

Understanding Europe's Fashion Data Universe

# GEN - Requirement No. 1

Deliverable number: D9.1

Version 4.0



Project Acronym: FashionBrain

Project Full Title: Understanding Europe's Fashion Data Universe

Call: H2020-ICT-2016-1

**Topic:** ICT-14-2016-2017, Big Data PPP: Cross-sectorial and

cross-lingual data integration and experimentation

Project URL: https://fashionbrain-project.eu

Deliverable type	Report (R)
Dissemination level	Confidential (CO)
Contractual Delivery Date	31 March 2017
Resubmission Delivery Date	4 February 2019
Number of pages	49, the last one being no. 43
Authors	Alessandro Checco, Jennifer Dick - USFD
Peer review	Peter Bath, Paul Reilly - USFD

# **Change Log**

Version	Date	Status	Partner	Remarks
1.0	5/06/2017	Final	USFD	Rejected 15/03/2018
2.0	20/04/2018	Resubmitted Final	USFD	Rejected 26/04/2018
3.0	04/06/2018	Resubmitted Final	USFD	Rejected 15/10/2018
3.1	18/01/2019	Revised Draft	USFD	
4.0	04/02/2019	Resubmitted Final	USFD	

# **Deliverable Description**

This deliverable describes the project ethics application process for each partner of the consortium, discusses the legal background of each country involved, and provides the approval letters of each ethics application.

#### **Abstract**

This deliverable describes the project's unified ethics application process for each partner and activity of the consortium, discusses the national legislation of each country involved, provides the approval letters of each ethics application and contains the relevant guidelines and policies given to each partner to help guide them through the process.

# **Table of Contents**

Lis	it of	Figures	V
Lis	t of	Acronyms and Abbreviations	vi
1.	Intro	oduction	1
2.	Nati	ional Legislation	2
	2.1.	Switzerland	2
		2.1.1. Discussion	4
	2.2.	Germany	5
	2.3.	UK Regulations	5
	2.4.	The Netherlands	6
3.	Ethi	cs Process	7
	3.1.	USFD Ethics Standards and Guidelines	7
	3.2.	Information Provided to the Applicants	8
	3.3.		
	3.4.	Application Procedure in Detail	11
	3.5.	Ethics Manager	12
4.	Fasł	nionBrain Ethics Applications	15
		Crowdsourcing	$15^{-1}$
	4.2.		
	4.3.		
	4.4.	Consent Forms	
			39
Α.	Guid	delines and Policies	40
			41

# List of Figures

2.1.	GDPR and FADP main differences	3
2.2.	Decision process on GDPR and FADP	4
3.1.	Online ethics application portal	Ć
3.2.	Example of peer-review online system to upload and manage consent forms and example of questions	
3.3.	Example of peer-review shown in the application form	11
3.4.	Ethics Manager review assignment system	13
3.5.	Ethics manager administration portal	13
3.6.	Ethics manager reviewers panel	14
3.7.	Ethics manager applications panel	14
4.1.	Document 68029	22
4.2	Document 68030	22

# List of Acronyms and Abbreviations

**DoA** Description of Action

**FADP** Federal Act of Data Protection

GDPR General Data Protection Regulation

**PFPDT** Préposé Fédéral à la Protection des Données et à la

Transparence (Swiss Federal Data Protection and Information

Commissioner)

UREC University Research Ethics Committee

## 1. Introduction

This deliverable describes the project ethics application process for each partner and research activity of the consortium, discusses the national legislation of each country involved, and provides the approval letters of each ethics application.

The partners involved in data collection (Fashwell, UNIFR and USFD) are located in Switzerland and UK. For this reason, the discussion is mainly focused on those countries, but all member countries have been considered in the evaluation and adherence to EU regulations. We refer to Section 2 for an analysis of partners' national legislation and procedures, and a description of the compliance of each partner to national and European laws.

The FashionBrain project has decided to follow a unified Ethics procedure (the FashionBrain Ethics Application process, based on the University of Sheffield GDPR compliant procedures) to ensure that all partners collecting and processing personal data comply with GDPR and the Modernised Convention 108 (we refer to Section 2 for an analysis of the consequences on national level). It is important to note that these processes are also in place because of broader ethical principles relating to research ethics, and adhere to the European Code of Conduct for Research Integrity.

The remainder of the deliverable is structured as follows. In Section 3, the FashionBrain Ethics Process is presented along with the Ethics Standards and Guidelines. Section 4 details the applications of the FashionBrain partners which are provided together with a report on the review process.

Finally, Appendix A provides links to relevant guidelines and policies documents.

# 2. National Legislation

The GDPR applies across the EU and its directives prevail over national law. Nonetheless, member states have the ability to introduce their own national laws based on specific derogations provided for by the GDPR. These derogations include national security, prevention and detection of crime, and also apply in certain other important situations ("opening clauses").

However, the nature of the data collection and data processing in FashionBrain does not fall into these categories, and thus, a unified GDPR procedure can be implemented for all EU consortium members<sup>1</sup>.

A separate distinction is needed for Switzerland, where the national act on data protection has significant differences with the GDPR. However, in the topics relevant to FashionBrain, GDPR policies are stricter (in some cases) than the Swiss data protection law, and the decision to follow the FashionBrain Ethics Process is the most appropriate. We refer to the rest of this section for a detailed analysis of this process.

As stated above, the FashionBrain project has decided to follow a unified Ethics procedure (the FashionBrain Ethics Application process, based on the University of Sheffield GDPR compliant procedures) to ensure that all partners collecting and processing personal data comply with GDPR and the Modernised Convention 108 as well as broader ethical principles relating to research ethics and the European Code of Conduct for Research Integrity, as explained in more detail in Section 3.

Regardless of this analysis, all participants will conform to and will process any personal data collected in the course of the project in compliance with (i) the GDPR and (ii) additional local law requirements.

A summary of the ethics procedures and national legislation on data protection of the consortium members follows.

#### 2.1. Switzerland: UNIFR and Fashwell

UNIFR and Fashwell are based in Switzerland, and are thus subject to the New Swiss Federal Act on Data Protection. Both partners are involved in data collection and processing that require ethics approval.

Partner UNIFR is also subject to their own internal policies, that are built around the Swiss Federal law:https://www.edoeb.admin.ch/edoeb/fr/home/protectio

<sup>&</sup>lt;sup>1</sup>Should any future data collection fall in one of these categories, then the matter will be raised with the FashionBrain Ethics Committee for consultation.



n-des-donnees/Internet\_und\_Computer/services-en-ligne/medias-sociaux/monitoring-des-medias-sociaux-et-protection-des-donnees.html.

More specifically, the institution follows the recommendations regarding social networks (PFPDT recommendation): https://www.edoeb.admin.ch/edoeb/fr/home/protection-des-donnees/dokumentation/revisions-de-la-loi-federale-sur-la-protection-des-donnees-lpd-.html.

