Horizon 2020

Understanding Europe's Fashion Data Universe

Project Work Plan and Quality Plan

Deliverable number: D8.1

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 January 2017
Resubmission Delivery Date	27 February 2019
Number of pages	43, the last one being no. 37
Authors	Jennifer Dick - USFD
Peer review	Alessandro Checco - USFD

Change Log

Version	Date	Status	Partner	Remarks
1.0	08/02/2017	Final	USFD	Rejected 15/03/2018
2.0	20/04/2018	Resubmitted Final	USFD	Rejected 26/04/2018
3.0	11/05/2018	Resubmitted Final	USFD	Rejected 15/10/2018
4.0	27/02/2019	Resubmitted Final	USFD	

Deliverable Description

This deliverable covers administrative and operational structures and approaches, information, contact points, document templates and use of logos. It will be accompanied by a quality plan that brings quality control aspects, risk management, conflict resolution and decision making procedures. This includes the risk plan and continuous risk monitoring plan. The monitoring of risks will be reported throughout the planned project in the periodic reports.

Abstract

The main goal of this deliverable is to provide a single point of reference on the quality assurance policies that will be used in the FashionBrain Project by all Consortium members for many day-to-day activities to aid in the effective implementation of the project. This deliverable covers administrative and operational structures and approaches, information, contact points as well as document and logo templates. It will be accompanied by a quality plan that brings quality control aspects, risk management, conflict resolution and decision making procedures which will be reported throughout the planned project in the periodic reports.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

API	Application Programming Interface
CA	Consortium Agreement
DMP	Data Management Plan
EC	European Commission
FaBIAM	FashionBrain Integrated Architecture
FAIR	Findable, Accessible, Interoperable and Reusable
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
JURI	European Parliament Committee on Legal Affairs
PDF	Portable Document Format
T&Cs	Terms and Conditions
UREC	University Research Ethics Committee
URL	Uniform Resource Locator
WP	Work Package

1. Introduction

The establishment of the FashionBrain Project Work Plan and Quality Plan (and its respective procedures) provides a single point of practical guidance to the coordinator, coordinating bodies and project partners during the lifetime of the project to ensure the quality of its outputs. The main goal of the Project Work Plan and Quality Plan is to deliver a clear set of guidelines for procedures and standards to be used in all major project activities including administrative and operational structures and approaches, information on reporting requirements, contact points as well as document and logo templates. Specifically the plan addresses quality control aspects, risk management, conflict resolution and decision making procedures.

In addition, the plan outlines the procedure to be followed by all project partners when when engaging in public engagement and dissemination activities (including presentations, publications, the project webpage and social media).

1.1. Scope of This Deliverable

This document is based on the terms and conditions established in the Grant Agreement; in particular Annex1: Description of Action; as well as in the Consortium Agreement specifications and requirements.

This document will also be complemented by Deliverables D7.2 - Project Websheet, D7.3 - Communication Plan, D7.5 - Draft Business Plan, D8.2 – Data Management Plan and D9.1 - D9.4 - General Ethics Requirements.

2. Project Work Plan

2.1. Management Structure

Project management activities will be aimed at securing an effective, efficient, and smooth execution of the Workplan, allowing the partners to fulfil their individual and collective aims established for the project. The management structure of the FashionBrain project (Figure 2.1) and its operational procedures are described in detail in several grant related documents:

Grant Agreement: Division of beneficiaries' roles and responsibilities

Description of Action: Management structure and work package delegation

Consortium Agreement: Operational procedures within the consortium

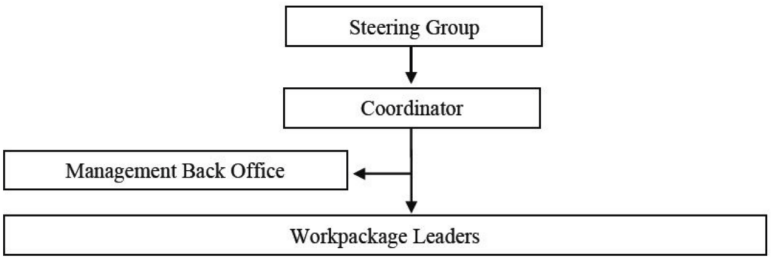

Figure 2.1: FashionBrain management structure.

