

Horizon 2020



Understanding Europe's Fashion Data Universe

GEN - Requirement No. 2

Deliverable number: D9.2

Version 4.0



Funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 732328

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	32, the last one being no. 27
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 26/04/2018
3.0	21/05/2018	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 process and establishment of the FashionBrain Ethics Committee and its actions for the duration of the project. It also provides the mid-term Ethics assessment written by the Committee.

Abstract

This deliverable describes the FashionBrain Ethics Committee constitution and its operations throughout the duration of the project. The Committee's main responsibility is to protect potential participants in the research, but also to ensure that the project is conforming to institutional, national and internationally accepted ethical guidelines. The committee may request modifications to research protocols or offer opinions on ongoing ethical issues in research to mitigate concerns where they exist. As well as regular communications and feedback to the consortium, the FashionBrain Ethics Committee has produced a mid-term ethics assessment that is available at the end of this document.

Table of Contents

List of Acronyms and Abbreviations v

1. Introduction 1

1.1. Scope of This Deliverable 1

2. The FashionBrain Ethics Committee 2

2.1. Selection Process 2

2.2. Ethics Committee Members 2

3. Meetings and Exchanges 9

A. Ethics Assessment 12

List of Acronyms and Abbreviations

AHRC	Arts and Humanities Research Council
AI	Artificial Intelligence
API	Application Programming Interface
CHIMR	Centre for Health Information Management Research
GDPR	General Data Protection Regulation
NHS	National Health Service
RCUK	Research Councils UK
SPERI	Sheffield Political Economy Research Institute
T&Cs	Terms and Conditions
UREC	University Research Ethics Committee

1. Introduction

This deliverable introduces the FashionBrain Ethics Committee, an independent group of experts that provides consultation on the ethical and data protection issues that could arise during the development of the project. The constitution and operations of the committee throughout the duration of the project will be summarised in this document.

The remainder of the deliverable is structured as follows. In Section 2, the FashionBrain Ethics Committee is presented. In Section 3, a chronological account of the actions performed by the Committee is reported. Finally, in Appendix A, the mid-term assessment produced by the committee is attached.

1.1. Scope of This Deliverable

We refer to Deliverable D9.1 for more information on the Ethics review process and a detailed report of all the ethics applications, to D9.3 for all ethical considerations within the project and to Deliverable D9.4 for a copy of all consent form documents.

2. The FashionBrain Ethics Committee

This section describes the composition of the FashionBrain Ethics Committee. The FashionBrain Ethics Committee is composed of Alessandro Bozzon, Natasha Whiteman, Paul Reilly, and Kristian Hildebrand, a group of independent and established scholars. Additionally, the committee can further consult with two members of the Sheffield University Research Ethics Committee (UREC) from the coordinating institute (USFD): Peter Bath (chair), and Jo Bates (Ethics coordinator). The FashionBrain Ethics Committee has also provided independent reviews of the Ethics applications, the process details of which are described in Deliverable D9.1.

2.1. Selection Process

Members of the committee were recommended and approved by all of the FashionBrain partners, and were ultimately chosen due to their expertise in a specific area related to the project objectives and outputs as follows:

Alessandro Bozzon leading expert in crowdsourcing and able to provide consultation on the issues related to data collection via human computation and crowdsourcing.

Kristian Hildebrand leading expert on computer vision and ideal for issues related to photo taking and AI.

Paul Reilly expert in social media and digital society and capable of reviewing the collection of data from social media and fashion blogs.

Natasha Whiteman an ethics expert in the context of digital media, thus ideally qualified to evaluate all issues.

2.2. Ethics Committee Members

In this section, a brief curriculum vitae for each of the members of the committee is presented.

Alessandro Bozzon

Assistant Professor in Web Information Systems
Delft University of Technology, The Netherlands

<http://alessandrobozzon.com/>

Since February 2013, Alessandro Bozzon has held an Assistant Professor (tenure track) position with the Web Information Systems group, at the Delft University of Technology. Since March 2014, he has held a Faculty Fellow position with the IBM Benelux Centre of Advanced Studies, where he is principal investigator of the “Inclusive Enterprise” research line. He is a member of the Delft Data Science initiative, with the “Human-Enhanced Data Management” research line. He is also a member of the AMS Amsterdam Institute for Advanced Metropolitan Solutions, with the “Social Data Science for Intelligent Cities” research line. He received his MSc in Computer Science & Engineering at Politecnico di Milano in 2005 (110/110 cum laude) and his PhD in Information Engineering (mark A) in April 2009, with a thesis on Model Driven Development of Search-based Web Applications (advisor Prof. Piero Fraternali).

