UNIVERSITY OF CANBERRA

Human Research Ethics Committee

Application for Approval

1. ABOUT THE PROJECT

Project Title:	Artisanal software development for text visualisation in real scenarios					
Type of Project:	◯ Teaching	◉ Research	◯ Clinical Trial	**Is your Project:**	◉ New	◯ Continuing
Proposed Start Date:	August 2014			**Anticipated Completion Date:**	May 2015	

Please note: The Committee cannot grant retrospective approval!

Is this project related to a previous application?	◯ Yes	◉ No	**If yes, original project number:**	
Does this project repeat a previous study?	◯ Yes	◉ No		
If yes, please explain why repetiton is necessary:				

1.1. ABOUT THE CHIEF INVESTIGATOR

Family Name:	Nualart	**Given Name:**	Jaume	**Title:**	Mr
Email - *Please provide a staff or student email address only:*	Jaume.Nualart@canberra.edu.au				
Faculty / University Research Centre:	Faculty of Arts & Design				
Staff ID:	s428998	and/or	**Student ID:**	u3091771	
Course of study (if a student):	PhD in Communication				

1.2. UNIVERSITY OF CANBERRA SUPERVISOR (if Chief Investigator is a student)

Family Name:	Whitelaw	**Given Name:**	Mitchell	**Title:**	Dr
Faculty / University Research Centre:	Faculty of Arts & Design				
Phone:	(02) 6201 5184	and	**Email:**	Mitchell.Whitelaw@canberra.edu.au	

1.3. QUALIFICATIONS AND EXPERIENCE OF CHIEF INVESTIGATOR

Qualifications:

2010-2012
Master of Advanced Studies (MAS) - In Spain: D.E.A.
Excellent (10 over 10) in PhD courses
Visualization tools applied to texts (media art, scientific & l egal texts).
Department of Public Law and Legal History,
School of Prevention and Integral Safety
UAB - Universitat Autònoma de Barcelona

1987-93
Licenciate (BA+MA) of Science (Chemistry)
Qualifications (7 over 10)
General Chemistry. Espec. Phisics-chemistry & electro-chemistry.
Science Faculty. Universitat Autònoma de Barcelona (UAB)

Relevant Research Experience:

2010
Lecturer / teacher
Workshop for the presentation od the CCCB archive: "The future of the archives"
Centre de Cultura Contemporània de Barcelona (CCCB)

2008
External expert for the European Commission.
The URL of the publication is: http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=2819
The PDF document: http://ftp.jrc.es/EURdoc/JRC54327.pdf
EUR 24063 EN
ISBN 978-92-79-13813-3
ISSN 1018-5593
DOI 10.2791/29872

2008-09
Software developer / Researcher
Developer of data visualization techniques for the digital archive of Arts

1.3.1. FOR STUDENTS ONLY

Degree and Training in Research:

PhD in Communication candidate
at Faculty of arts and Design. University of Canberra

2013
HDR Seminars at University of Canberra
- Ethics application
- Qualitative analysis
- Quantitative analysis

1.4. CO-INVESTIGATOR 1

Family Name:		**Given Name:**		**Title:**	
Faculty / University Research Centre:					
Name of external organisation:					
Phone:		and		**Email:**	
Qualifications:					
Relevant Experience:					

1.4.1. CO-INVESTIGATOR 2

Family Name:		**Given Name:**		**Title:**	
Faculty / University Research Centre:					
Name of external organisation:					
Phone:		and		**Email:**	
Qualifications:					
Relevant Experience:					

1.4.2. CO-INVESTIGATOR 3

Family Name:		**Given Name:**		**Title:**	
Faculty / University Research Centre:					
Name of external organisation:					
Phone:		and		**Email:**	
Qualifications:					
Relevant Experience:					

1.5. DETAILS OF ANY OTHERS INVOLVED IN THE RESEARCH (IF KNOWN)

Name(s):	
Role(s):	
Qualifications:	
Relevant Experience:	

1.6. SITE(S) WHERE THE RESEARCH WILL BE CONDUCTED

Address 1:	Level 2 - Canberra Research Laboratory. Tower A. 7 London Circuit. Canberra City ACT 2601. Australia
Address 2:	
Address 3:	
Address 4:	

1.7. FUNDING AND REVIEW

Source of funds and amount:	
RM Number (if applicable):	

Do the funding and/or commercial and intellectual property arrangements place you in a conflict of interest as a researcher? If so, please provide details below.