The management structure ensures i) administrative and financial coordination; ii) efficient communication between project partners and stakeholders; iii) effective methods for mitigating risk and resolving conflicts; iv) partners are provided with the support and tools to fulfil their work package role and responsibilities; and v) compliance of the EC Grant Agreement.[1]

A brief overview of the management structure is as follows:

Steering Committee: This committee is the top management executive team of the FashionBrain project and the ultimate decision making body for the purpose of the Project. Each representative will have been delegated the proper authority to make decisions on behalf of their organisation to deliver strategic decisions; resolve project-wide issues; agree on project changes; define the schedule

[1] It is important to note that each partner may also have their own institutional management structure that will run alongside the FashionBrain management structure to ensure quality assurance and compliance to local, national and international standards.

of activities; and evaluate and validate project progress and outputs. The Steering Group will be comprised of:

- Project Coordinator (chair)
- Management Back Office / Project Manager
- Work Package Leaders

Project Coordinator: The Project Coordinator will be operationally responsible for the day-to-day coordination within the project (with particular attention given to the quality and timely submission of Deliverables and reports) and will act as the liaison between the Consortium and the EC. Due to unforeseen circumstance, the FashionBrain project has had several changes to the role as listed below.

- Gianluca Demartini (Initial)
- Paul Clough (Interim)
- Alessandro Checco (Final)

Management Back Office: The Management Back Office will mainly be comprised of the FashionBrain Project Manager to support the administrative and financial requirements as specified in the Grant Agreement and within the approved budget. Where appropriate, other project support staff from the coordinating institute may also assist on the project in order to adhere to the institutional standard operating procedure and quality assurance process.

- Kathryn Mackellar (Project Manager)
- The Information School (Departmental Professional Services & Technical Staff)
- Finance (Faculty Finance Administration)
- Human Resources (Faculty Human Resource Administration)
- Research Services (Institutional Research Management)

Work Package Leaders: For each WP, the assigned leader will ensure the proper planning activities and delivery of results, as well as contribution to the optimal coordination of parallel WP tasks according to the Work Plan.

- WP 1/3: Alessandro Checco - USFD
- WP 2: Mourad Khayati - UNIFR
- WP 4: Alexander Loeser - BEUTH
- WP 5: Matthias Dantone - Fashwell
- WP 6: Roland Vollgraf / Alan Akbik - Zalando
- WP 7/8/9: Alessandro Checco / Kathryn Mackellar - USFD

2.2. Communication

The FashionBrain project's internal communication strategy aims at ensuring the appropriate transparency and cooperation among partners as well as the timely generation, collection and storage of project information. The strategy includes three main methods of communication:

- E-mail exchanges
- Consortium meetings
- Document sharing

2.2.1. Mailing Lists

Day-to-day communication will be email-based. Secure electronic mailing lists have been created for the purpose of ensuring that relevant information is circulated among all key stakeholders and avoids unfortunate circumstances where someone may be missed by typing addresses manually. In order to send and receive communications, to and from the mailing list address, the stakeholder's email address must be added by the Project Manager (or Project Coordinator).

FashionBrain@sheffield.ac.uk addresses all consortium members actively involved in day-to-day coordination of project activities.

fashionbrain-admin@sheffield.ac.uk addresses all relevant administrative contacts for each of the partner institutions.

fashionbrain-tech@sheffield.ac.uk addresses all relevant technical research and support contacts for the project.

2.2.2. Meetings

The dynamic nature of the project makes meetings a necessity to boost understanding and efficiency. virtual and face-to-face meetings will take place over the lifetime of the project to monitor and plan its progress. The Project Coordinator will chair all meetings (unless otherwise agreed upon).