Since beginning his career, his research has focused on the design, implementation and evaluation of novel methods and tools for Web information systems engineering and Web data management. The scientific work, which has resulted in high-profile publications, has always been supported by substantial implementation and demonstration efforts, and by a continuous interaction with industry. During his doctoral studies he was a visiting researcher at the L3S Research Centre (Hannover, Germany, 2007), and at FAST Search & Transfer – now Microsoft Development Centre Norway (Oslo, Norway, 2008). There, he had the opportunity to deepen his knowledge of Web scale and enterprise-class search systems. He had an extensive research collaboration with WebModels S.R.L., a startup of Politecnico di Milano, where he served as research consultant, analyst, project leader in an European project, and lecturer in industrial courses.

Kristian Hildebrand

Professor in Graphics and Interactive Systems
Beuth University, Berlin, Germany

<http://hildebrand.beuth-hochschule.de>

Kristian Hildebrand received his PhD from Technische Universität Berlin in 2013 under the supervision of Prof. Dr.-Ing. Marc Alexa. He holds a Diploma from the Computer Science and Media programme of the Bauhaus University Weimar. During his studies, he was a visiting researcher at the University of British Columbia in Vancouver and the Max-Planck-Institut Informatik in Saarbrücken. From 2006-2008 he worked as a software developer at art+com AG in Berlin where he participated in the development of several innovative media installations. From August-October 2012, he was a visiting researcher at Disney Research Zurich. Between 2014 and 2015 he worked as Principal Research Engineer at DISDAR GmbH developing web-based machine learning technologies. He is also co-founder of the art technology platform kunstmatrix. His research interests span Digital Fabrication, Computer Graphics, Computer Vision and Machine Learning.

Paul Reilly

Senior Lecturer in Social Media & Digital Society
Information School, University of Sheffield
<https://www.sheffield.ac.uk/is/staff/reilly>

Paul Reilly was appointed Senior Lecturer in Social Media & Digital Society in the Information School at the University of Sheffield in October 2015. He was previously a lecturer in Media and Communication at the University of Leicester (November 2009-September 2015) and a Part-Time Lecturer (Teaching only) at the University of Glasgow (2007-2010). His PhD (Glasgow, 2008) focused on the ways in which loyalist and republican groups used their websites to frame the Northern Irish peace process. Paul Reilly research focuses on the study of online political communication, with a focus on three key areas: (1) the use of social media by citizens to create and share acts of sousveillance (inverse surveillance); (2) the ways in which digital media can be used to crowdsource crisis information; and (3) the use of new media to reduce sectarian tensions and promote better community relations in divided societies such as Northern Ireland.

University Responsibilities

- Officer for PGR Student Affairs, Faculty of Social Sciences (October 2016-present)
- Deputy Director of Research, Information School (March 2016-present)
- Member of Digital Societies Research Group

Qualifications & Accreditations

- Senior Fellow, Higher Education Academy (July 2015)
- Postgraduate Certificate in Academic Practice in Higher Education (Distinction, awarded March 2013)
- Fellow, Higher Education Academy (November 2012)

Natasha Whiteman

Associate Professor of Media and Communication
University of Leicester, UK

<https://www2.le.ac.uk/departments/media/people/natasha-whiteman>

Natasha joined the University of Leicester in 2008. Prior to this, she was based the Institute of Education, University of London where she obtained her PhD, taught on a range of Masters and Doctoral level courses and was the Programme Leader for the MRes in Educational and Social Research. She is a member of the Association of Internet Researchers, Fan Studies Network and the British Sociological Association. She is the module leader for the following undergraduate modules: Creative Audiences, and Media and Communication Research in Practice. She also contributes to a number of undergraduate modules including Digital Games & Culture and The Media, Celebrity and Fan Culture. Her research explores ethical issues relating to media research, digital environments and media consumption activity, with a particular interest in the study of online fan and videogame cultures. She recently published a comment piece in The Conversation on the importance of ethical review, in the light of the Cambridge Analytica “controversy” <https://theconversation.com/facebook-data-why-ethical-reviews-matter-in-academic-research-94205>. Other recent work includes a British Academy funded project “The ethics of contemporary media consumption: An exploratory study of consumers’ ethical decision-making in respect of the use of the Internet to access and exchange media content,” and a monograph on the ethics of online research (please see below). Her earlier research examined social activity within online fan communities. Her PhD explored the workings of two fan communities (the video game site Silent Hill Heaven and television fan site City of Angel) over a three year period. Using text analysis of online posting activity, her thesis presented a sociological analysis of the patterning of consumption and identification with fan objects demonstrated on the public forums of these sites, and how these related to the ways that participants learned to be ‘good’ community members/successful fans.