Details:	
Constraints on publication if any:	
Describe any peer review of the proposed research:	

2. RESEARCH DESIGN AND METHODOLOGY

This section is to provide members of the HREC with clear understanding about the need for the research and the approach adopted. Please use language that can be understood by those outside of the discipline or profession.

2.1. RATIONALE AND LITERATURE REVIEW

The purpose here is to understand how the problem being investigated fits with other research in the area (see NS 1.1(c))

In the age of data, tools to study, manage, and discover facts in the data are crucial. The growth in accessible online data is significant; and most of this data is text (Heer, 2010). This project is about the process of developing visualization and exploration tools for textual data. We think that now is the time to push for a more extended use of visual tools to manage, understand and discover facts through studying data. At the same time it is more important than ever that we defend the democratization and transparency of data.

This project is a creative production research project, ie, a practice led research. This is a feasible and fruitful research strategy that will help achieve these goals. The research project will produce both theoretical and practical assets as outputs.

The theoretical study will propose a classification of different approaches to text visualization and exploration. The aim of this classification is provide tools for a global conversation, introducing common names and definitions.

The practical side of the research is the creation of three to five software artifacts, working with individual texts and texts collections as data, and based in real scenarios including digital archives and museum exhibitions.

My personal background as free software developer lead me to use the expression “artisanal development” to describe my approach here. “Artisanal” is used in the sense of handmade, taking care as if the project were a personal project; feeling part of the project; living and assuming the project as a personal challenge. Artisanal software is also related to open source, open techniques, handmade coding, and the originality of the work. Artisanal software development refers to a hyper-customised software project, rather than a mass-produced or “industrial” approach. It also refers to an organic, and spontaneous approach based in a non-corporate style of working and influenced by creative labs, hackerspaces and free software communities(). I use this term inspired by Mark Bernstein (2004, Atzenbeck [2008]) and Christopher Schanck (Schanck 2009).

At the time of writing the initiated projects are:
Visference: this is an alternative interface for accepted papers of an academic conference. This will be applied to the 282 accepted papers in JMLR Workshop and Conference Proceedings (Volume 28 : Proceedings of The 30th International Conference on Machine Learning). This project is developed with Dr. Wray Buntine (former NICTA supervisor of my research).
Area: this is a representation of a collection of items, e.g., digital archives. Area will be applied to all published papers (over five hundred) of the Information Research journal, an open access journal edited by Prof. T.D. Wilson (http://informationr.net/ir/)Crossreads: this is an application that explores collections of texts and invites the user to read them by breaking the expected linear narrative into short pieces. the dataset will be about sixty written texts by the artist Eugeni Bonet since 1970’s. It will be presented in MACBA (Museum of Contemporary Art of Barcelona) as part of the exhibition The listening eye - Eugeni Bonet: screens, projections, writings (MACBA, 2014).

2.2. RESEARCH AIMS OR QUESTIONS

Main question:
How can we create text visualisation and exploration approaches developed in an artisanal manner and applied to real scenarios with real users?

Why text visualization and exploration approaches?

As outlined above, with access to increasing volumes of textual data, the greater the need for tools to use, manage and understand data. Visualization and exploration techniques work hand in hand with data science, advancing in the comprehension of the new universe of data evolving today.

Why real scenarios?

Another choice is to work only in real world scenarios, making the project realistic in resources and in timing. In the literature review we found a lack of text visualisation applied in real scenarios. In this research we will work with real scenarios because we intend to understand how text visualization and exploration can be applied in concrete cases, rather than just playful demos.

.

2.3. RESEARCH APPROACH, METHODS AND INSTRUMENTS

Please note that the National Statement requires additional ethical matters to be considered in particular types of research such as clinical trials, the collection of human samples, genetic testing, cellular therapy, ionising radiation, research on gametes or the creation of embryos (see NS 3.3-3.6).

