Horizon 2020

fashion
BRAIN project

Understanding Europe's Fashion Data Universe

GEN - Requirement No. 3

Deliverable number: D9.3

Version 3.0

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	15, the last one being no. 9
Authors	Alessandro Checco, Jennifer Dick - USFD
Peer review	Peter Bath, Paul Reilly - USFD

Change Log

Version	Date	Status	Partner	Remarks
1.0	31/03/2017	Final	USFD	Rejected 15/03/2018
2.0	20/04/2018	Resubmitted Final	USFD	Rejected 15/10/2018
2.1	22/01/2019	Revised Draft	USFD	
3.0	04/02/2019	Resubmitted Final	USFD	

Deliverable Description

This document provides a thorough analysis of the ethics issues raised by this project and the measures that will be taken to ensure compliance with the ethical standards under H2020.

Abstract

This deliverable provides an analysis of the ethics issues raised by the FashionBrain project and the measures that have been taken to ensure compliance with the ethical standards under H2020, as reported by each applicant and in the Ethics Committee's reviews.

Table of Contents

List of Figures

List of Acronyms and Abbreviations

AMT	Amazon Mechanical Turk (www.mturk.com), micro-task crowdsourcing platform
API	Application Programming Interface
F8	Figure Eight (www.figure-eight.com), micro-task crowdsourcing platform
GDPR	General Data Protection Regulation
HTML	Hypertext Markup Language
ID	Identification
oAuth	Open Authorization
T&Cs	Terms and Conditions

1 Introduction

This deliverable provides an analysis of the ethics issues raised by the FashionBrain project and the measures that have been taken to ensure compliance with the ethical standards of H2020, as reported by each applicant and in the Ethics committee's reviews.

The remainder of the deliverable is structured as follows. In Section 2, the ethics issues raised by the project are analysed. In Section 3, a synthesis on the compliance of the project to GDPR principles and The European Code of Conduct for Research Integrity is provided. In Section 4, the FashionBrain Consent Manager is presented.

1.1 Scope of This Deliverable

We refer to Deliverable D9.1 for more information on the Ethics review process, to Deliverable D9.2 for an independent ethics assessment and information on the constitution and operations of the FashionBrain Ethics Committee, and to Deliverable D9.4 for a copy of all the consent form documents.

2 Ethics Dimensions

Presented below is a summary of the different issues and measures taken on ethics and data protection for the three FashionBrain activities for which data collections and processing were reviewed by the Ethics Committee: **Crowdsourcing**, **Focused Sampling**, and **Social Media**. We refer to Deliverables D9.1 and D9.4 for the ethics application documents and consent forms, respectively.

For the most part, a standardised process was utilised for the three activities. However, where there were differences between the three data collection processes, it is specifically highlighted and detailed in the text.

2.1 Risk Indicators

For all three activities, the risk level was considered to be **low**, in that:

- they are not human-interventional study;
- they do not involve human tissue;
- they are not clinical trials;
- they are not social care research;
- they do not involve adults who lack the capacity to consent;
- they do not involve research on groups that are described in the list of 'Proscribed terrorist groups or organisations'.

2.1.1 Personal Safety of Participants

All three activities:

- do not involve potentially vulnerable participant;
- do not involve potentially highly sensitive topics.

The potential for physical and/or psychological harm/distress to the participants has been estimated to not be greater than ones encountered in everyday life.

When the participants are involved in the data collection process, to ensure appropriate protection and well-being of the participants, the data collection can be interrupted by the participant at any time (this is relevant to **Crowdsourcing** and **Focused Sampling** activity).

2.1.2 Personal Safety of Consortium Members

For all three activities, no personal safety issues (greater than ones encountered in everyday life) have been raised for the consortium members.

2.2 Participants

Focused Sampling In this action, the participants are Twitter users. The crowdsourcing platform AMT or F8 will be only used as a recruitment tool: participants will be users that already have an account on the crowdsourcing platform and actively select out tasks from the crowdsourcing marketplace. In addition, this task will be offered to volunteers via an open invitation on Twitter.

Crowdsourcing In this action, the participants will be users that have an account on the crowdsourcing platform F8 or AMT and that actively select out tasks from the crowdsourcing marketplace.

Social Media In this action, the participants are fashion bloggers and Instagram users.

2.2.1 Recruitment

Focused Sampling Paid participants will be recruited on the (paid) crowdsourcing platform F8 and on AMT via an open call in the platform marketplace.

Paid participants will select the crowdsourcing task from the search engine of the crowdsourcing platform. They will be provided with the consent form, that includes a description of the project's research goals.

Volunteer participants will access an HTML 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.