Sheffield University Research Ethics Committee (UREC)

Peter Bath

Professor of Health Informatics
Information School, University of Sheffield, UK
<https://www.sheffield.ac.uk/is/staff/bath>

Peter has been interested in information science and health informatics for 23 years. Following his first degree in Applied Biology, he studied for his MSc at the University of Sheffield. At the end of the MSc, he remained in Sheffield and completed his PhD in chemical structure. He was appointed Research Fellow (1994), then Lecturer (1996) and Professor of Health Informatics in the Information School in 2013. His research interests are in Health Informatics and specifically, how patients, carers and health professionals seek, obtain and share information and advice in relation to their health and well-being through online digital resources. He has undertaken collaborative research with a range of organisations, including NHS Direct, the World Health Organisation, Macmillan Cancer Support and has been involved in a wide range of inter-disciplinary funded projects, including computer science, geriatric medicine, nursing, palliative care, biomedical engineering, psychiatry, public health medicine, epidemiology, health economics, statistics.

University Responsibilities

- Head of School
- Chair, University Research Ethics Committee (UREC)
- Director, Centre for Health Information Management Research (CHIMR)
- Deputy Programme Coordinator for the Health Informatics programme.
- Module Coordinator for the MSc in Data Science
- Programme Coordinator for the BSc in Informatics

Sheffield University Research Ethics Committee (UREC)

Jo Bates

Senior Lecturer in Information Politics and Policy
Information School, University of Sheffield, UK
<https://www.sheffield.ac.uk/is/staff/bates>

After completing a degree in Politics and American Studies at Keele University, Jo took a number of years away from academia working in the Housing Benefits department of Manchester City Council. She then returned to study for an MA in American Studies at Nottingham University, before moving into the information field via an NHS librarian graduate traineeship. Whilst completing her librarianship qualification at Manchester Metropolitan University, Jo became increasingly interested in the social shaping in information environments. This led her to staying on at Manchester to complete her PhD on the politics of the UK's Open Government Data initiative. She joined the Information School at Sheffield as a Lecturer in 2012. Her research examines the socio-material factors that influence the production and use of data, and that enable and restrict the movement of data between different people and organisations. This includes examining how government policies and legislation shape how data move between different sites of practice. Jo's recent AHRC-funded research project - The Secret Life of a Weather Datum - was cited by RCUK in their submission to the 2015 UK government inquiry on the Big Data Dilemma. She also gave oral evidence as an invited international expert on Open Government Data to a public hearing of the Canadian Federal Parliament in 2014.

University Responsibilities

- Information School Ethics Coordinator
- Faculty of Social Science Digital Society Network Steering Group
- Module Coordinator: Introduction to Health Informatics
- Member of the Digital Societies Research Group
- Associate Fellow of the Sheffield Political Economy Research Institute (SPERI)

3. Meetings and Exchanges

This section details the formal meetings and less formal email or telephone communications and exchanges between the FashionBrain consortium and the Ethics Committee.

These meetings assess whether progress is being made toward remedying any existing or new ethical concerns that have arisen from previous meeting and with respect to the research being undertaken in the project. The Ethics Committee may offer opinions and suggest modifications to research protocols to mitigate concerns where they exist.

Induction

The initial induction meeting took place on the 6th of February 2017. The purpose of this initial meeting was to introduce each Committee member to the specific details of the project as described in the Description of Action, potential ethical concerns involved in research activities and their commitment requirements over the lifetime of the project. Moreover, the document from the University of Sheffield “Ethics Policy Governing Research Involving Human Participant Personal Data and Human Tissue” (https://www.sheffield.ac.uk/polopoly_fs/1.755691!/file/Ethics_Policy_Senate_Approved.pdf) has been provided to the members of the committee.

Exchanges

The following briefly describes the communication exchanges between the activity leaders and the Ethics Committee that took place over the course of each review process for each research activity following the Induction. Copies of each ethics application with actions taken can be found in Deliverable 9.1, Section 4.

USFD - Crowdsourcing Experiments where participants were asked to perform tasks related to the identification of fashion products, to express their opinions on fashion and the collection of demographic data.

The activity leader (Alessandro Checco) completed a draft of an ethics consent form on 28 March 2017 to the Sheffield Ethics System. Natasha Whiteman, Alessandro Bozzon, and Paul Reilly were selected as reviewers. Issues raised by the reviewers on 31 March and 8 April 2017, were amended and remedied by the group leader on 7 April, 10 April and 15 April 2017, to the satisfaction of the reviewers. Final approval was granted on 23 May 2017.

Fashwell - Collect social Media Posts and analyse publicly-available data from the World's biggest fashion bloggers and match them to products.

The activity leader (Matthias Dantone) completed a draft of an ethics consent form on 13 June 2017 to the Sheffield Ethics System. Alessandro Bozzon, Paul Reilly, and Kristian Hildebrand were selected as reviewers. Issues raised by the reviewers on 13 June and 23 June 2017, were amended and remedied by the group leader on 21 July 2017, to the satisfaction of the reviewers. Final approval was granted on 19 September 2017.

USFD and UNIFR - Twitter Connectivity Analysis for Fashion to understand how users are connected to influencers and to model user sociability, expertise and motivation.