Main criteria		GDPR	E-FADP
Category	Specifications		
General rules	Geographical range	Entities in the EU and/or entities processing data of natural persons within the EU	Entities based in Switzerland
	Fines	Up to EUR 20 million or 4% of sales (the higher of the two amounts constitutes the maximum penalty)	Up to CHF 250,000 (individuals)
Principles for p personal data	rocessing	The processing of personal data is generally prohibited, unless there is a legal basis	The processing of personal data is generally allowed, unless the privacy of an affected person is violated
		The following processing principles should be considered for all personal data being processed:  1. Purpose 2. Data minimisation 3. Accuracy 4. Storage limitation 5. Integrity and confidentiality 6. Accountability	The following processing principles should be considered for all personal data being processed:  1. Purpose 2. Data minimisation 3. Accuracy 4. Storage limitation 5. Integrity and confidentiality
Directory of processing activiti		Management of a data inventory for the GDPR and E-F	ADP
Data breach	Deadline	Within 72 hours	"ASAP"
	Receiver	Supervisory authority and, under certain conditions, the data subjects	Swiss supervisory authority and, upon request, the persons concerned
Privacy rights		A data subject has the right to transfer personal data	There is no right of transferability
		<ol> <li>Right to rectify inaccurate personal data</li> <li>Right to notification</li> <li>Right to restrict processing</li> <li>Right to be forgotten – erasure</li> <li>Right to object</li> <li>Right to revoke consent</li> <li>Right not to be the subject of exclusively automated processing</li> </ol>	Only right to information (the remaining rights are not explicitly formulated in the E-FADP, but are otherwise regulated and enforceable by the Swiss legal system)

Figure 2.1: GDPR and FADP main differences.

As shown in Figure 2.1 (source Christian Wyss: https://www.inblf.com/wp-content/uploads/2018/04/GDPR-Switzerland.pdf), there are some differences between GDPR and FADP. The main differences that are relevant to this this project are:

- The GDPR provides stricter requirements regarding the consent of data subjects, the right to object, the right to erasure, and the right to restriction of processing.
- The GDPR provides for more rigid requirements regard data breach notifications.
- The FADP does not include a right of data portability.
- The GDPR's threshold to notify data protection authorities in the case of data breaches seems lower than the threshold under the FADP.
- There is no mandatory data protection officer under the FADP.

Partner Fashwell has customers in Europe for which, at least in principle, the data processing could concern the offering of services. For this reason, as shown in Figure 2.2 (source www.pwc.ch), GDPR is applicable.

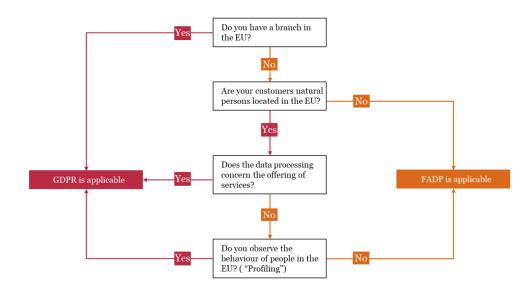


Figure 2.2: Decision process on GDPR and FADP.

#### 2.1.1. Discussion

While the potential of entering into categories where the differences between GDPR and FADP become relevant for this project are small, for the reasons stated above, it has been decided that Fashwell and Friburg be required to follow the more restrictive FashionBrain Ethics Application Policy, that guarantees compliance with GDPR and broader ethical principles relating to research ethics in adherence with the European Code of Conduct for Research Integrity. We refer to Section 3 for more details on the application process.

## 2.2. Germany: Beuth and Zalando

As a legal body owed by the state of Berlin, Beuth (and similarly Zalando) follows the very strict data privacy rules in Germany and of the state of Berlin.

Beuth guidelines for research project are explained in the following: http://www.beuth-hochschule.de/571/, and details on data privacy are available at: http://www.beuth-hochschule.de/datenschutz.

Moreover, Beuth follows §35 of the Berliner Datenschutzgesetz (state law) providing rules for dealing with data in research projects for public bodies in the state of Berlin. http://gesetze.berlin.de/jportal/;jsessionid=DD604CFE33BF8597A 0FD62DB75E60BE3.jp23?quelle=jlink&query=DSG%20BE&psml=bsbeprod.psml&m ax=true&aiz=true#jlr-DSGBE2018pP35.

Analogously, §27 of the Bundesdatenschutzgsetz (Federal law) regulates how to deal with data privacy in research projects: https://dejure.org/gesetze/BDSG/27.html.<sup>2</sup>

#### **GDPR** implementation in Germany

The German legislator has made extensive use of opening clauses and introduced a number of provisions<sup>3</sup>. However none of these provisions are relevant to FashionBrain. Should any future data collection fall in one of these categories, then the matter will be raised with the FashionBrain Ethics Committee for consultation.

#### 2.3. UK: USFD

Having received Royal Assent on May 23, 2018, the UK Data Protection Bill is now an Act of Parliament.

The Data Protection Act 2018 implements the GDPR and replaces the UK Data Protection Act 1998.

Notable provisions that make use of the ability of Member States to implement different measures from the GDPR are as follows:

- Section 9 of the Act sets out the age of consent in relation to information society services at 13 years old, instead of 16 years old.
- Exemptions from certain rights and obligations set out in the GDPR when it comes to certain criminal and immigration matters, for example, as well as for reasons of freedom of expression and information (e.g., for journalistic, academic, artistic, literary purposes), among a number of other diverse areas.

 $<sup>^2</sup>$ Since all these documents tackle state and Federal law, no translation exists.

<sup>&</sup>lt;sup>3</sup>as detailed here: https://www.twobirds.com/en/news/articles/2017/germany/germany-is-the-first-eu-member-state-to-enact-new-data-protection-act-to-align-with-the-gdpr

However, these differences do not affect this project as the research undertaken does not fall into these categories. Should any future data collection fall in one of these categories, then the matter will be raised with the FashionBrain Ethics Committee for consultation. For these reasons, we consider the UK regulation fully compliant with the GDPR as well as the European Code of Conduct for Research Integrity, and have built the FashionBrain Ethics Process on top of the existing University of Sheffield GDPR compliant procedures, as explained in detail in Section 3. Further information is available at: https://www.insideprivacy.com/international/united-kingdom/the-uk-adopts-data-protection-act-2018/.

#### 2.4. The Netherlands: MDBS

On 22 May 2018, the GDPR Implementation Act aligned the Dutch Data Protection regulation with the GDPR. There are minor differences between the new Dutch Data Protection Authority and the GDPR:

- The age of consent. As allowed by GDPR, different countries can have different age of consent. The Dutch regulation keeps it at 16.
- The role and independence of the Dutch Data Protection Authority (it has given additional powers).
- Regulation on penalties (higher in some cases).
- Some special cases like social security and public interest related data collection are clarified in the dutch law.

Again, these differences do not affect this project because it is does not fall into any of these categories. Should any future data collection fall in one of these categories, then the matter will be raised with the FashionBrain Ethics Committee for consultation. For more information, refer to: https://blogdroiteuropeen.files.wordpress.com/2018/06/paul-1.pdf.