General Methodology

This is a multidisciplinary practice led research project, undertaken through practice (practice as an element of the research design). This concept is an adaptation of the classification that Frayling defines in visual arts (Frayling 1993).

The methodology is based on the study and lessons learnt from a number of text visualisation tools that will be developed as part of the project. Following Linda Candy (Candy et al. 2006), the main focus of this research is to advance knowledge about practice and it seeks to learn about practice through making the artefacts. It can be considered a practice-led research rather than a practice-based one, where the contribution to knowledge is demonstrated through artefacts such as images,
music, designs, models, software.

The project cycle structure for each artifact development is:
- Description of the scenario and context
- Prototyping and production
- Tools evaluation
- Reflection as methodology to improve the results.
- Informs the next project cycle

In the literature review we studied forty-nine examples of text visualisation; we couldn't find published user evaluation for any case. For this reason, we need to create our own user evaluation, in order to be able to validate the features of the developed artifacts.

In order to address the research question we will study how users accept and use the created artifacts. For the case of Area and Visforgeo, the evaluation will focus on new features introduced through text visualisation, by comparison with existing or conventional

2.4. RESEARCH PRODUCTS

My research will produce a number of software packages that will be published with open licences and accessible to everyone.
In addiction, as mentioned below, the research will produce a theoretical study on text visualization and exploration classification. All this will be explained deeply in an exegesis.

With the evaluation experiment we will validate the usability and understanding of the created artifacts.

2.5. BENEFITS AND OTHER IMPACTS OF THE RESEARCH

In addition to the creation of software, ancillary aims include:
• Engage/encourage groups of people to work with text visualisation tools in real scenarios.
• Propose a methodology for the successful implementation of text visualisation tools.
• Help to identify cases in which text visualisation can be adopted and implemented.

3. PARTICIPANTS, RELATIONSHIPS, FEEDBACK AND CONSENT

Please note that the National Statement focuses on the principle of respect for persons. In addition to the participants recruited for the research, there is a need to consider the impact on others who may be affected by the research (see NS 1.1 (d), 1.6-1.9, 1.10, Chapter 2.1)

3.1. PLEASE DESCRIBE WHAT PARTICIPANTS WILL BE REQUIRED TO DO OR AGREE TO HAVE DONE TO THEM.

As outlined in 2.3, we can divide the participants into two groups: standard users and experts.

a) Standard users
The tools will be installed in real scenarios, so standard users will be, for each case:
- Visference: a reader of proceedings of a conference on machine learning.
- Area: a reader of an academic information science journal.

For the standard users we will conduct an anonymous qualitative online questionnaire. Initially, participants will see an introductory text explaining the test process. Here we will introduce the context of the particular artifact's data, and we will recommend that participants spend some time visiting the pre-existing (standard) interface, and then visiting the new interface that incorporates the developed artifact. This visit will be a web guided tour to help users become familiar with the tool and to understand what the interface is about and what the tool does. Then the user will have time to try the tool; recommended time will be from five to ten minutes. Finally the participant will answer some qualitative questions about the features and usability of the tool. This will take about ten minutes. The participants will be able to use the artifacts in order to answer the questions. We expect the time will vary from ten to twenty minutes. In total the user can expect to spend a minimum of twenty minutes and maximum of thirty to forty minutes doing the evaluation.

The questionnaires will be mainly multiple-choice questions, with two free text questions about features of the developed artifact.

Age and gender data will be collected from each standard user.

b) Experts
An expert for each experiment will be a person who knows about the represented data and project's domain. Experts will be invited to participate in a semi-structured interview. The interview, when possible, will be made in person. Data recorded will include a video recording of the computer screen while using the artifact, and an audio recording of the conversation. The screen capture video will enable correlation between on-screen activity and the interview content. In case the expert is not available for an in person interview

3.2. PLEASE OUTLINE THE RISK OF ANY POSSIBLE DISCOMFORT OR HARM FOR PARTICPANTS AND YOUR STRATEGIES FOR MINIMISING THIS RISK.