The activity leaders (Alessandro Checco and Jie Yang) completed a draft of an ethics consent form on 12 October 2018 to the Sheffield Ethics System. Alessandro Bozzon, Paul Reilly and Daniel Rose (Consent Manager - refer to D9.1 for further information) were selected as reviewers. Issues raised by the reviewers on 21 October and 24 October were amended and remedied by the group leaders on 8 November 2018, to the satisfaction of the reviewers. Final approval was granted on 12 November 2018.

Meeting 1

Meeting 1 took place on 11 December 2018. The purpose of this meeting was to discuss the first draft of the Ethics Assessment, and to require clarification to the consortium. The main comments and actions resulting from this meeting were as follows:

- The Ethics Committee asked some clarifications on the technology behind the Consent Manager.
- The Ethics Committee discussed how to clarify some details on data security in the Ethics Assessment.
- The Ethics Committee discussed on how to clarify some details on GDPR in the Ethics Assessment.
- The Ethics Committee asked for more details on the Ethics application related to the T3.3 amendment.

Meeting 2

Meeting 2 took place on 16 January 2019. The purpose of this meeting was to ratify the mid-term Ethics Assessment. The main comments and actions resulting from this meeting were as follows:

- The Ethics Committee ratified the added last section of the mid-term Ethics assessment.
- No other actions were required.

Future Meetings

Two further meetings are scheduled for March 2019 and December 2019, meetings 3 and 4, respectively.

- Meeting 3 will focus on the evaluation of the status of the project and the analysis of potential new issues.
- Meeting 4 will address the final ethics report required to be submitted to complete the ethics governance and formal reporting procedures.

A. Ethics Assessment

The mid-term Ethics Assessment of the FashionBrain Ethics Committee is attached below.

Ethics Assessment Change Log

Version	Date	Status	Remarks
1.0	10/12/2018	Draft	Initial version
1.1	15/01/2019	Final	<ul style="list-style-type: none">• Added GDPR and data security details• Added a new ethics assessment for T3.3 amendment• Added an analysis of the Consent Manager• Clarification on what has already been done regarding ethics committee assessments and what new actions are being evaluated or finalised.

FashionBrain Ethics Assessment

Alessandro Checco, FashionBrain Ethics Committee

Version 1.1

Introduction

This document is an assessment of the ethics status of the FashionBrain project, to certify that the current *modus operandi* is in agreement with the EU H2020 Ethics and data protection regulations, spirit and best practices¹. It has been countersigned by the independent FashionBrain Ethics committee, and it has been written with feedback and advice from ethics experts from the University Research Ethics Committee (UREC) at the University of Sheffield. It is the result of regular meetings between the members of the FashionBrain project and the FashionBrain ethics committee².

In order to evaluate the ethics status of the project, the FashionBrain Ethics committee also has access to deliverable D9.1 which summarises the outcomes of all of the ethics applications for the project in terms of all ethics dimensions identified, and D9.4 which contains the consent forms and application forms processed by the FashionBrain Ethics Committee through the FashionBrain online Ethics Application System³.

The document is structured as follow: In Section 1, the self-assessment presented in the Description of the Action (DoA) by the FashionBrain consortium is evaluated. In Section 2, the ethics issues identified at the beginning of the project are discussed. In Section 3, the actions taken to address these issues are examined. In Section 4, new issues and actions that arose during the development of the project are identified and discussed. Section 5 is the evaluation of those new actions, that have been discussed during the second Ethics committee meeting. Section 6 concludes the document.

¹ The Ethics committee will evaluate the project accordance to the Modernised Convention 108 and the EU General Data Protection Regulation (GDPR).

² The FashionBrain Ethics committee constitution and the meetings schedules are reported in detail in deliverable D9.2.

³ The FashionBrain online Ethics Application System and the FashionBrain Ethics application procedure is presented in D9.1.

Table of Contents

1. DoA Self-Assessment	3
1.1 National Legal and Ethics Requirements	3
1.2 Ethics Issues	3
1.3 Security	4
2. FashionBrain Ethics Issues Assessment	5
2.1 Crowdsourcing Experiments	5
Action Required	5
2.2 Collecting and analysing social media posts	6
Action Required	6
2.3 Interviews with stakeholders	6
Action Required	6
3. Evaluation of Actions	6
3.1 Crowdsourcing Experiments	7
3.2 Collect social media posts	8
4. New Issues	9
4.1 Amendment to T5.2 - Social Media Post Linking	9
Action Required	10
4.2 Amendment to T3.3 - Fashion Influencers	10
Action Required	10
4.3 Consent Manager	10
5. Evaluation of New Actions	11
5.1 Social Media Post Linking	11
5.2 Consent Manager	12
6. Conclusions	13

1. DoA Self-Assessment

The DoA ethics self-assessment (Section 5.1 of DoA part B) has been written at the beginning of the project and describes, in general terms, how the consortium intends to meet the EU and national ethics and data protection regulations, which are the ethics issues that will arise from the project tasks, and how the consortium intends to address such ethics issues. In this section we will discuss the self-assessment document in relation to the Modernised Convention 108 and the GDPR, and refer to other sections when the topic has evolved over the duration of the project.