### 3. FashionBrain Ethics Process

After consultation with the FashionBrain Ethics Committee and with the University of Sheffield Research Ethics Committee (UREC), in particular with the Chair, Professor Peter Bath, and the Senate Faculty Representative Dr. Jo Bates, it has been decided that the FashionBrain project will follow a unique Ethics application procedure (see Section 2 for further details on the rationale of this decision) based on the University of Sheffield Research Ethics Process and additionally, having the experts of the FashionBrain Ethics Committee serve as reviewers (as detailed in Deliverable D9.2).

#### 3.1. USFD Ethics Standards and Guidelines

The Sheffield UREC defined the Policy on Good Research and Innovation Practices<sup>1</sup>, that together with the Ethics Policy Governing Research Involving Human Participants Personal Data and Human Tissue<sup>2</sup> guarantees compliance with the GDPR and the The European Code of Conduct for Research Integrity.

The UREC approach to research ethics requires that all research involving human participants, personal data, or human tissue should be reviewed, and research ethics approval obtained, before data gathering commences. This approach applies to anyone who conduct, or contribute to, research activities involving human participants, personal data or human tissue, whether these take place within or outside University premises and facilities.

Researchers have a duty of care towards all individuals whom their research may affect, not just those who are directly involved as participants; the potential for harm or distress to any such individuals should be considered at the outset, and appropriate steps taken to mitigate this risk where necessary. A further detailed discussion of safety and well-being can be found in Research Ethics Policy Note no. 3 (see Appendix).

The Ethics Review Procedure aims to achieve an appropriate balance between carrying out the ethical review of projects in a sufficiently rigorous way to effectively protect the welfare, dignity and rights of human participants, whilst also being risk-aware, flexible and as user-friendly as possible in order to minimise project disruption.

The procedure is based on the following guiding standards:

<sup>2</sup>https://www.sheffield.ac.uk/polopoly\_fs/1.755691!/file/Ethics\_Policy\_Senate\_App roved.pdf



<sup>1</sup>https://www.sheffield.ac.uk/polopoly\_fs/1.671066!/file/GRIPPolicy.pdf

**Quality:** competent and consistent decision-making by ethics reviewers should be enabled and encouraged.

**Effectiveness:** the dignity, rights, safety and well-being of participants and data controllers must be protected.

**Ease of application:** the procedure is designed to be as simple and prompt as possible, while maintaining high standards.

**Efficiency:** on average, the system should provide a decision on an ethics application within 10 working days.

**Independence:** ethics reviewers must not have any conflict of interest with respect to an application they review.

**Proportionality:** the detail and depth of the ethics review of any particular project should be in proportion to the estimated level of risk posed to prospective participants. This is not a straightforward matter; where possible the applicants should take into account potential participants' likely perceptions of risk.

**Transparency:** applicants should receive sufficiently detailed, critical and constructive feedback from reviewers to explain the decision made; this should also be able to satisfy the requirements of external scrutiny, if ever required.

Although ethics approval is required before any data collection involving human participants commences, applicants are expected to consider the ethical implications of the project at all stages as even a well thought-out project may come across unexpected ethical challenges after approval has been obtained. Applicants should constantly reflect on the ethics of their project. If changes are made to the project after approval has been obtained, it may be necessary to obtain re-approval in certain circumstances.

Research ethics applications, supporting documents and review decisions will be automatically retained within the online Ethics Application System and may be used for audit purposes. The Ethics Manager is responsible of keeping this record (see Section 3.5 for more information).

# 3.2. Information Provided to the Applicants

Ethics applications (as detailed in Section 4) have been prepared by the research activity leaders who describe the way in which activities to be performed are compliant to ethical standards. The applicants have been provided with relevant documentation, in particular:

- All activity leaders have been provided with the USFD Ethics Guidelines.
- All activity leaders have been provided with The USFD Research Ethics Policy Notes no. 1-4, and 6, regarding respectively:
- – Defining human research participants and personal data



- Principles of consent
- Participant and researcher safety and well-being
- Principles of anonymity, confidentiality and data protection
- Research involving vulnerable people
- Activity leaders having activities related to social media have been provided with a document on Research Involving Social Media Data.
- Activity leaders have been provided with the Application Guide and links to the whole USFD UREC Policies.

We refer to the appendix for links to these documents.

Moreover, each applicant has been provided with a useful flowchart to streamline the decision process and a guide to the online application system.

Refer to the Appendix for the links to these documents.

## 3.3. Application Process

#### **Submission**

The applicant has to log-in to the Ethics Application System to start an ethics application. They will be presented with a series of online forms, where they will have to answer a series of questions (as detailed in Section 3.4) and they will have to provide accompanying documentation. Figure 3.1 shows the online portal used

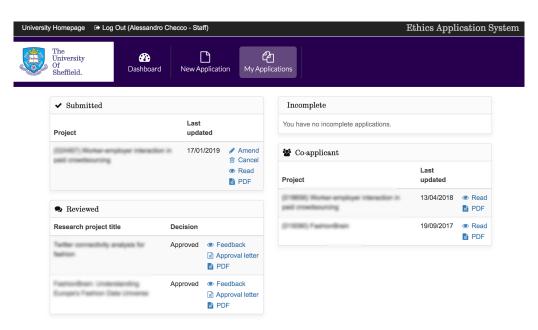
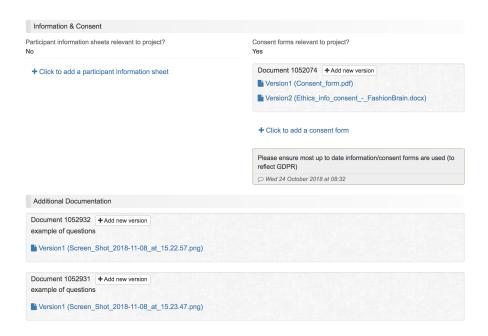


Figure 3.1: Online ethics application portal.



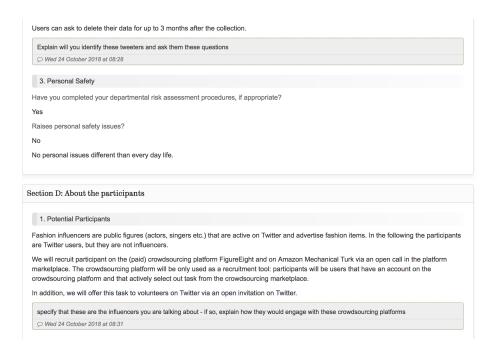
by the applicants. In Figure 3.2, an example of the consent form upload system is shown which can also keep track of the revisions to such documents.



**Figure 3.2:** Example of peer-review online system to upload and manage consent forms and example of questions.

#### Revisions

After all of the reviewers have completed their review, the applicant will be able to see the comments in the portal, in anonymous form, as shown in Figure 3.3. In the instance where the revisions are considered mandatory, they will have to provide a revision, which will go back to the reviewers for evaluation. This procedure will go on until no further mandatory revisions are required.