Crowdsourcing Participants will be recruited on the (paid) crowdsourcing platform F8 and on AMT via an open call in the platform marketplace.

Paid participants will select the crowdsourcing task from the search engine of the crowdsourcing platform. They will be provided with the consent form that includes a description of the project's research goals.

Social Media The interaction with the participant will happen only through their publicly available data points (available on Social Media) which will be used to generate conclusions about trends based on the information collected. Strong and influential fashion bloggers will be identified based on the number of followers they have as well as publicly available site visit information.

2.2.2 Consent

For the actions **Crowdsourcing** and **Focused Sampling**, informed consent will be obtained from the participants. In addition to the project information, possible ethical issues (data use and privacy) and researcher contact information, the consent form will contain the address to contact to withdraw consent. The participant will have the opportunity to leave the task at any time and can request to have their personal data removed for up to 3 months after its collection[1]. The participant will have to sign the form (or click the "I agree" button in the online version) in order to perform the task.

For the action **Social Media**, only implicit consent will be obtained[2], so we refer to Section 4 for a more detailed analysis of consent for this action.

2.2.3 Erasure

For the action **Focused Sampling** and **Crowdsourcing**, the participants are informed on their consent form on the address of the research administrator to request erasure of the data within 3 months after collection phase. After this point the data will have been anonymised, so that it will not be possible to identify the data in order to delete it.

Regarding **Social Media** actions, we refer to Section 4 for a more detailed explanation on how they can request erasure of the data.

2.2.4 Payment

Payment procedure is different for each action.

Focused Sampling The action will have both volunteer participants that will not be paid, and crowdsourcing participants that will be paid through the crowdsourcing platforms Figure Eight or Amazon Mechanical Turk. The paid participants will be paid the equivalent of UK minimum wage.

Crowdsourcing Participants will be paid through the crowdsourcing platforms Figure Eight or Amazon Mechanical Turk. The paid participants will be paid the equivalent of UK minimum wage.

Social Media Participants will not be paid.

2.3 About the Data

Focused Sampling and **Crowdsourcing** actions will use personal data (user IDs) only during the collection phase: all personal data will be erased during the data

---

[1]This finite window of time is necessary as personal identifiers will be destroyed after that period.

[2]users agreed to the T&Cs of the social media platform they are using about access of their data via the platform Application Programming Interface (API).

processing phase. **Social Media** actions will use personal data such as Instagram metadata and images.

2.3.1 Data Controller and Data Processors

We clarify here which project partner organisation(s) will act as Data Controller and which as Data Processor(s).

Focused Sampling UNIFR is the sole data controller, USFD and UNIFR are both data processors.

Crowdsourcing USFD is the sole data controller, the members of the FashionBrain Consortium are all data processors.

Social Media Fashwell and USFD are data controllers, the members of the FashionBrain Consortium are all data processors.

2.3.2 Data Confidentiality

This section shows what measures are in place to ensure confidentiality of personal data.

Focused Sampling 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[3] connectivity, together with a model of user sociability, expertise, and motivation. These confidentiality measures will be communicated to the participants in the informed consent form.

Crowdsourcing No personal identifying information will be collected aside from the user IDs (needed to allow future erasure of the data); the online experiment will be anonymous. Anonymous user identifiers will be used to integrate data together and user IDs will be stripped during data processing. These confidentiality measures will be communicated to the participants in the informed consent form.

Social Media Only image metadata and text will be stored. The actual images will be reachable through the Instagram API but will not be stored. Erasure from Instagram will be mirrored in the dataset. Moreover, request of erasure can be made through the Consent Manager (Section 4.1). These confidentiality measures are published in the FashionBrain website (https://fashionbrain-project.eu/data-ethics-and-privacy/).

---

[3]Fashion influencers are public figures (actors, singers etc.) that are active on Twitter and advertise fashion items.

2.3.3 Data Storage and Security

This section explains who will have access to the data generated at each stage of the research, and in what form.

The data recorded will be encrypted and securely stored on password protected computers at the Data Controller and a copy may also be stored on other members' laptops for analysis purposes. The encrypted and password protected data will be backed up on an external drive kept in a locked drawer in the data controller's safe storage unit.

Apart from **Social Media** actions (for which we refer to Section 4), all identifiable data will be completely anonymised in the data processing phase (user IDs will be destroyed).

3 Legal basis for processing of personal data - GDPR

3.1 Lawfulness

For the actions **Crowdsourcing** and **Focused Sampling**, the data processing can be considered lawful as consent will be obtained, and the processed data will be completely anonymised after the collection phase.