1.1 National Legal and Ethics Requirements

The first section of the self-assessment answers the question: *“Describe how the proposal meets the national legal and ethics requirements of the country or countries where the tasks raising ethical issues are to be carried out.”*

In the self-assessment, the consortium claims the project will conform with EU regulations, in particular:

“Where a host organization or its Member State does not have a specific regulation and procedure in place, the Project Coordinator will develop a procedure in line with those used by other partners, and serve as the reviewing body so that all research activities are reviewed for research ethics, to ensure that they comply with EU regulations.”

Each member has the responsibility to guarantee that their own national and internal Ethics procedures are respected. Keeping track of each internal application process would be not practical. For this reason, the FashionBrain project used an online central Ethics Application System (<https://ethics.ris.shef.ac.uk/>, presented in detail in deliverable D9.1), that provides a centralised system for the FashionBrain Ethics committee to review the applications of all project member. The ethics committee evaluated each application with respect to their adherence to EU regulations and ethics standards. Moreover, the Ethics Committee agreed to follow The University of Sheffield Research Ethics Policy Notes⁴.

Moreover, in the DoA the consortium claims that:

“In cases where a participant reveals illegal information during project meetings, the appropriate agency will be contacted and the participant will be advised that such action has taken place; this is highly unlikely given our research objectives and the types of interactions that we envisage with participants.”

From the information gathered, no participant has revealed any illegal information during the duration of the project.

1.2 Ethics Issues

In the DoA, the consortium claims that the FashionBrain project may collect personal information from online social media and, overall, will be using three types of data:

1. Data that is restricted by institutional policy and regulations, e.g., product catalogs of private corporations;
2. Data from previously collected data sets, e.g., logs and other data from human experiments or publicly available test collections;

⁴ More details on The University of Sheffield Research Ethics Policy Notes are available in deliverable D9.1.

3. Newly collected data that may come from a) self-reported data from interviews or questionnaires, b) unobtrusive observation (in person and/or via computer) using log files, and screen capture video., c) the text of articles, blogs, and social media posts.

For each of these data types, the actions proposed to perform the data collection and processing are in agreement with standard practices. However, more detail is required on the specific datasets and experiments performed, and for this reason we refer to Sections 2 and 3, where we examine the issues in more details, together with the actions taken to address them.

1.3 Security

The proposal claims that:

“The FashionBrain proposal does not involve activities or results raising security issues.” and that “The FashionBrain proposal does not involve 'EU-classified information' as background or results.”.

These claims have been confirmed from all the information gathered during the FashionBrain Ethics Committee’s work.

2. FashionBrain Ethics Issues Assessment

An initial analysis of the potential ethical risks from activities being carried out during the lifetime of the FashionBrain project was undertaken at the beginning of the project. Three main areas where a potential ethics risk could arise had been initially identified:

1. Running crowdsourcing experiments where participants are asked to perform tasks related to the identification of fashion products, to express their opinions on fashion and the collection of demographic data.
2. Collecting and analysing publicly-available social media posts from the World's biggest fashion blogger and matching them to products.
3. Performing interviews in Zalando with both stakeholders who need to take decisions based on available data and with data scientists who create data products. The interviews serve to better understand challenges faced when dealing with data exchanges both internal and external to Zalando.

In this section, we will assess each issue, evaluate the ethics risks and describe the actions required, actions that will then be analysed in Section 3.

Whenever more details on the data collection and data processing are required, the activity leader is required to submit an ethics application through the online Ethics Application System, explaining:

- aims and objectives of the research,
- data collection methodology,
- potential participants composition,
- recruiting process of potential participants,
- how consent will be obtained,
- payment of the participants (if any),
- data processing details,
- details on data storage and data security procedures.

Moreover, the activity leader should explain whether the data collection process involves **potentially highly sensitive topics** or **potentially vulnerable participants**, what is the potential for **physical and/or psychological harm/distress to the participants** and the well-being of the participants will be protected⁵.

Finally, the activity leader should explain whether the data collection and data processing raises personal safety issues to the researchers.

2.1 Crowdsourcing Experiments

The University of Sheffield is leading activities that involve crowdsourcing experiments. In particular, the issue has been identified as:

Running crowdsourcing experiments where participants are asked to perform tasks related to the identification of fashion products, to express their opinions on fashion and the collection of demographic data.

Action Required

The activity leader is required to submit an ethics application through the online Ethics Application System.

⁵ We refer to deliverable D9.1 for details on the review process and ethics policy documents.

2.2 Collecting and analysing social media posts

Project partner Fashwell is leading activities that involve social media data collection. In particular, the issue has been identified as:

Collecting and analysing publicly-available social media posts from the World's biggest fashion blogger and matching them to products.