**Figure 3.3:** Example of peer-review shown in the application form.

## 3.4. Application Procedure in Detail

The following details the series of questions an applicant has to answer and documentation they have to provide in the online form.

- 1. Proposed project duration: the applicant has to indicate the starting and ending date of the data collection and processing.
- 2. Suitability: the applicant has to answer to the following questions:
  - Involve health research?
  - Human-interventional study?
  - Likely to lead to publication in a peer-reviewed journal?
  - Involves human tissue?
  - Clinical trial?
  - Social care research?
  - Involves adults who lack the capacity to consent?
  - Involves research on groups that are on the list of 'Proscribed terrorist groups or organisations?
- 3. Indicators of risk: the applicant has to answer to the following questions:
  - Involves potentially vulnerable participants?
  - Involves potentially highly sensitive topics?
- 4. Summary of research: the applicant has to provide in detail the following information

- Aims & Objectives of the project: what is the goal of the project?
- Methodology: which methodology of data collection and processing will be used?
- Personal Safety: are there personal safety issues for the researcher?
- 5. About the participants: the applicant should provide the following information about the participants:
  - Potential Participants: who are they?
  - Recruiting Potential Participants: what is the process of recruiting participants?
  - Consent: Will informed consent be obtained from the participants?
  - Payment: Will financial/in kind payments be offered to participants?
  - Potential Harm to Participants: What is the potential for physical and/or psychological harm/distress to the participants? How will this be managed to ensure appropriate protection and well-being of the participants?
- 6. About the data
  - Data Processing: Will you be processing (i.e. collecting, recording, storing, or otherwise using) personal data as part of this project?
  - Please outline how your data will be managed and stored securely, in line with good practice and relevant funder requirements.
- 7. Supporting documentation: Consent forms relevant to project? Participant information sheets relevant to project? Additional Documentation? The applicant can upload here these documents.

# 3.5. Ethics Manager

Kathryn Mackellar and Daniel Rose are the Ethics Manager(s). They have the role of assigning reviewers and manage the review process: set up a lead reviewer, review deadlines etc. The reviewers are selected from the FashionBrain Ethics committee (in accordance with their relevant expertise) and from the University of Sheffield, if it is considered necessary, as shown in Figure 3.4.

For each application of the FashionBrain project, at least two members of the FashionBrain Ethics Committee have served as reviewers, and every application has been reviewed by at least three reviewers.

When all the mandatory revisions required by the reviewers have been completed satisfactory, the ethics manager will collect the documentation and notify the applicant.

The management portal functionalities are shown in Figures 3.5, 3.6, and 3.7: the administrator can communicate with the reviewers and the applicants, keep track

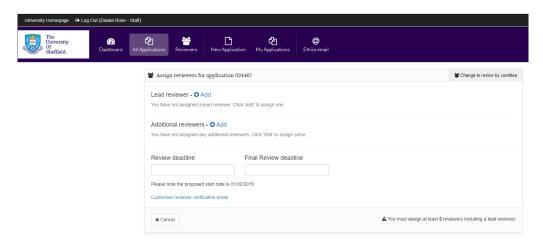


Figure 3.4: Ethics Manager review assignment system.

of the status of all reviews and assign additional reviewers when needed.

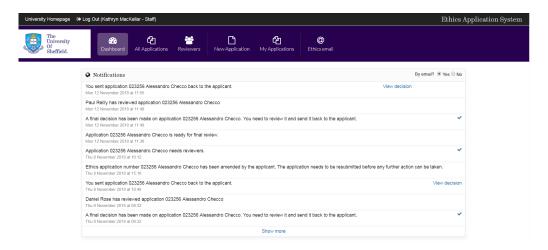


Figure 3.5: Ethics manager administration portal.

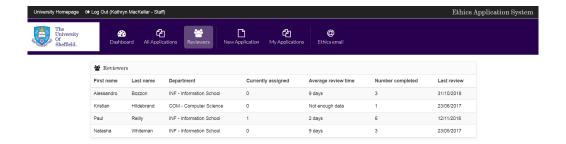


Figure 3.6: Ethics manager reviewers panel.

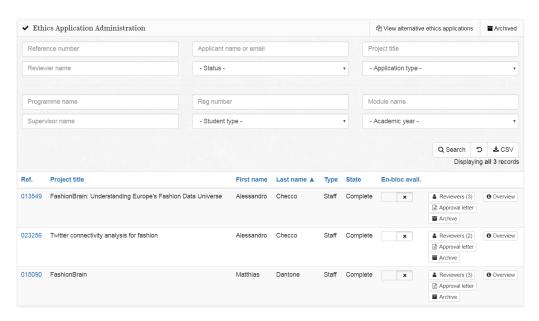


Figure 3.7: Ethics manager applications panel.

# 4. FashionBrain Ethics Applications

As explained in Deliverable D9.2, the Ethics Committee, together with the members of the consortium, agreed on the need to submit the following ethics applications, that mostly involve USFD, UNIFR, and Fashwell.

Separate participant information sheets were not used for each activity application as the majority of the information was already included in the application form itself.

## 4.1. Crowdsourcing

USFD is the lead applicant for all activities involving crowdsourcing operations within the project, that will also involve Zalando and UNIFR. This application covers all the operations in the project in which there is data collection employing crowd workers (Crowd-based Fashion Data Sampling, Curation and Cleansing).

Application Activity	Crowdsourcing
Activity Leader	Alessandro Checco - USFD
Draft Date	27/03/2017
Application Description	See attached application and approval letter (below)
Ethics Reviewers	Natasha Whiteman, Alessandro Bozzon, and Paul Reilly
Required Actions	• Added further details about the experimental tasks
	including participant instructions; types of documents;
	types of images or tweets shown to participants; and
	types of questions asked
	• Added further details regarding potential participants
	• Added other sites involved in project (aside from
	Amazon Mechanical Turk)
	Added details regarding participant payment
	• Added information regarding data confidentiality and
	storage measures
	• Modified consent form to reflect types of data shared
	within project
	• Added participant data withdrawal process

The application and the approval letter for this action is attached as follows. Note that documents 036156 and 036157 are identical, so 036156 has been discarded. Information on the participants were included in the application form.



# Ethics Application 013549

Section A: Applicant details
Created: Mon 27 March 2017 at 16:45
First name: Alessandro
Last name: Checco
Email: a.checco@sheffield.ac.uk
Generic research application: No Last updated: 23/05/2017
Department: Information School
Date application started: Mon 27 March 2017 at 16:45
Applying as: Staff member
Research project title: FashionBrain: Understanding Europe's Fashion Data Universe

# Section B: Basic information

1. Co-Applicants(s)

None

2: Proposed project duration

Proposed start date:

Thu 1 June 2017

Proposed end date:



Tue 31 December 2019 3: URMS number (where applicable) URMS number 147955 4: Suitability Takes place outside UK? Involves NHS? No Healthcare research? ESRC funded? Involves adults who lack the capacity to consent? Led by another UK institution? Involves human tissue? No Clinical trial? No Social care research? 5: Vulnerabilities Involves potentially vulnerable participants? Involves potentially highly sensitive topics? No Section C: Summary of research 1. Aims & Objectives The role of Sheffield University in the FashionBrain project is in the area of Crowdsourcing and Human Computation applied to Fashion data. The goals are to improve the quality of Human

Computation applications by developing better quality and less expensive crowdsourcing techniques; to develop novel crowdsourcing quality assurance techniques; to understand dimensions like bias and subjectivity and to deal with it by developing priming and training approaches for crowdsourcing.