No risks are expected

3.3. PLEASE OUTLINE ANY POSSIBLE RISK TO YOU AS THE RESEARCHER AND YOUR STRATEGIES FOR MINIMISING THIS RISK.

NOT APPLICABLE

3.4. PLEASE OUTLINE ANY POSSIBLE RISKS TO OTHERS ARISING FROM THIS RESEARCH.

NOT APPLICABLE

3.5. PLEASE DESCRIBE HOW YOU WILL SELECT, RECRUIT AND CONTACT PARTICIPANTS.

1) Standard users:
I will publish open calls and send it to mailing lists from University of Canberra, NICTA and, potentially,communities from Europe, Africa and America. As we said, the participants in the online questionnaires may be potential visitors of each case. For the two cases - Area and Visference - we will invite people from appropriate academic communities.

Access to the questionnaires will be open and an initial agreement will be presented to visitors explaining the kind of user that the test expects.

The questionnaire's web address will not be included in the existing real scenarios, but from my personal blog and from the calls that will be sent to mailing lists and published in web sites. This will enable a more selective recruitment of participants.

2) Experts:
Experts will be invited individually using the network of colleagues for each case.

3.6. IF APPLICABLE, PLEASE DESCRIBE HOW YOU WILL OBTAIN APPROVAL TO ACCESS PARTICIPANTS.

NOT APPLICABLE

3.7. HOW MANY PARTICIPANTS WILL YOU RECRUIT? WHAT IS THE RATIONALE FOR THIS NUMBER?

1) Online questionnaires:
For the questionnaires we will try to get as many participants as possible. Usually around 50 random participants is good enough to get no more than 5% error (Niles, 2006).

2) Semi-structured interviews
For the semi-structured interviews with experts we will do at least five interviews per artifact. In this case we will follow the sometimes controversial Nielsen (Nielsen, J., & Molich, R. 1990) experiment because we want, mainly, to detect new ideas to improve the current prototypes.

References:
Nielsen, J., & Molich, R. (1990, March). Heuristic evaluation of user interfaces. In Proceedings of the SIGCHI conference on Human factors

3.8. ARE THERE LIKELY TO BE PARTICIPANTS FOR WHOM THERE ARE SPECIFIC ETHICAL CONSIDERATIONS?

- ◯ Children and young people
- ◯ People in dependent or unequal relationships
- ◯ Women who are pregnant and the human foetus
- ◯ People unable to give consent for health or other reasons
- ◯ People with a cognitive impairment, intellectual disability or mental illness
- ◯ Aboriginal and Torres Strait Islanders
- ◉ People in other countries
- ◯ People who are incarcerated
- ◯ People for whom English is a second language
- ◯ People who may be involved in illegal activities

3.8.1. IF SO, HOW ARE THE SPECIFIC ETHICAL CONSIDERATIONS BEING ADDRESSED?

NOT APPLICABLE

3.9. ARE ANY CATEGORIES OF PARTICIPANT SPECIFICALLY EXCLUDED? IF SO, PLEASE PROVIDE A RATIONALE.

NOT APPLICABLE

3.10. PLEASE DESCRIBE ANY PAYMENT OR COMPENSATION TO PARTICIPANTS.

NOT APPLICABLE

3.11. PLEASE DESCRIBE ANY PRE-EXISTING RELATIONSHIP WITH PARTICIPANTS AND ANY ETHICAL CONSIDERATIONS THAT NEED TO BE ADDRESSED AS A RESULT OF THIS RELATIONSHIP.

For the standard users any pre-existing relationship will be unknown because the online questionnaire will not collect data about the identity of the participants aside age and gender. The questionnaire will make it clear that participants may withdraw from the survey at any time without penalty.

For the experts, it is possible that I will have a pre-existing professional relationship with participants. In that case I will ask the expert to be clear in his/her opinions because this will help to improve the tool. I will not recruit specifically to avoid pre-existing relationships. If there is any reason indicating that a participant could feel pressured to participate in the experiment, the participant will not be recruited.

3.12. HOW WILL YOU OBTAIN PARTICIPANTS' CONSENT TO PARTICIPATE? HOW WILL YOU ENSURE THAT CONSENT IS VOLUNTARY? HOW WILL YOU ENSURE THEIR RIGHT TO WITHDRAW FROM THE RESEARCH WITHOUT PENALTY AND WITHOUT FEELING DISCOMFORT?