Action Required

The activity leader is required to submit an ethics application through the online Ethics Application System.

2.3 Interviews with stakeholders

Zalando is leading activities that involve interviews. In particular, the issue has been originally identified as:

Performing interviews in Zalando with both stakeholders who need to take decisions based on available data and with data scientists who create data products. The interviews serve to better understand challenges faced when dealing with data exchanges both internal and external to Zalando.

Action Required

It has been decided that this issue does not require an ethics application with consent form, as the individuals in Zalando to be interviewed are Zalando employees and they were to be informally consulted in order to define future challenges. As established in The University of Sheffield Research Ethics Policy Note no. 1⁶, this kind of consultation does not constitute research in itself, but rather is only contributing to the design of a research project, with no intention to disseminate the data as academic research.

It is worth noticing that even if no ethics application is needed, personal data must be obtained and handled in compliance with the Modernised Convention 108 and the GDPR. In this specific case, no personal data has to be collected: no individual will be identifiable by the generic output obtained from the interviews, nor it would be possible to single them out in a way which makes it possible to find out who the data subject is by conducting further research.

3. Evaluation of Actions

For the issues identified in Section 2, the actions required have been carried out through the FashionBrain online Ethics Application System, hosted by the University of Sheffield, and reviewed by the FashionBrain Ethics Committee. We summarise in this section the salient points of each action with respect to data protection, consent and ethical risks. For a more detailed analysis of all the ethics dimensions identified we refer to deliverable D9.3.

⁶ The University of Sheffield's Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue: Version 7.4 (Note no. 1), available at https://www.sheffield.ac.uk/polopoly_fs/1.112746!/file/Research-Ethics-Policy-Note-1.pdf.

3.1 Crowdsourcing Experiments

The activity leader (Alessandro Checco) presented through the FashionBrain online Ethics Application System an ethics application with a draft of an ethics consent form. The review process has been successfully completed as of the 8th of April 2017.

This is a summary of the main ethics dimensions regarding this applications:

- **Participants:** Any adult (anyone aged 18 years or over) who volunteers to complete the task on the crowdsourcing platform, recruited via submitting tasks/jobs to two popular crowdsourcing platforms: Amazon Mechanical Turk and Crowdfunder.
- **Data collection methodology:** crowdsourcing experiments will be run where participants will be asked to perform simple tasks like judging the relevance of a document or classifying an image using an online form 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 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.

- **Consent:** A consent form will be presented, that contains information on which data will be collected, how the data will be used, and 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). They are available in deliverable D9.4.
- **Data security:** 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 micro-task id and their worker id: this information combination is only known by the worker and does not jeopardize their anonymity.

The data will be encrypted and 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, in accordance with the University of Sheffield policies. The encrypted 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.

- **Risks for the participants:** The crowdsourcing studies will not involve any psychologically distressing search tasks. The risks of participating are the same as those experienced in everyday life.
- **GDPR Lawfulness:** The data processing can be considered lawful because consent will be obtained, and the processed data will be completely anonymised after the collection phase. The participants interact with the researchers through a crowdsourcing marketplace, that guarantees another layer of protection for the participants. The processed datasets will not contain any personal data.

For a more detailed report of the application and the consent form we refer to D9.1 (ethics reviews) and D9.4 (general information)

Discussion

The Ethics Committee believes this application to be in observance with the Ethics and data protection standards and it does not present critical risks.

3.2 Collect social media posts

The activity leader (Matthias Dantone) presented through the FashionBrain online Ethics Application System an ethics application with a draft of an ethics consent form. The review process has been finalised as of the 19th of September 2017.

This is a summary of the main ethics dimensions regarding this applications:

- **Participants:** In this research Fashwell is collecting publicly available social media posts from influencers and fashion bloggers, that are public figures who publish content for public use.
- **Data collection methodology:** Fashwell will collect publically available data points from fashion bloggers that are available in Social Media and blogs which will be used to generate conclusions about fashion trends. Fashwell will be able to identify influential fashion bloggers based on the number of followers and publically available site visit information. The only information that will be used in the study is data that can be derived from an image (the products and their attributes and the setting for the image) together with publicly available text data points like captions and hashtags. Fashwell will remove all personal information (usernames) from the collected data.
- **Consent:** consent will not be obtained (except from the implicit gatekeeper consent obtained by using Instagram API and following its T&C), as the data will be anonymised.
- **Data security:** The consortium is storing only the response we collect using the Instagram's Application Programming Interface (API). The encrypted and password protected data is stored on the consortium Cloud and only members of this research project have access to it. The images will not be stored locally: Fashwell will only store the metadata and text. Based on the Instagram API Terms and Conditions (T&Cs), it is necessary to delete any content that is no longer made available on their network. Fashwell will check periodically if the collected content is still active or needs to be deleted as per these T&Cs.
- **Risks for the participants:** The risks of participating classify as "minimal risk" where the probability and magnitude of harm or discomfort anticipated in the research are no greater in and of themselves from those ordinarily encountered in the performance of routine physical or psychological daily life / tasks.
- **GDPR Lawfulness:** some of data collected are personal data. The data processing can be considered lawful because of legitimate interest, for the following reasons:
 - a. The participant would expect this processing to take place, as the data are published for commercial reasons.
 - b. The processing benefit the participants, as increased exposure of the items advertised will increase the value of their operations.