#### 2. Methodology

We will run crowdsourcing experiments where participants will be asked to perform simple tasks like judging the relevance of a document or classifying an image using an online experiment available through either a crowdsourcing platform or via a standard web browser. Their judgements, actions and time taken will be recorded using logging software. Participants may also be asked to complete a questionnaire relating to the task and their crowd work experiences. The kind of documents, images or tweets shown will have high variability: from rating of items in term of aesthetics, to classification of taxonomies of items. The documents will not contain topics of sensitive nature.

A task will be placed on a popular crowdsourcing site (e.g. Amazon Mechanical Turk), where users can choose to accept the task if they wish. When the participant accepts the task (or accesses the system web page), they will be asked to read the Information Sheet and Consent Form and indicate their agreement and this is the first page of the system.

Then they may be presented with an entry questionnaire to gather demographic information such as age, gender, education/profession, proficiency in English. Next, the participant will be presented with a series of micro-tasks, where they will need to complete a web form to, for example, judge the relevance of a document (text, image or video) to a topic. Instructions on how to complete the task will be presented first also including examples of correctly completed tasks, followed by a data item (e.g., text, image) which needs to be analysed. After each task, participants may be required to complete a questionnaire related to the task workload and their familiarity with the topic.

As the participant uses the online system, behaviour will be recorded using logging software to capture mouse clicks, scrolling, time taken, etc. The logs will be analysed quantitatively to examine how the factors identified are important to the work process.

An example of a micro-task follows:

#### **INSTRUCTIONS:**

Please label the follow piece of clothing according to their usage (wedding dress, casual etc.)

[IMAGE]

[INPUT TEXT]

#### 3. Personal Safety

Raises personal safety issues? No

Pesonal safety management

- not entered -

#### Section D: About the participants

#### 1. Potential Participants



Any adult (anyone aged 18 years or over) who volunteers to complete the task on the crowdsourcing platform.

#### 2. Recruiting Potential Participants

Via submitting tasks/jobs to two popular crowdsourcing platforms: Amazon Mechanical Turk and Crowdflower.

#### 2.1 Advertising methods

Will the study be advertised using the volunteer lists for staff or students maintained by CiCS? No - not entered -

#### 3. Consent

Will informed consent be obtained from the participants? (i.e. the proposed process) Yes

The Information Sheet and Consent Form will be presented to participants as the part of the experiment. Before starting the experiment itself, participants will be asked to confirm that they have read and understood this document. They will also be invited to ask any further questions about the study. The expected completion time of a micro-task will be estimated by the researcher and via a pilot experiment. This information will allow to compute a per-task payment equivalent to the minimum hourly wage. The researcher will in fact hire the crowdsourcing platform service, which will in turn hire the workers. This will ensure that the anonymity of the worker is preserved when they are paid. If some worker will complete the task in a longer than expected time, the crowdsourcing platform provides a service to compensate such cases with additional payments.

#### 4. Payment

Will financial/in kind payments be offered to participants? Yes

Payment will be offered and will depend on the time the tasks take to complete: the task maximum completion time will be estimated (using previous tasks statistics and researchers dry run) and used to compute the per task payment so that the hourly wage will at least equal to UK minimum wage.

This payment will not contravene the anonymity granted to participants because the crowdsourcing platform acts as financially intermediary, so the researcher will not need to know any identity detail to be able to pay the worker.

#### 5. Potential Harm to Participants

What is the potential for physical and/or psychological harm/distress to the participants?

The risks of participating are the same as those experienced in everyday life.

How will this be managed to ensure appropriate protection and well-being of the participants?

The crowdsourcing studies will also not involve psychological distressing search tasks. The risks of participating are the same as those experienced in everyday life.

#### Section E: About the data

#### 1. Data Confidentiality Measures

No personal identifying information will be collected; the online experiment will be anonymous. Anonymous user identifiers will be used to integrate data together.

If a participant requests to be withdrawn from the project, they will have to provide their microtask id and their worker id: this information combination is only known by the worker and does not jeopardize their anonymity.

#### 2. Data Storage

The data will be stored securely on password protected computers at The University of Sheffield. A copy will also be kept on the Sheffield RA's university-issued laptop for analysis purposes in encrypted form. The data will be backed up using an external hard drive which will be kept in a locked drawer in the IR lab at Sheffield. The PI, the RA, and PGR students will have access to the data during the project.

#### Section F: Supporting documentation

#### Information & Consent

Participant information sheets relevant to project?

Yes

#### **Participant Information Sheets**

 Ethics\_info\_consent\_fashionbrain.docx (Document 036156)

Consent forms relevant to project?

Yes

#### **Consent Forms**

 Ethics\_info\_consent\_fashionbrain.docx (Document 036157)

#### **Additional Documentation**

None

#### **External Documentation**



Offical notes		
- not entered -		
Section G: Declaration		
Signed by: Alessandro Checco Date signed: Mon 15 May 2017 at 01:42		
,		

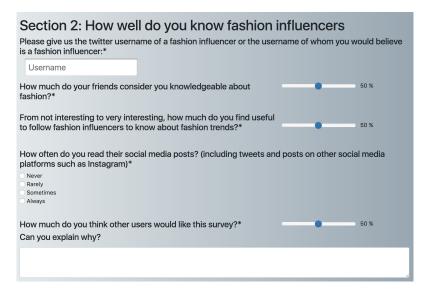


Figure 4.1: Document 68029.



Figure 4.2: Document 68030.



Downloaded: 17/09/2018 Approved: 23/05/2017

Alessandro Checco FashionBrain Project

Dear Alessandro

**PROJECT TITLE:** FashionBrain: Understanding Europe's Fashion Data Universe **APPLICATION:** Reference Number 013549

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 23/05/2017 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 013549 (dated 15/05/2017).
- Participant information sheet 1028999 version 2 (15/05/2017).
- Participant consent form 1029000 version 2 (15/05/2017).

If during the course of the project you need to <u>deviate significantly from the above-approved documentation</u> please inform me since written approval will be required.

Yours sincerely

Kathryn MacKellar Ethics Administrator FashionBrain Project



## 4.2. Social Media

Fashwell is leading activities that involve image recognition from social media.