Steering Committee

The Steering Committee will meet on a monthly basis, by telephone/VoIP conferences and / or at (semi)annual face-to-face meetings that will coincide with scheduled project reviews or technical meetings as much as possible. Regardless of the venue, official agendas and minutes (including attendance) will be produced for all meetings and later deposited in the shared Google Drive filestore (Section 2.2.3).

In the case of face-to-face meetings, the hosting partner is responsible for the logistical organisation (in coordination with Project Coordinator and Manager),

although each institution is responsible for actually making and paying for their own travel and accommodation arrangements (unless otherwise agreed upon).

Where meetings are held external to any of the partners' facilities, the Project Coordinator and Manager will assume the meeting logistics.

Management Back Office

The Project Coordinator and Manager will meet following every Steering Committee meeting and at least once in-between (minimum twice per month) to discuss additional issues and confirm planning objectives. Additional project support staff may also attend these meetings when necessary to aid in the administrative tasks.

During critical periods during the project, the frequency of meetings may increase to once per week to ensure that progress towards administrative and technical milestones is being met.

Ad-hoc Meetings

Meetings are not limited to the above. Any member may request an additional meeting with the entire steering committee or the Project Coordinator for any number of reasons (i.e. missed monthly meeting, clarification of WP tasks, new issues). Meetings may be not be minuted depending on the nature of the discussion but will be logged by the Project Coordinator for future reference.

2.2.3. Shared Google Drive Filestore

A project working space platform has been set up for network collaboration and documents distribution, as well as data storage, using a password-protected Google Drive Enterprise account that has an integrated backup solution. Google Drive Enterprise edition gives unlimited cloud storage, Google's suite of collaborative Docs editors and administrative controls to enhance enterprise productivity and collaboration. This is managed by Project Coordinator and Manager and the cost of this solution is covered by USFD.[2]

The FashionBrain shared Google Drive filestore is structured as follows[3]:

- Administration
- BDV
- Datasets
- Deliverables

[2] USFD are satisfied that the security controls put in place by Google are sufficient to protect University data. A summary of their assessment can be found at: https://www.sheffield.ac.uk/polopoly_fs/1.254519!/file/FactSheet-DataSecurityandPrivacywithGoogle.pdf.

[3] Additional folders and subfolders can be created upon request to the Management Team.

- Description of Action &Agreements
- Dissemination
- Ethics Documents
- Interim Review
- Logos
- Meeting
- Photographs
- Publications / Conferences
- Timesheets
- Webpage Biographies
- Website Backup

Monitored access will only be granted and available to members of the project consortium using an existing google account (or newly created). In certain instances, it may be necessary to provide access to those outside of the consortium; which will be dealt with on a case-by-case basis and only following approval by the steering committee.

3. Quality Plan

All partners are responsible for the quality assurance of the FashionBrain project. The quality of project activities and outputs will be monitored by the WP leaders and the Project Coordinator according to the roles and responsibilities as defined in Section 2.

A Project Progress Assessment Plan has also been established. For the duration of the project, the Coordinator will compile and analyse data on indicators against which the progress and outcomes of each WP and the project as a whole will be monitored and assessed to evidence its short-term and long-term impact.

3.1. Logos

A project logo (or creative device) has been designed to aid and standardize project identification and branding across the consortium (Figures 3.1 and 3.2). The project logo(s) are available to partners via the shared Google Drive filestore (Section 2.2.3) and must be used on all project-related dissemination and reports.

Figure 3.1: Project logo (white).

Figure 3.2: Project logo (black).

3.2. Funding Statement

All project-related dissemination and reports must include the following statement:

The FashionBrain project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732328.

Where appropriate, the EU logo (Figure 3.3), which is available to partners via the shared Google Drive filestore (Section 2.2.3), should also be used.

Figure 3.3: EU logo.