- c. There is no position of power of Fashwell over the participants.
- d. The participants are not in a vulnerable class.
- e. It is possible to stop processing at any time on request of the participant.

For a more detailed report of the application and the consent form we refer to D9.1 (ethics reviews) and D9.4 (general information).

Discussion

The data to be used will be classified as publicly available, as only forum data that is openly accessible to the public, visible to web browsers and search engine results. However, as suggested by the University of Sheffield Research Ethics Policy Note no. 14 on Research Involving Social Media Data, social media research **often constitutes a grey area**, where the distinction between public and private is often unclear and may depend on user perceptions and forum culture as well as legal definitions.

In terms of gaining consent from the individual forum users whose posts may be studied, there are some practical and ethical issues. As discussed in the University of Sheffield Research Ethics Policy Note no. 14, and the British Psychological Society's guidance on Internet-mediated research ethics (British Psychological Society, 2017), an important distinction that influences whether individual consent should be sought, is whether or not the forum users might reasonably be expected to know that their posts are public, not private, and that their posts might be read by strangers. This is definitely the case in this instance, since fashion influencer are actively seeking publicity with their posts for commercial reasons.

Despite that, the Ethics Committee recommends the consortium to follow the recommendation received during the interim review, and **build a consent manager that will allow users to remove their data**⁷, if they wanted to do so. It is important to note that the risk is still considered to be small, as the content is akin to advertisement content, and because Fashwell anonymise the data, does not store the images (only metadata), and already mirrors any data removal that should happen on Instagram.

4. New Issues

In this, section, we evaluate new issues that have arisen during the project or after the interim review, and discuss the actions taken so far as well as suggestions for the future.

4.1 Amendment to T5.2 - Social Media Post Linking

The original version of T5.2 and related deliverable D5.2 had the objective of re-identifying social media profiles and posts across different social media platforms, with the final goal of linking the semantically similar posts across different platforms, to be able to improve product recognition and predict new trends. However, after the analysis in D3.3 and the preliminary work in WP5 and WP6, the consortium observed that:

the objective of product recognition can be achieved without the need of linking profile across different platforms: by using computer vision and deep learning to perform the product linking, the consortium would not need to store the actual images of the outfits. The images and the metadata of the images should be considered as personal data and therefore have special data protection constraints. By extracting non-private data from the images and not storing the images, the project would significantly reduce the privacy concerns. The proposed system will be more general and more powerful than a mere cross-profile linking, because it will rely on an abstract product representation, and thus will be able to work on any social medium.

⁷ See Section 4.3 for additional discussion on the Consent Manager.

For this reason, T5.2 has been amended from “Social Media Post Linking” to “Product Taxonomy linking”:

This task focuses on identifying products on different social media platforms. This task consist of different steps. We will start with crawling the different social media profiles from influencers and fashion bloggers. In T5.1 we extract the identity of the products and in this task we go one significantly step further. We plan to extract additional characteristic features (e.g. colors, necktype, occasion) by using an attribute taxonomy (e.g. color:red, neck-type:v- neck) of fashion-products. The extraction is based on a defined taxonomy (using the results from T1.3) and will be performed automatically using computer vision and deep learning. The proposed methods are going to be used for D5.2 and D5.3. The extracted attributes will improve T5.3 in the following ways: The output of T5.2 will be used to significantly improve the accuracy of the TimeSVD_vc algorithm (implemented in D5.3) which uses the visual features of the fashion items to perform the prediction. The results of this tasks will be reported in D5.2.

Similarly, D5.2 has been amended from “Social Media Post Linking” to “Product Taxonomy Linking”:

This deliverable extends D5.1 with a demo that integrates the products social media posts linking, that means that we recognise products from different social media channels.

Action Required

While the original application would require careful lawfulness and data protection analysis, this amendment does not cause ethics concerns, as the new form of data collection and data processing proposed falls under the same conditions examined in Section 3.2.

4.2 Amendment to T3.3 - Fashion Influencers

After the work in D3.3, the FashionBrain consortium observed that:

The objective of predicting new trends does not need the early detection of fashion influencers. The fashion industry is rather interested in keeping track of new products, observing already successful influencers. For this reason, T3.3 and related deliverable D3.3 will be amended: Instead of detecting new influencers, T3.3 will explore the way influencers are connected to each other and to the public, i.e. explore the social graph of existing influencers.