Application Activity	Collect Social Media Posts
Activity Leader	Matthias Dantone - Fashwell
Draft Date	13/06/2017
Application Description	See attached application and approval letter (below)
Ethics Reviewers	Alessandro Bozzon, Paul Reilly and Kristian Hildebrand
Required Actions	Provided clarification on methodology to evaluate
	collected data
	• Added further details regarding identifying and finding
	potential participants/fashion bloggers
	• Added detail to justify decision not to seek informed
	consent from project blogs etc.
	• Explained handling of public comments / data on
	public blogs
	• Added details regarding participant data presentation
	and anonymisation
	• Added information regarding data confidentiality and
	storage measures

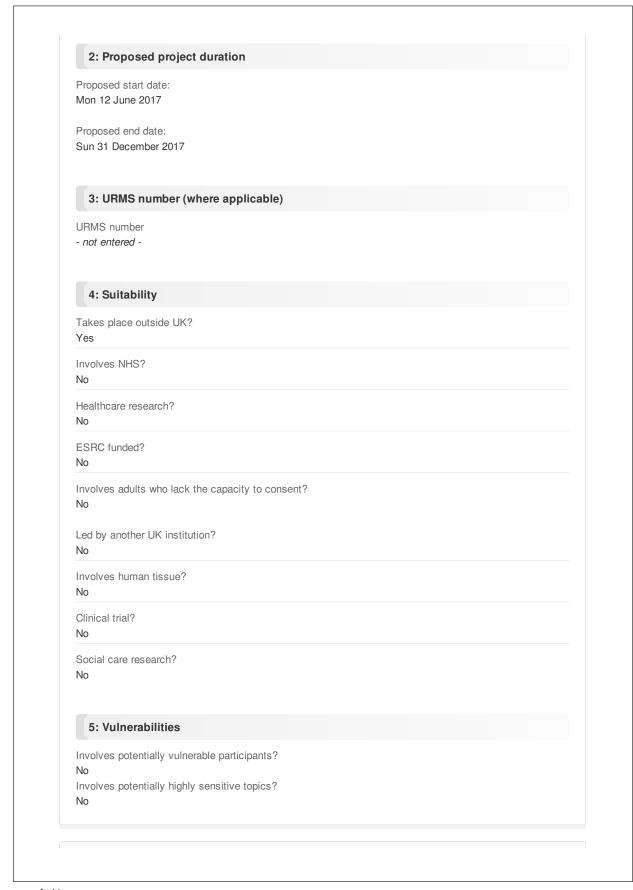
The application and the approval letter for this action is attached as follows. Note that document 040221 was a summary of the FashionBrain proposal (DoA) that the reviewers already previously received.



# Ethics Application 015090

Section A: Applicant details	
Created: Mon 12 June 2017 at 14:45	
First name: Matthias	
Last name: Dantone	
Email: m.dantone@sheffield.ac.uk	
Generic research application: No Last updated: 19/09/2017	
Department: Information School	
Date application started: Mon 12 June 2017 at 14:45	
Applying as: Staff member	
Research project title: FashionBrain	

# 1. Co-Applicants(s) Name Email Gianluca Demartini g.demartini@sheffield.ac.uk Alessandro Checco a.checco@sheffield.ac.uk Kathryn MacKellar k.mackellar@sheffield.ac.uk



#### Section C: Summary of research

#### 1. Aims & Objectives

The overall FashionBrain project aims at combining data from different sources to support different fashion industry players by predicting upcoming fashion trends from social media as well as by providing personalized recommendations and advanced fashion item search to customers.

In this specific research we are planing to:

- a) collect public available social media posts from the world biggest fashion blogger.
- b) analysing these posts and match the fashion product the person is wearing on that picture to specific database of products
- c) perform analysis on these kind of matches

Summary: "We download social media images from fashion blogger and try to figure out what they wear".

Similar research has been done by:

- Yamaguchi et al. 2013 CVPR http://ieeexplore.ieee.org/abstract/document/6248101/? reload=true

#### 2. Methodology

The methodology in which we will evaluate data collected during this project, can be divided into three areas:

We plan to experiment with various techniques for object detection and trend prediction. The employees at Fashwell will use benchmarks to evaluate the success of different techniques. The creation of the benchmarks and the basis for all of our algorithms will involve human subjects to label and create the ground truth for all of our data sets. We have a team that is responsible for data collection, and then our team of engineers has built all of the technology to extract information from images. We have a trained algorithm that can 1) detect objects in images 2) match them to similar objects and 3) draw in other similar products. From this technology we can understand the products in images and different attributes of them, in addition to product information and draw conculsions of what are trending objects and colors.

#### 3. Personal Safety

Raises personal safety issues? No

Pesonal safety management

- not entered -

#### Section D: About the participants

#### 1. Potential Participants

We do not plan to contact any fashion bloggers. Our interaction will be only with their publically available data points that are available in Social Media and blogs which we will use to generate conclusions about trends based on the information we collect. We will be able to identify strong

fashion bloggers based on the number of followers and publically available site visit information. The reason we will analyze their traffic is so we can ensure the bloggers we start to analyze further are significant influencers in the Fashion space. We plan to select 100 significant and diverse Fashion bloggers to use in our data set. We will use their publically available information, as these are public figures who product public content daily for public use. The only information we will use in the study is data we can derive from an image (the products and their attributes and the setting for the image) and then we will use the publically available text data points like captions and hashtags. After we collect data from this context, it will not be used again on a singled out basis. All of the data will be compiled into conclusions about fashion products and what is popular.

#### 2. Recruiting Potential Participants

We don't contact the participants. We download their last Instagram Posts via the Instagram API.

#### 2.1 Advertising methods

Will the study be advertised using the volunteer lists for staff or students maintained by CiCS? No

- not entered -

#### 3. Consent

Will informed consent be obtained from the participants? (i.e. the proposed process) No

Influencer and blogger are persons of public interest and everything they post publicly is done in that way to increase their reach and support their business. In this project we are only interested in the content that is made to be shared publicly like their images, post captions and hashtags. The blogger will be anonymized by not using the username, and we will only use the content to draw conclusions for data sets. The information we will be collecting is the public content published by the influencer and will not reach out directly to influencers to obtain the content as it is publically available. We do not collect or analyze any other publicly available content like comments or likes or information that their viewers or users post as comments on the images or on their blog. Therefore, we do not find it necessary to contact any participant for consent.

#### 4. Payment

Will financial/in kind payments be offered to participants? No

- not entered -

#### 5. Potential Harm to Participants

What is the potential for physical and/or psychological harm/distress to the participants?

There is no potential for that.

How will this be managed to ensure appropriate protection and well-being of the participants? See above.



#### Section E: About the data

#### 1. Data Confidentiality Measures

In this research we are collecting publicly available posts from influencers. For this we are using the official Instagram API. The information we collect is the following:

- the main image in a post (there are sometimes multiple images and we just use the first)
- the number of likes
- the caption (without links to other users or usernames)
- the metadata (this includes the Instagram ID and timestamp)

We are not planning to present the data in any publication or reports. We are going to remove all personal information (e.g. usernames, email address, ..) from the collected data. All references to other users of the platforms will be removed.