Online survey participants will indicate their consent by clicking a button on the online form. Participants will be able to withdraw from the survey at any time as the consent text will state.

Expert users will be invited formally to participate and informed about their voluntary participation. After that, in the begining of the interview they will be read a consent agreement and will indicate their acceptance by voice, to be recorded with the rest of the interview.

3.13. DO YOU INTEND TO WITHHOLD OR DISGUISE THE PURPOSE OF THE RESEARCH IN ANY WAY? IF SO, PLEASE PROVIDE REASONS.

No

3.14. WILL YOU PROVIDE ANY FEEDBACK TO PARTICIPANTS ABOUT THE RESULTS OF THE RESEARCH? IF SO, PLEASE ADVISE IN WHAT FORM FEEDBACK WILL BE PROVIDED.

Not specific feedback to participants. Only the public final exegesis publication, which will show the final evaluation of each software tool..

4. DATA

It is important that privacy and respect are the principles underlying the collection, storage and use of data. Data should be reliable, retrievable and replicable if necessary. (See Section 2 of the Australian Code for the responsible Conduct of research)

4.1. HOW WILL YOU PROTECT THE CONFIDENTIALITY AND PRIVACY OF PARTICIPANTS?

1) Standard users: participants will be completely anonymous; there will be no way to link specific responses to an individual.

2) Experts: while the identities of the expert users will be known to the researcher, these responses will be anonymised in the recorded research data.

4.2. WILL OTHERS BE ABLE TO IDENTIFY RESEARCH PARTICIPANTS FROM PUBLISHED DATA OR OTHER SOURCES (E.G. INTERPRETERS, TRANSLATORS, OBSERVERS)?

No

4.3. WILL THE DATA BE INDIVIDUALLY IDENTIFIABLE, NON-IDENTIFIABLE *(COLLECTED IN A WAY THAT NO ONE KNOWS THE NAME OF THE PARTICIPANTS)* OR RE-IDENTIFIABLE *(CODED IN SOME WAY SO THAT ONLY THE RESEARCH TEAM CAN CONNECT THE DATA TO SPECIFIC PARTICIPANTS)*?

Data will be non-identifiable.

4.4. HOW WILL YOU ENSURE THE SECURITY OF THE DATA?

The raw data generated in this evaluation experiments will be accessible only for me and eventually for my supervisor, Dr Mitchell Whitelaw. The data will be encrypted in a dedicated server owned by me that uses a secure OS, Ubuntu GNU/Linux that has all open source tools to encrypt. The data will be encrypted using openSSL (http://www.openssl.org/). Only I will know the encryption code.

Following the UC policy on research data (https://guard.canberra.edu.au/policy/policy.php?pol_id=3276) the data will be physically stored in two hard disks. One of them will be in the UC and the other will be under my protection. Both hard disks protected by appropriate security measures in lockable cabinets.

4.5. IN WHAT FORM WILL THE DATA BE STORED?

1) Questionnaires: in CSV and TXT files and then encrypted.

2) Semi-structured interviews: video files (OGG format).
The data coming from 1 and 2 will be transformed and converted to the required formats for the quantitative and qualitative analysis. The formats will be standard data interchange formats: JSON, XML, CSV, or similar.

4.6. WHERE WILL THE DATA BE STORED?

After the experiment is done, the data will be stored in two hard drives in my office at NICTA in a lockable cabinet.

Inline with university policy, an unencrypted copy of the research data will also be stored securely at the University of Canberra, for a minimum of five years after the research project is completed.

4.7. WHO WILL HAVE ACCESS TO THE RAW DATA?

Primarily, only I will have access to the raw data. However, I may need to provide access to my supervisor, Dr Mitchell Whitelaw and potential assessors upon submission of my thesis to the University of Canberra and NICTA.