For this reason, T3.3 “Focused sampling: Crowdsourcing fashion data source search” has been amended:

A special type of crowdsourcing tasks will be developed in the context of this project, that will allow to model sociability, expertise, and motivation of fashion users when using social media. The goal is to explore the way fashion influencers are connected to each other and to the public. In this task we will develop an additional set of task interfaces and questionnaires that will run on a focused sample of the crowd population. The results of these tasks will be D3.3 and crowd-generated models that will be used in WP5, 6, and, 7.

Action Required

The amendment to T3.3 **does raise** additional Ethics concerns in terms of data protection and consent, because of the different way the data will be used (that requires consent from the participants). For these reasons **a new ethics application and consent form** has to be submitted to the online Ethics Application System, and has been reviewed by the FashionBrain Ethics Committee in the second meeting, as shown in Section 5.

4.3 Consent Manager

In the cases in which public data are obtained without consent (under the limiting conditions that make this action acceptable, as explained in Section 3.2), Art. 17 of GDPR (right to erasure) should be guaranteed. For this reason, it is recommended to developing a Consent Manager and making it available on the project website as a tool that allows the public to submit an opt-out from FashionBrain data collection. The user will need to submit certain details enabling the system to identify and remove collected content from that user. A potential solution would be, for example, to use a digital identification method (oAuth) to guarantee a user is the proprietor of the specific account for which they are asking for the removal of the data. From the date of receipt of the opt-out submission, FashionBrain will destroy all past communications, and the account ID will be added to a static blacklist table to prevent future data collection from that account.

Along with a Consent Manager on the website, it is recommended that a privacy page is added, to explain what kind of public data is being collected and what is being done with this data. This will guarantee the adherence to the principles of fairness and transparency of GDPR (for the lawfulness principles we refer to the specific ethics applications). In Section 5, the implementation of the consent manager has been analysed.

5. Evaluation of New Actions

After the interim review, two new actions have been required, as presented in Section 4. We evaluate those actions in this section.

Moreover, after the second meeting with the Ethics Committee, Section 4 has been updated with more information about GDPR lawfulness.

5.1 Focused Sampling

Partner UNIFR presented through the FashionBrain online Ethics Application System an ethics application with a draft of an ethics consent form. The review process has been finalised as of the 12th of November 2018.

This is a summary of the main ethics dimensions regarding this applications:

- **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.

Partner UNIFR 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, UNIFR will advertise this task to volunteers on Twitter via an open invitation on Twitter.

- **Data collection methodology:** Participants will be required to name a fashion influencer and a Twitter user that might know them. Additionally, some fashion questions and some generic demographic/psychometric questions will be asked to understand the participants' motivation and expertise.

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.

- **Consent:** 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.

- **Data security:** The encrypted data recorded will be securely stored on password protected computers at University of Fribourg. A copy in encrypted form 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.

All identifiable data (Twitter user id and bio) will be completely anonymised after the data collection. 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 anonymisation, even the researchers will not be able to associate a user to a piece of data.

- **Risks for the participants:** The risks of participating classify as "minimal risk" where the probability and magnitude of harm or discomfort anticipated in the research are no greater in and of themselves from those ordinarily encountered in the performance of routine physical or psychological daily life / tasks.

- **GDPR Lawfulness:** The data processing can be considered lawful because consent will be obtained, and the processed data will be completely anonymised after the collection phase.

For a more detailed report of the application and the consent form we refer to D9.1 (ethics applications) and D9.4 (consent form).

Discussion

No additional discussions.

5.2 Consent Manager

The consent manager has been published https://fashionbrain-project.eu/consent_manager/: it uses the Instagram Application Programming Interface (API) to be able 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 work by redirecting the authentication to Instagram through the API and 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.

Along with a Consent Manager on the website, a privacy page (<https://fashionbrain-project.eu/data-ethics-and-privacy/>) has been added, to explain what kind of public data is being collected and what is being done with this data. 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.

The Ethics Committee consider this solution adequate.

6. Conclusions

The FashionBrain Ethics Committee considers the online Ethics Application System a sufficient tool to perform the review process of the ethics applications and consent form review.

For all the tasks that require collection of personal data, the consortium has followed the regulations and best practices of the Modernised Convention 108 and the GDPR.

The changes of the project DoA have been reviewed and approved by the Ethics Committee on the 11th of December 2018, with the requirement for some additional actions: the development of a Consent Manager as detailed in Section 4.3, more information on GDPR compliance, and a review of the ethics application of the Social Media Post Linking task.

On the 16th of January 2019, these additional actions have been completed.

In conclusion, the Ethics Committee assesses the FashionBrain project data collection and processing as satisfactory in accordance to the Modernised Convention 108 and the GDPR.

The Ethics Committee will continue to have regular meetings with the consortium as reported in D9.2, to continuously assess the Ethics status of the project and to serve as independent reviewers for the ethics applications needed.