#### 2. Data Storage

The consortium is storing only the response we collect using the Instagram-API. The data is stored on the Consortium Google Cloud and only members of this research project have access to it. We are unable to save the data on the university repository, because not all consortium members have access to it. We are not planning to store any image locally, and will only store the metadata and text. Based on the Instagram API Terms and Conditions, we need to delete any content that is no longer made available on their network. We will check periodically if the collected content is still active or needs to be deleted per these T&Cs.

#### Section F: Supporting documentation

#### **Information & Consent**

Participant information sheets relevant to project?

No

Consent forms relevant to project?

Nο

#### **Additional Documentation**

732328-stage2--SEALED-PROPOSAL.pdf (Document 040221)
 Horizon 2020 Project Proposal

#### **External Documentation**

- not entered -

#### Offical notes



Section G: Declaration			
Signed by:			
Matthias Dantone Date signed:			
Tue 19 September 2017 at 0	7:18		



Downloaded: 17/09/2018 Approved: 19/09/2017

Matthias Dantone FashionBrain Project

Dear Matthias

PROJECT TITLE: FashionBrain

APPLICATION: Reference Number 015090

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 19/09/2017 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

• University research ethics application form 015090 (dated 19/09/2017).

If during the course of the project you need to <u>deviate significantly from the above-approved documentation</u> please inform me since written approval will be required.

Yours sincerely

Kathryn MacKellar Ethics Administrator FashionBrain Project

# 4.3. Focused Sampling: Twitter Connectivity Analysis for Fashion

As detailed in Deliverable D9.3, an amendment to T3.3 (Focused sampling: Crowdsourcing fashion data source search) required an additional application process as the data collection and processing required were not compatible with the application described in Section 4.1. The lead applicant for this application is USFD, while the operations and the application was written in strict collaboration with UNIFR.

Application Activity	Twitter Connectivity Analysis for Fashion	
Activity Leader	Alessandro Checco - USFD	
	Jie Yang - UNIFR	
Draft Date	12/10/2018	
Application Description	See attached application and approval letter (below)	
Ethics Reviewers	Alessandro Bozzon, Paul Reilly, Daniel Rose	
Required Actions	• Added further details regarding identifying and	
	potential twitter users to question	
	• Specified who influencers are and how they engage	
	with platform	
	• Updated consent forms to reflect GDPR	

The application and the approval letter for this action is attached as follows.



# Application 023256

# Section A: Applicant details Date application started: Fri 12 October 2018 at 18:50 First name: Alessandro Last name: Checco Email: a.checco@sheffield.ac.uk Generic research application: Last updated: 12/11/2018 Department: FashionBrain Project Applying as: Staff member Research project title: Twitter connectivity analysis for fashion Has your research project undergone academic review, in accordance with the appropriate process? Similar applications: - not entered

#### Section B: Basic information

lame	Email	
ie Yang (University of Fribourg)	jie@exascale.info	
Proposed project duration		
art date (of data collection): u 25 October 2018		
ticipated end date (of project) t 12 October 2019		
3: Project code (where applicable)		
3: Project code (where applicable)		

#### Suitability

Takes place outside UK?

No

Involves NHS?

Nο

Human-interventional study?

Nο

ESRC funded?

No

Likely to lead to publication in a peer-reviewed journal?

Yes

Led by another UK institution?

Nο

Involves human tissue?

No

Clinical trial?

Nο

Social care research?

Nο

Involves adults who lack the capacity to consent?

Nο

Involves research on groups that are on the Home Office list of 'Proscribed terrorist groups or organisations?

No

#### Indicators of risk

Involves potentially vulnerable participants?

No

Involves potentially highly sensitive topics?

No

### Section C: Summary of research

#### 1. Aims & Objectives

The purposes of this research are to understand the structure of the connectivity of fashion influencers in Twitter (how users are connected to influencers) and to model user sociability, expertise, and motivation.

Fashion influencers are public figures (actors, singers etc.) that are active on Twitter and advertise fashion items. In the following the participants are Twitter users, but they are not influencers.

#### 2. Methodology

Participants will be recruited through an open call on Twitter and on crowdsourcing platforms, as explained in more detail in the recruitment section. Participants are Twitter users.

We will ask participant A to name one fashion influencer X (like a singer, actor, etc) and recommend a friend B (i.e. a Twitter follower of A) who they think might know the fashion influencer X. We will also ask participant A some fashion questions and some generic demographic/psychometric questions to understand their motivation and expertise.

After that, we will contact the named followers (B) and repeat the process to explore the connectivity path. We will identify the twitters by the user ID and it will be used to track the connectivity towards the fashion influencer named.

Some examples of the questions asked are attached in the following.

All user identification data (user ID) will be anonymised after the collection phase and kept in a University of Sheffield Google drive folder (to be used also by the external collaborators, that will not keep this data outside of the Google folder).

No data about the users will be published: the result will be a mathematical model of sociability, expertise, and motivation and of fashion

social graph connectivity

Users can ask to delete their data for up to 3 months after the collection.

#### 3. Personal Safety

Have you completed your departmental risk assessment procedures, if appropriate?

Yes

Raises personal safety issues?

Nο

No personal issues different than every day life.

#### Section D: About the participants

#### 1. Potential Participants

Fashion influencers are public figures (actors, singers etc.) that are active on Twitter and advertise fashion items. In the following the participants are Twitter users, but they are not influencers.

We will recruit participant on the (paid) crowdsourcing platform FigureEight and on Amazon Mechanical Turk via an open call in the platform marketplace. The crowdsourcing platform will be only used as a recruitment tool: participants will be users that have an account on the crowdsourcing platform and that actively select out task from the crowdsourcing marketplace.

In addition, we will offer this task to volunteers on Twitter via an open invitation on Twitter.

#### 2. Recruiting Potential Participants

Paid participants will select the crowdsourcing task from the search engine of the crowdsourcing platform, they will be provided with the consent form and a description of the research goals.

Volunteer participants will access to the link advertised on Twitter.

Some participants will be recommended by previous participants, and will be contacted using the private message (direct message) functionality of Twitter.

#### 2.1. Advertising methods

Will the study be advertised using the volunteer lists for staff or students maintained by CiCS? No

- not entered -

#### 3. Consent

Will informed consent be obtained from the participants? (i.e. the proposed process) Yes

A consent form will be presented, with explanation on how to withdraw consent (the participant can leave the task at any time and they can require to delete the data for up to 3 months after the collection), how to contact the researcher, and how the data will be used.

The participant will have to agree with the form in order to perform the task.

### 4. Payment

Will financial/in kind payments be offered to participants? Yes

We will have both volunteer participants that will not be paid, and crowdsourcing participant that will be paid through the crowdsourcing platforms Figure Eight or Amazon Mechanical Turk. The paid participants will be paid UK minimum wage.

#### 5. Potential Harm to Participants

What is the potential for physical and/or psychological harm/distress to the participants?

not bigger than the one encountered in everyday life.

How will this be managed to ensure appropriate protection and well-being of the participants?

participants can interrupt the task at any time.