4.8. DO YOU ANTICIPATE USING THE DATA IN A FUTURE PROJECT?

No

4.9. WILL THE DATA BE ARCHIVED OR DESTROYED? IF DESTROYED, PLEASE PROVIDE A DATE.

It will be archived

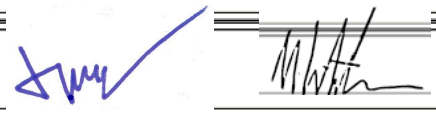
4.10. IF THE DATA WILL BE ARCHIVED, WHO WILL HAVE ACCESS TO IT AND WHAT CONDITIONS WILL BE ATTACHED?

After the completion of the research the data will be made available to enable validation and replication of the results.

5. DECLARATIONS AND SIGNATURES *(please print this page and ensure that all researchers, supervisors and the ADR sign it, please do not insert any electronic signatures)*

I/we certify that:

- All information is truthful and as complete as possible.
- I/we have had access to and read the *National Statement on Ethical Conduct in Human Research*, and that the research will be conducted in accordance with the national Statement and in accordance with the ethical arrangements of the organizations involved.
- I/we have consulted any relevant legislation and regulations, and the research will be conducted in accordance with these.
- I/we will immediately report to the HREC anything which might warrant review of the ethical approval of the proposal (NS 5.5.3).
- I/we will inform the HREC, giving reasons, if the research project is discontinued before the expected date of completion (NS 5.5.6, 5.5.8b).
- I/we will adhere to the conditions of approval stipulated by the HREC and will cooperate with HREC monitoring requirements, including the provision of annual progress reports and final reports as required.

Chief Investigator/Co-Investigator(s)/ Supervisor name:	Jaume Nualart, Mitchell Whitelaw
Signature:	
Date:	11th June 2014

I certify that:

- I am familiar with this project and endorse its undertaking;
- the resources required to undertake this project are available;
- the researchers have the skill and expertise to undertake this project appropriately or will undergo appropriate training as specified in this application.

Associate Dean Research or nominee (name):	
Signature:	
Date:	

6. ATTACHMENTS CHECKLIST

- ◉ Participant Information Form(s) and Consent Form(s)
- ◯ Recruitment Material
- ◉ Questionnaires, surveys, interview questions, test items
- ◉ Organisational and/or institutional approvals
- ◯ Relevant agreements or contracts
- ◯ Other relevant information or documentation

NOTES FOR APPLICANTS

When completing this Application Form, you should always refer to the Human Research Ethics Manual. Please go to http://www.canberra.edu.au/research/ethics/human-ethics-manual to download a copy.

INSTRUCTIONS TO APPLICANTS

- Please answer all questions. If a question is not applicable then please state "not applicable" in the relevant box.
- Please save the form when you have finished and, if approved by your supervisor and/or ADR, forward an electronic copy to the Research Ethics & Compliance Officer.
- Please print the signature page only and ensure that all researcher(s), your supervisor (if applicable) and the Associate Dean Research sign it. A scanned copy of the signature page should then be sent to the Research Ethics & Compliance Officer.
- To be considered at a particular meeting, completed application forms must be submitted to the Research Ethics and Compliance officer by the scheduled closing date listed on the Research Services Office website.
- University insurance must be arranged for each project involving clinical trials (definition: http://www.nhmrc.gov.au/health-ethics/human-research-ethics/clinical-trials]. Complete the Clinical Trials Insurance Data Collection Form and return to insurance@canberra.edu.au.
- MAC users - Please complete the form using Adobe Reader and please do not use Mac Preview. If you use Mac Preview then your entered form data may disappear.

CHECKLIST FOR SUBMISSION OF APPLICATION

Having completed the application form please check the following:

- Full details of Chief Investigator and Supervisor have been provided.
- All questions have been answered and the language used can be understood by a layperson.
- Attachments are clearly identified, numbered and included with the application.
- Does your Participant Information Form conform to UC Guidelines?
- Letters of approval from cooperating institutions, e.g. schools and government agencies, have been included (if applicable).
- Storage of data has been stated as being at the University of Canberra.
- Private addresses and phone numbers have not been used as means for participants to contact the Researcher.
- Follow-up counselling has been identified if necessary, and the counselling service identified.
- Application has been signed by Researcher/s, Supervisor and Associate Dean Research (or nominee).
- Starting date of the research postdates the meeting at which the application will be considered.
- The relevant closing date for receipt of applications has been noted.