For the action **Social Media**, some of data collected are personal data. The data processing can be considered lawful because of legitimate interest, and for the following reasons:

- The participant would expect this processing to take place, as the data are published for commercial reasons.
- The processing will benefit the participants, as increased exposure of the items advertised will increase the value of their operations.
- There is no position of power for the Data Controller over the participants.
- The participants are not in a vulnerable class.
- It is possible to stop processing at any time on request of the participant (see Section 4).
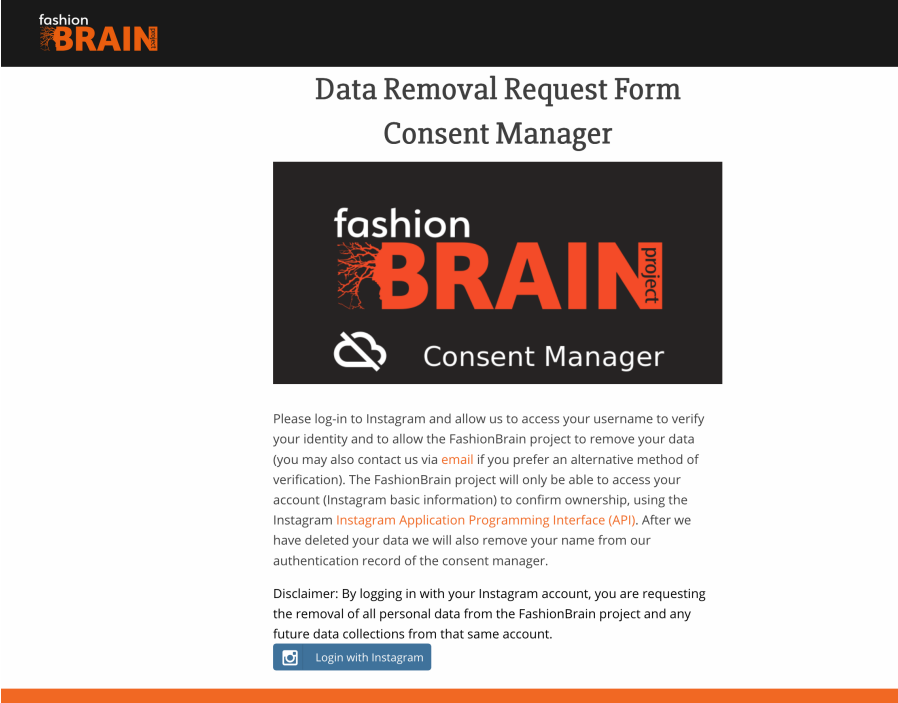
3.2 Fairness

As detailed in Section 2, consideration has been given to how data processing may affect the participating individuals and no adverse impact has been identified. The consortium will only handle participants' data in ways they would reasonably expect. Finally, the consortium does not purposely deceive or mislead participants in instances when personal data is collected as the project information including data ethics and privacy policies are disseminated via consent forms and also published on the public website .

3.3 Transparency

As explained in detail in Section 2, the research applicants comply with the participants' rights to be informed via the consent form.

4 Consent Manager

In the cases in which public data are obtained without consent, Art. 17 of GDPR (right to erasure) should be guaranteed. For this reason, the EU reviewers and the FashionBrain Ethics Committee recommended the development of a Consent Manager to allows the public to request an opt-out from FashionBrain data collection.

Figure 4.1: FashionBrain Consent Manager.

The Consent Manager has been published on the project website: https://fashionbrain-project.eu/consent_manager (Figure 4.1). It uses the Instagram Application Programming Interface (API) to verify the identity of the data owner that is making the opt-out request from FashionBrain data collection, without the risk of disclosing any other personal data. Technically, this works by redirecting the authentication to Instagram through the API and subsequently receiving a token back (`access_token`) from Instagram that allows minimal access to the Instagram profile (`id`, Instagram name). A request sent through this page will activate the data removal and blacklisting procedure: from the date of receipt of the opt-out submission, FashionBrain will destroy all past records belonging to the requester,

and the account ID will be added to a static blacklist table to prevent future data collection from that account. As previously explained, the Consent Manager is in place for cases in which public data are obtained without consent from the FashionBrain Project. Instead, implicit consent has been given on another platform when the users agreed to its T&Cs. For that reason, the owners of the data will not be personally notified of the existence of the Consent Manager. However, should they wish to have their current data or any future data removed, they will be able to do so via the website's Consent Manager. The consortium will advertise this page to reach the widest audience possible. By providing this transparency, the project is demonstrating and communicating strong ethical principles to the public.

Along with a Consent Manager on the website, a privacy page have been added to the FashionBrain website to explain what kind of public data is being collected and what is being done with this data, to guarantee the adherence to the principles of fairness and transparency of GDPR (https://fashionbrain-project.eu/data-ethics-and-privacy).