#### Section E: About the data

#### 1. Data Processing

Will you be processing (i.e. collecting, recording, storing, or otherwise using) personal data as part of this project? (Personal data is any information relating to an identified or identifiable living person).

Voo

Which organisation(s) will act as Data Controller?

University of Sheffield only

#### 2. Legal basis for processing of personal data

If, following discussion with the UREC, you wish to use an alternative legal basis, please provide details of the legal basis, and the reasons for applying it, below:

- not entered -

Will you be processing (i.e. collecting, recording, storing, or otherwise using) 'Special Category' personal data?

#### 3. Data Confidentiality

What measures will be put in place to ensure confidentiality of personal data, where appropriate?

All user data will be anonymised and the research output will not contain any personal data: it will be a mathematical model of fashion influencers connectivity, together with a model of user sociability, expertise, and motivation.

We will communicate our confidentiality measures to the participants in the informed consent form.

#### 4. Data Storage and Security

In general terms, who will have access to the data generated at each stage of the research, and in what form

The data recorded will be securely stored on password protected computers at University of Fribourg. A copy will be stored on the researcher's university laptop for analysis purposes and it will be backed up on an external drive kept in a locked drawer in the Exascale Infolab at UNIFR.

What steps will be taken to ensure the security of data processed during the project, including any identifiable personal data, other than those already described earlier in this form?

All identifiable data (twitter user id and bio) will be completely anonymised after the data collection.

Will all identifiable personal data be destroyed once the project has ended?

Yes

Please outline when this will take place (this should take into account regulatory and funder requirements).

all identifiable personal data will be anonymised after 3 months from the end of the data collection: the time window during which users can require deletion of their data. After data even the researchers will not be able to associate a user to a piece of data.

#### Section F: Supporting documentation

#### Information & Consent

Participant information sheets relevant to project?

No

Consent forms relevant to project?

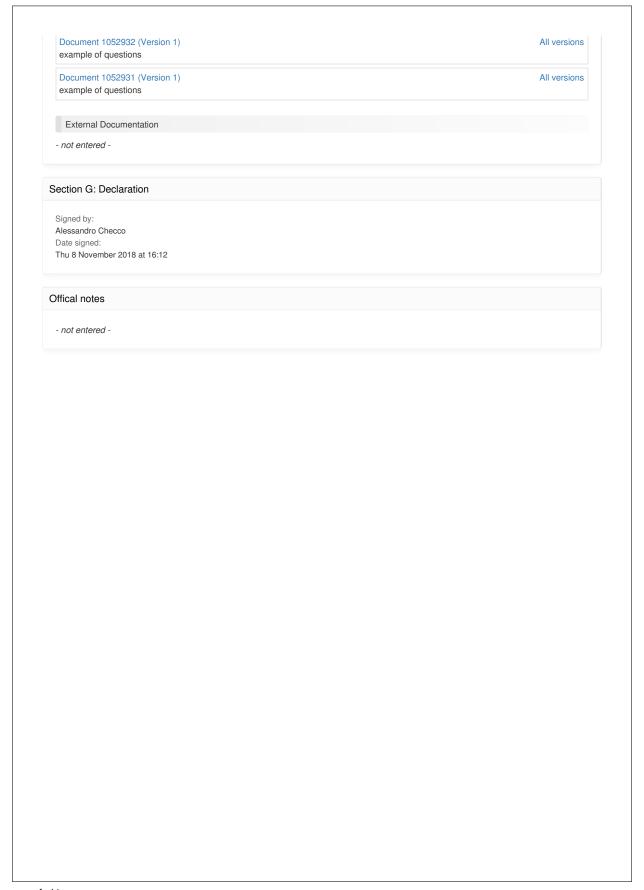
Yes

Document 1052074 (Version 2)

All versions

Additional Documentation







Downloaded: 12/11/2018 Approved: 12/11/2018

Alessandro Checco FashionBrain Project

Dear Alessandro

**PROJECT TITLE:** Twitter connectivity analysis for fashion **APPLICATION:** Reference Number 023256

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 12/11/2018 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 023256 (dated 08/11/2018).
- Participant consent form 1052074 version 2 (02/11/2018).

If during the course of the project you need to <u>deviate significantly from the above-approved documentation</u> please inform me since written approval will be required.

Yours sincerely

Kathryn MacKellar Ethics Administrator FashionBrain Project



# 4.4. Consent Forms

As specified in the Description of Actions, we refer to Deliverable D9.4 for all consent forms.

# 4.5. Compliance to EU Laws

We refer to Deliverable D9.2 for a detailed assessment of compliance to EU regulations by the FashionBrain Ethics Committee.

# A. Guidelines and Policies

The following documents have been provided to all research activity applicants in order to familiarise them with the FashionBrain Ethics policies and guidelines.

## Ethics Policy Governing Research Involving Human Participants Personal Data and Human Tissue

This main document clarifies the UREC ethos and approach to ethical data collection and processing: https://www.sheffield.ac.uk/polopoly\_fs/1.755691!/file /Ethics\_Policy\_Senate\_Approved.pdf.

# Research Ethics Policy Note no. 1 - Defining Human Research Participants, Personal Data and Human Tissue

An important document to clarify which data should be considered personal data, and which individuals should be considered participants: https://www.sheffiel d.ac.uk/polopoly\_fs/1.112746!/file/Research-Ethics-Policy-Note-1.pdf.

### Research Ethics Policy Note no. 2 – Principles of Transparency and Consent

This document clarifies the principles of transparency and consent: https://www. sheffield.ac.uk/polopoly\_fs/1.112749!/file/Research-Ethics-Policy-No te-2.pdf.

# Research Ethics Policy Note no. 3 – Participant and Researcher Safety and Well-being

This document details the guidelines and policies to guarantee well-being for both participants and researchers: https://www.sheffield.ac.uk/polopoly\_fs/1.11 2751!/file/Research-Ethics-Policy-Note-3.pdf.

#### Principles of Anonymity, Confidentiality and Data Protection

This document clarifies the principles of data confidentiality and data protection: https://www.sheffield.ac.uk/polopoly\_fs/1.112753!/file/Research-Ethi cs-Policy-Note-4.pdf.

## Research Involving Vulnerable People

This document is important as it allows the applicant to identify potential vulnerable participants: https://www.sheffield.ac.uk/polopoly\_fs/1.112756!/file/Research-Ethics-Policy-Note-6.pdf.

# Research Ethics Policy Note no. 14 - Research Involving Social Media Data

For applicants working on social media, this document is fundamental, especially to clarify the ideas of public domain and consent in this context (see also Deliverable D9.2 for more information): https://www.sheffield.ac.uk/polopoly\_fs/1.67 0954!/file/Research-Ethics-Policy-Note-14.pdf.

# A.1. Guidance on the Application Process

The applicants were provided with the following guide to the online application process: https://www.sheffield.ac.uk/polopoly\_fs/1.563235!/file/Online\_Ethics\_Guide\_Staff\_Applicant.pdf and with the following flowchart:

