external guidelines “established within the cultural sector” that “may [govern] ethical practices”, but neither verifies the existence of nor points to any specific guidelines to which artists can turn (Article 2.6). Likewise, university ethics boards are not equipped to evaluate artists’ proposals because they lack familiarity with research-creation processes (Bolt, p.195; Cox et al., p.22) and their values and interests differ from those of the artistic community (Bolt, p.196-7). While ethics boards might be concerned with preventing harm or discomfort to participants, for artists, “the risk of harm or discomfort may be one of the aims of the work” (Bolt, p.196). Susan Cox observes that “REBs in many institutions...are increasingly oriented towards minimizing risk and avoiding institutional liability,” concerns which can compromise the quality of research-creation (p.41). Even when institutions and artists seem to agree on core values that must be respected, “many artists and artistic researchers have a different understanding [of those values] to that which is embodied in [national policy statements]” since “art's beneficence lies in its capacity to create trouble” (Bolt, p.197).

In addition, institutionally based artists are, along with the larger research community, faced with the looming privacy issues of anonymized data sets and the many new techniques available for their de-anonymization. Again, there is no requirement for ethical review of research that utilizes anonymized data sets for secondary use (Rossi, p.59). While anonymization has historically been adequate in protecting the identity of data subjects, big data now compromises the protective measures it offers. The threat lies in both the wide range of information that can be inferred through location-tracking data (Baron), and the ease with which individuals in anonymous data sets can be re-identified by cross-referencing demographic information (Rocher). The inadequacy of current privacy laws and ethical guidelines fails to assist artists in making informed decisions on ethical and privacy issues and makes ethical guidelines specific to research-creation necessary.
Potential Risks and Harms

Any research, to be ethical, must meet the core values of respect for participants, merit and integrity of research, just treatment of participants, and beneficence of the work (Bolt p.19-0-1). Work that utilizes anonymous data sets risks, in the era of big data, disclosing the identity and personal information of participants, (Baron, p.1,2; McKibbin, p.3) thereby compromising the public’s respect for the artist, for the institution they are affiliated with, and potentially damaging the reputation of the research community at large (McKibbin, p.7). A breach of privacy does not have to be intentional to be damaging; it can be an unanticipated result of the digital ecosystem which has emerged in the 21st century (McKibbin, p.4). In the KTVR Ethics eSymposium, Katrina Ingram presented a short history of Google’s moves into the highly lucrative health care market. Ingram detailed how the issue with Google is not with any one of its companies or technologies, but when they are considered as an ecosystem, each collecting and feeding off consumer data. It is this data ecosystem that undermines current privacy protections, posing potential ethical issues for researchers. An example is when Marilène Oliver, one of the authors, uploaded a medical scan of her head to Facebook. The identity of the scan was not identifiable to the human eye, but Facebook’s facial recognition technology identified it as an image of her. While a person uploading their own scan data does not present an ethical challenge, the situation is radically different when the data is being uploaded by or in affiliation with an artist who is using images from a public or anonymous data set. The privacy violation may come from a visitor to the exhibition uploading images to their social media account which accidentally triggers the identification of the anonymous data subject.

Data that can be legally and ethically used in research can now breach a core tenant of ethical research because of the capabilities of machine learning, facial recognition, and big data (Hassan, p.50). The issue is complex: does Facebook have the right to identify every photo posted to its platforms? Should ethics boards recognize the potential for re-identification and indicate a need for a no-photography policy for exhibition spaces showing the work? Is there a need for software that corrupts data—such as Fawkes—so that identification is impossible and that facial recognition software will recognize as private? (Hassan, p.52) Where does responsibility lie when the breach of privacy involves multiple participants and how is a breach dealt with in a productive, transparent and positive way that moves the field forward rather than stifle research that seeks to ask difficult and contentious questions? As the RCR Toolkit suggests, “the adoption of a proactive and transparent attitude, as well as a continuous reflexive view of the research process, are key elements in preventing breaches of [responsible conduct in research]” (Noury et al, p.12).

There are also issues when considering the inconsistency of privacy protection. As Dr Remigus Nwabueze has detailed, privacy protection is not extended to the dead in all jurisdictions (p.26). Early in 2021, Microsoft received a patent for software that uses the data of deceased persons to create chat bots (Brown). Our digital corpus is already available to paying customers of data brokers (Parks); once a person has died, there is, depending on your place of residency, limited or no protection of your personal data. According to Dr Ubaka Ogbogu, it is not even clear whether an individual has ownership or control over their own data since privacy is not the same as ownership (p.33), and the TCPS notes that “Canadian law does not provide clear recognition of property rights in human biological materials” (TCPS Article 9.19). One might argue that publicly available data is not owned by anyone. When considering research-creation, how do ethics boards approach big data? If the data is publicly available or if it is anonymized, researchers are not obliged to submit to ethical review. But is the use of this data ethical? Would a project such as Your Data Body be right in using the data of deceased persons, and would the potential shock of the use of that data be justified if it provoked a public conversation around the availability and use of digital data? Another project, I Know Where Your Cat Lives, scrapes photo sharing websites for images of cats (Mundy et al.) The images are then pinned to a satellite map using the photo’s location metadata. While the images are not of people, it pushes against notions of privacy, provoking viewers to consider the implications of making their data public. The site also provides instructions on how to remove an image by changing their privacy settings, thereby satisfying both privacy requirements (using publicly available data) and consent requirements (providing a way to withdraw consent). It is when art pushes up against boundaries that it can be the most powerful, and with our increasingly complex digital life,
artists need guidance when judging when their work pushes hard enough to provoke and when it crosses ethical boundaries in unacceptable ways.

**Literature**

Currently, there exists a scoping review of the literature on the issues in research-creation (Voarino et al.) as well as two sets of guidelines developed to aid researcher-creation practitioners and ethics boards to navigate the evaluation of the ethics of a project. *Guidelines for Ethical Visual Research Methods* focuses on visual research creation methods, and the *Toolkit for RCRC*, which evolved out of the review by Voarino et al, focuses on research-creation in the Canadian and Quebecois context but is widely applicable.

The Guidelines for Ethical Visual Research Methods focuses on the ethics of research-participant relationships formed through the practice of visual research methods and provides six categories with guiding questions. The categories are confidentiality, consent, representation and audience(s), fuzzy boundaries, authorship and ownership, and minimizing harms (Cox et al., p.8). While confidentiality, minimizing harms, and consent are standard tenets to research ethics, the methodological differences between academic research and research-creation result in different concerns for each category. While digital technology makes confidentiality more difficult to guarantee in any research, some participants in research-creation projects may prefer to be identified, making anonymity undesirable and antithetical to the project’s goals.

Regarding minimizing harm, it is important to recognize the potential for discomfort and vulnerability of participants in research-creation projects. The importance is not always in avoiding these harms but in preparing participants for the possibility and in providing ways to support both participants and researchers. Research-creation can complicate obligations surrounding consent since visual presentation of the work at the dissemination stage can present unexpected consequences. While academic research views consent as a one-time event, research-creation may require a continual renewing of consent throughout the life of the work and a process for how to facilitate the withdrawal of consent.

Other categories are more unique to research-creation. Fuzzy boundaries arise when the roles of researchers and participants are less clearly defined, for instance when participants hold more expertise in the given topic than the researchers. In these cases, it is important for the roles and expectations of different actors to be well defined early in the work. When research-creation is co-created, issues of authorship and ownership can arise, from acknowledgement of participation, ownership of the final work, and the consent to use of the final work and the establishing of protocols to accommodate the withdrawal of consent. Finally, issues with representation and audience can include the integrating of expectations of participants, providing adequate guidance to audiences to address or avoid problematic readings of the work, and considering how audience members—particularly if disseminated online—may use or interact with the work in unexpected or problematic ways.

Focussing on the disconnect between academic institutions’ ethical guidelines and the nature of research-creation and building on the scoping review, the *RCRC Toolkit* aims to bring research-creation and responsible conduct of research closer together (Noury et al., p.15). They identified six categories of issues that exist in responsible conduct of research creation. In descending order of the frequency, they categorize the issues as those which arise 1. from the nature of the research-creation approach, 2. from the position which research-creation practitioners hold in relation to the work, 3. from the mismatch of academic training with the research-creation process, 4. from the funding of research-creation projects, 5. from the dissemination of research-creation knowledge, and 6. from the conflicts of authorship, commitment, and interests within research-creation production (Voarino et al, p.317-9).

Issues arising from the research-creation approach include conflicts between research-creation and academic research ethics and the difficulty of balancing the needs of participants (informed consent, minimizing harms, etc.) with the integrity of artistic expression. Issues arising from the position which research-creation practitioners hold in relation to the work when the work involves collaboration, when participants in the work have dual expertise, when the work is intended to create new spaces, provide a voice to a specific community,
or direct attention to the research-creation practitioner’s ethical position, when the work makes participants or researchers vulnerable, or how research-creation practitioners relate to their role as researcher-creator. Issues arising from the mismatch of academic training with the research-creation process stem from the integration and acceptance of research-creation into academia and the accompanying need to develop appropriate student-supervisor relationships and ethics training that meet the specific needs of research-creation. Issues in the funding of research-creation projects arise from both a lack of funding for the arts and the need for research-creators to tailor their funding applications or their work to criteria that do not fit well the process of research-creation. Issues arising from the dissemination of research-creation knowledge include the issues of disseminating the work in unorthodox ways, or in ways that do not meet the standards of academic research, as well as issues of ownership and authorship in relation to collaborative work. Issues arising from conflicts of authorship, commitment, and interests within research-creation production can stem from the needs of participants and the needs of the research-creation practitioner, the multiple roles that research-creation practitioners can hold (as student or supervisor and artist, as a member of the artistic community and a member of an academic institution), and the potential conflict between the needs of each of their roles.

Building on the six categories identified in the review by Voarino et al, the **Toolkit for RCRC** provides a checklist for researchers and ethics boards to consider when designing and evaluating research-creation projects. Overall, they suggest “the adoption of a proactive and transparent attitude, as well as a continuous reflexive view of the research process, are key elements in preventing breaches of [responsible conduct in research]” (Noury et al, p.12).

Together, these resources provide a basis of ethical guidance. However, in part due to the quickly evolving nature of digital technology, none of them adequately address evolving technology and the unanticipated impact that it could have on the work. The goal of KTVR therefore is to provide further guidance in the use of personal data in an era of big data and vast digital ecosystems in order for artists to better anticipate and resolve issues in an ethical and transparent way.

**Working Towards a Solution**

The guidelines being developed by KTVR aim to provide guidance to artists working outside of institutions who have fewer resources to draw on, and to address some of the gaps that exist in the resources available to research-creation practitioners within institutions. Specifically, these guidelines provide a summary of existing visual research ethics (as given above here), before providing an overview of privacy and consent laws in Western countries so that artists understand the complexity of the laws that govern their use of personal and medical data, and an understanding of various policies governing research involving human data subjects in order to assess the source of the data with which they are working.

The final section of the KTVR guidelines take the form of 8 self-reflexive questions that are accompanied by an explanation of terms, best practices and discussion of potential harms and fuzzy boundaries. The questions/categories are:

1. Anonymization: Is the data really anonymous? Should it be and why?
2. Provenance: Where does the data come from? Who/what has it been passed through?
3. Demographic makeup: Who does the data represent? In what ways and from which perspectives was this data being recorded?
4. Consent: Who gave permission to use the data? For what purposes and under what conditions was consent given? Are there different levels of consent depending on how the data is going to be used?
5. Access & Licensing: How can you get the data? What do you have to do to get it? Who may gain and/or suffer from your use of this data?
6. Participation and authorship: Are data subjects participants / collaborators? Are you providing space for their perspectives to become known?
7. Sustainability: Does new data really need to be collected? Can existing data be used?
8. **Audiences:** Who is your work for? How do different cultures in the project understand the data you’ve gathered and the concept of data more generally?

9. **Artificial Intelligence:** Is AI used in the project and if so what kind? Are you in control of the data? Do you understand what is happening to it?

10. **Dissemination:** How is the work being made public? Is there a risk it will be documented and shared out of context? If so, do your participants understand the complexities of this, especially with regards to facial recognition software that might allow them to be automatically tagged on social media?

Moving forward, we also intend to invite a range of diverse artists to work with the tools created as part of the KTVR project so that they too can make VR artworks using medical and personal data for a public exhibition. Throughout the project we will discuss the legal and ethical issues of working with personal data as artistic material with the artists and build a repository of reflections for future researchers and artists.

## Conclusion

Working to understand and reflect upon legal and ethical issues entangled with the KTVR project has emboldened those involved to take more risks, be more provocative and to recognise the VR artworks as spaces to actively reflect upon data ethics. As a result of Ingram’s presentation and the lively discussions with participants afterwards, (Ingram) it was decided to add more personal biometric data including dental, retinal and facial scans as well as voice (which is now being used by banks to verify identity) to *My Data Body*.

*Your Data Body* was also directly influenced by the KTVR Ethics eSymposium and Dr Nwabeze’s attention to the lack of data protection for the dead. Whereas originally the intention was for the conversational AI in the VR work to be an ‘advisor’ for users, it was decided instead to make it possible for users to converse with certain key datasets about data ownership and privacy. One of the key datasets chosen is the Visible Human Project (VHP)--a highly used, open access scan dataset from Joseph Paul Jernigan, who was convicted and executed in Texas (Pandya). Utilising role-playing strategies to create a database of conversations to train the AI, we made a character profile which includes what the VHP knows about their data, feels about their data and wonders about their data. Repeatedly pretending to be the VHP has proven to be a process that elicits considerable empathy and engagement. Although there are concerns about Microsoft’s potential use of the data of deceased persons, (Brown) the activity of creating a conversational AI based on what we imagine the VHP knows, feels and wonders has proven to be a valuable and sensitive research tool in the project. Equally important, it also allows us to engage in dialogue about the ethical use of the data of deceased persons and to question how that data will be used in the future.
References