3.3. Work Packages and Deliverables

3.3.1. Work Packages

Work Package (WP) describe the specific objectives of the action that will be achieved throughout each individual part of the project and the specific tasks that will be carried out within each to achieve the set objectives:

WP 1 will run at the beginning of the project to make sure the activities are well aligned by gathering data integration requirements and survey existing solutions to avoid replication of work and to benefit from existing data integration infrastructure and datasets.

WP 2 is the data infrastructure layer that will develop custom data integration techniques based on the results of WP1 requirement analysis and initial developments. In this WP we will also deploy our data infrastructure and develop efficient database primitives for our time series methods.

WP 3 will focus on human computation solutions for data integration in the fashion industry. We will complement WP2 data infrastructure with novel crowdsourcing techniques to deal with the limitation of machine-based data integration solutions (e.g., identification of fashion items in social media images).

WP 4 will look at data integration challenges over structured (i.e., databases) and unstructured content (i.e., textual content) by developing custom solutions for our application scenarios in WP 5 and 6. Partners will investigate how to apply deep learning techniques on CPU- and GPU-based hardware.

WP 5 is the integration and evaluation Work Package. In this WP, we will leverage methodologies developed in WP 2, 3, and 4 to deal with the social media fashion universe by annotating data and identifying upcoming fashion trends.

WP 6 is our second application Work Package together with WP5 which will use data and results from WP 2, 3 and 4 to provide textual search functionality in the domain of fashion related images, for the benefit of key stakeholders in the Fashion data value chain.

WP 7 will make sure project results are disseminated to key stakeholders outside of the project consortium including both academic and industry organizations.

WP 8 will deal with the overall project management making sure all tasks run smoothly and partners deliver results as defined in the proposal.

WP 9 will deal with ethical issues deriving from the collection, use, and secondary analysis of personal data during the project or previously.

3.3.2. Deliverables

Deliverables are additional outputs (e.g. information, special report, a technical diagram brochure, list, a software milestone or other building block of the project), identified in Annex 1 of the grant agreement, that are submitted to the EC in accordance with predetermined timing and conditions.

Template

A standardized deliverables template (https://www.overleaf.com/read/ssfytmgqvqbr) using Overleaf[1] for was introduced following the midterm review (September 2018). This collaborative writing and publishing system uses LaTeX[2] and allows project members to create, edit and share reports online while minimising version inconsistencies and software package conflicts. The template (PDF version) is attached in Appendix A.

Versioning

The versioning system will follow the Major.Minor numbering rule, similar to software versioning systems, where drafts will be 0.x (e.g. v0.1) and final versions will be numbered x.0 (e.g. v1.0).

Deliverables Review Process

Internal reviews are carried out for each deliverable before submission to the EC. Reviews may be carried out by other WP leaders, the Project Coordinator or project support staff depending on the nature and technical specification of each. WP leaders are reminded approximately 1 month in advance of the deliverable due date (usually during the steering committee meeting). Overleaf drafts are due for review approximately 2 weeks prior to submission. Drafts are shared with the Project Coordinator and Manager who then distribute to the relevant peer reviewer for comments. Comments should be provided approximately 3-5 days within receiving the draft. This will allow time for revisions by the authors and a final review by the Project Coordinator and Manager before depositing a PDF copy to the shared Google Drive filestore and submitting it to the EC.

List of Deliverables

Table 3.1 shows the deliverables planned and already achieved by the project; the deliverable leader (main partner) and submission dates (for each version). Deliverables intended for public dissemination are published on the website (with

[1] https://www.overleaf.com/about.
[2] https://www.latex-project.org/.

PDF download links for the most up-to-date version) on the Project Achievements page (Refer to D7.2 - Project Website and D7.3 - Communication Plan).

Deliverable No.	Title	Leader	Due Date	Delivery Date	Type of Deliverable	Dissemination Level
1.1 (WP1)	Survey document of existing datasets and data integration solutions	USFD	M6 JUN 17	v1-JUL 17 v2-APR 18 v3-MAY 18 v4-FEB 19	Report	Public
1.2 (WP1)	Requirement analysis document WP1	Zalando	M6 JUN 17	v1-JUL 17 v2-JUN 18 v3-FEB 19	Report	Confidential
1.3 (WP1)	FashionBrain Ontology/Taxonomy	UNIFR	M12 DEC 17	v1-FEB 18 v2-APR 18	Report	Confidential
1.4 (WP1)	Software Requirements: SSM library for time series modelling and trend prediction	Zalando	M12 DEC 17	v1-DEC 17 v2-JUN 18 v3-FEB 19	Report	Public
2.1 (WP2)	Named Entity Recognition and Linking methods	UNIFR	M18 JUN 18	v1-JUN 18	Other	Public
2.2 (WP2)	Requirement analysis document WP2	UNIFR	M8 AUG 17	v1-SEP 17 v2-APR 18 v3-MAY 18 v4-FEB 19	Report	Confidential
2.3 (WP2)	Data integration solution	MDBS	M24 DEC 18	v1-DEC 18 v2-FEB 19	Other	Public
2.4 (WP2)	Time Series Operators for MonetDB	MDBS	M24 DEC 18	v1-DEC 18 v2-FEB 19	Other	Public
2.5 (WP2)	Library of trained Deep Learning models	Zalando	M18 JUN 18	v1-JUN 18	Other	Confidential
3.1 (WP3)	A set of crowdsourcing interfaces	USFD	M12 DEC 17	v1-FEB 18 v2-APR 18 v3-MAY 18	Other	Public
3.2 (WP3)	A set of aggregation algorithms and their experimental evaluation	UNIFR	M24 DEC 18	v1-DEC 18 v2-FEB 19	Other	Public
3.3 (WP3)	Surveys design and crowdsourcing tasks	USFD	M36 DEC 19		Report	Public
3.4 (WP3)	Report on how to quantify and address ambiguity, subjectivity and other biases in crowdsourcing for fashion	Zalando	M36 DEC 19		Report	Confidential

Deliverable No.	Title	Leader	Due Date	Delivery Date	Type of Deliverable	Dissemination Level
4.1 (WP4)	Report on text joins	BEUTH	M15 MAR 18	v1-APR 18 v2-FEB 19	Report	Public
4.2 (WP4)	Demo on text joins	BEUTH	M18 JUN 18	v1-JUN 18	Demonstrator	Public
4.3 (WP4)	Relation Extraction with Stacked Deep Learning	BEUTH	M30 JUN 19		Report	Public
4.4 (WP4)	Demo on Relation Extraction with Stacked Deep Learning	BEUTH	M36 DEC 19		Demonstrator	Public
5.1 (WP5)	Scalable Crowdsourced Social Media Annotation	Fashwell	M18 JUN 18	v1-JUN 18 v2-FEB 19	Demonstrator	Public
5.2 (WP5)	Product Taxonomy Linking	Fashwell	M24 DEC 18	v1-DEC 18 v2-FEB 19	Demonstrator	Public
5.3 (WP5)	Early Demo on Fashion Trend Prediction	UNIFR	M18 JUN 18	v1-JUN 18 v2-FEB 19	Demonstrator	Public
5.4 (WP5)	The classification algorithm and its evaluation on fashion time series	UNIFR	M18 JUN 18	v1-JUN 18 v2-FEB 19	Other	Public
5.5 (WP5)	Demo on Fashion Trend Prediction	UNIFR	M36 DEC 19		Demonstrator	Public
6.1 (WP6)	Dataset of annotated images	Zalando	M12 DEC 17	v1-APR 18 v2-FEB 19	Other	Confidential
6.2 (WP6)	Entity linkage data model	Zalando	M18 JUN 18	v1-JUN 18	Report	Confidential
6.3 (WP6)	Early Demo on textual image search	Zalando	M18 JUN 18	v1-JUN 18 v2-FEB 19	Demonstrator	Public
6.4 (WP6)	Enriched image dataset	Zalando	M24 DEC 18	v1-DEC 18 v2-FEB 19	Other	Confidential
6.5 (WP6)	Demo on textual image search	Zalando	M33 SEP 19		Demonstrator	Public
7.1 (WP7)	Project factsheet	Zalando	M1 JAN 17	v1-FEB 17 v2-APR 18 v3-FEB 19	Report	Public
7.2 (WP7)	Project Web site	USFD	M1 JAN 17	v1-MAR 17 v2-APR 18 v3-FEB19	Websites, patents filling, etc.	Public
7.3 (WP7)	Communication Plan	USFD	M6 JUN 17	v1-JAN 18 v2-JUL 18 v3-FEB 19	Report	Public
7.4 (WP7)	Showcase specification	Fashwell	M6 JUN 17	v1-JUL 17 v2-APR 18 v3-FEB 19	Report	Public
7.5 (WP7)	Draft business plan	USFD	M18 JUN 18	v1-JUN 18 v2-FEB 19	Report	Confidential
7.6 (WP7)	Showcase specification and dissemination summary	Fashwell	M36 DEC 19		Report	Public
7.6 (WP7)	Final business plan	USFD	M36 DEC 19		Report	Confidential

Deliverable No.	Title	Leader	Due Date	Delivery Date	Type of Deliverable	Dissemination Level
8.1 (WP8)	Project Work Plan & Quality Plan	USFD	M1 JAN 17	v1-FEB 17 v2-APR 18 v3-MAY 18 v4-FEB 19	Report	Confidential
8.2 (WP8)	Data Management Plan	USFD	M6 JUN 17	v1-JAN 18 v2-APR 18 v3-FEB 19 v4-FEB 19	ORDP	Confidential
9.1 (WP9)	GEN - Requirement No. 1	USFD	M3 MAR 17	v1-JUN 17 v2-APR 18 v3-JUN 18 v4-FEB 19	Ethics	Confidential
9.2 (WP9)	GEN - Requirement No. 2	USFD	M3 MAR 17	v1-MAR 17 v2-APR 18 v3-MAY 18 v4-FEB 19	Ethics	Confidential
9.3 (WP9)	GEN - Requirement No. 3	USFD	M3 MAR 17	v1-MAR 17 v2-APR 18 v3-FEB 19	Ethics	Confidential
9.4 (WP9)	GEN - Requirement No. 4	USFD	M3 MAR 17	v1-MAR 17 v2-APR 18 v3-FEB 19	Ethics	Confidential

Table 3.1: FashionBrain deliverables.

4. Risk Management Plan

The risk management plan contains an analysis of likely risks with both high and low impact, as well as mitigation strategies to ensure the project is still able to achieve (or review / modify) its objectives should problems arise.

4.1. Management of Intellectual Property

Explicit rules on how to access Background and Foreground intellectual property, and how to ensure the protection of Intellectual Property Rights (IPR) have been established and addressed within the Consortium Agreement (CA) which has been agreed and signed by all project partners (and their respective institutions). Project partners shall grant royalty-free access to Background and Foreground intellectual property where it is technically essential for carrying out the tasks of the project. Intellectual property which is explicitly excluded from the project, will be identified in the CA. However, any and all IPR in any Background and/or Foreground intellectual property made available by a project partner in the course of the project shall remain vested exclusively in that project partner. Foreground intellectual property shall be owned by the project partner carrying out said work. If any Foreground intellectual property (or resulting invention) is created jointly by at least two project partners such that the contribution of each of the project partners cannot be distinguished from that of the other, such work (and patent applications filed) will be jointly owned by the contributing project partners.

4.2. Management and Protection of Knowledge

The FashionBrain project will adhere to the following policy (https://www.sheffield.ac.uk/library/openaccess/what) with respect to open access publications, including both green (self-archiving / free) and gold (journal website / fee) routes. Within the FashionBrain project, the process is as follows:

- Details of all research outputs will be provided to the Project Coordinator / Manager (USFD).
- A copy of the final accepted peer-reviewed manuscript (for all journal articles and conference papers accepted for publication within the scope of the FashionBrain project), along with evidence of the date of acceptance, will be deposited in a personal repository on-line (e.g., personal or university website);

within the FashionBrain shared Google Drive filestore; and on the website, as soon as possible (and at the latest, upon publication).

- Other types of research outputs (e.g., book chapters, reports, creative outputs, and grey literature) will also be deposited in a personal repository on-line (e.g., personal or university website); within the FashionBrain shared Google Drive filestore; and on the website, as soon as possible and / or where agreement to do so can be reached with the respective publisher.

Although the publication of outputs as open access is a key consideration within the project as a condition of funding, the coordinating partner, USFD, is committed to the principles of open access to research outputs in order to ensure that its world-class research is available as widely as possible to maximize its impact within both the academic research community and more widely within society.[1]

4.3. Conflict Resolution

The potential for conflict within a multidisciplinary project is assumed and expected to be high as it involves individuals from different backgrounds and organisational cultures working together to complete a complex set of tasks. Day-to-day conflicts may relate to differences in priorities, resource allocation, methodology choices, ways of working, or expectations of results. The conflict resolution mechanism described below aims to minimize the impact of any disagreement by ensuring a timely resolution.

- The primary mechanism for decision-making within the project will be by consensus.
- Conflicts within a WP shall, in the first instance, be mediated by the Project Coordinator and if necessary, by the Steering Committee.
- Conflicts between the Project Coordinator and a Consortium Member, or between Consortium Members, shall be referred directly to the Steering Committee for mediation (although the Coordinator may first be invited to mediate disputes between Consortium Members).
- In case of non-performance of any of the partners, the Steering Committee shall have the power to remove the offending partner from the project consortium by a vote of unanimity minus one.[2]

[1] http://librarysupport.shef.ac.uk/OpenAccessPositionstatement.pdf.

[2] In such circumstances the provisions of Grant Agreement guidelines will apply, as well as relevant non-conflicting provisions made in the CA.

4.4. General Ethics Requirements

The FashionBrain project has decided to follow a unified Ethics procedure 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. 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.

Deliverable 9.1 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.

Deliverable 9.2 summarizes 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 self-assessment (available in the appendix of D9.2) to certify that the current modus operandi is in agreement with the EU H2020 Ethics and data protection regulations, spirit and best practices This document also includes feedback and advice from ethics experts from the University Research Ethics Committee (UREC) (USFD).

Deliverable 9.3 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. In the cases in which public data are obtained without consent (and instead, implicit consent has been given on another platform when the users agreed to its T&Cs), 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 which allows the public to request an opt-out from FashionBrain data collection.

The Consent Manager has been published on the project website which will be advertised 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 has 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.

Deliverable 9.4 contains the informed consent forms used in the project verifying that research participants are provided with a full explanation of the project, what their participation entails, what type of data will be collected and any risks that may be involved. By conveying this information and obtaining their express permission, the project is ensuring compliance with EU data protection rules.

4.5. Data Management Plan

The Data Management Plan (DMP) of the FashionBrain project outlines what data will be generated or collected, how data will be managed, the standards in use, the workflow to make the data accessible for use, reuse and verification, and which plans for data sharing and preservation exist ensuring that data are well-managed.

Deliverable 8.2 outlines the initial DMP, which is in line with the H2020 guidelines for data management plan creation and identifies the initial classes of datasets of the project. The document presents an overview of the FashionBrain integrated architecture (FaBIAM); summarises the data used and collected by the consortium; describe the data processing workflow, discusses the findability, accessibility, interoperability and reuse (FAIR) of the data; and presents the resources allocation for the management of the data. It also addresses the security of the data as well as the ethical aspects of the data collection. Links to institutional and consortium policies are also provided.

The DMP has been reviewed by the Research Data Management team (USFD) and the FashionBrain Ethics Committee.

4.6. Go-to-market Risks

Risks specific to the Fashionbrain go-to-market business plan, including losses and failures of marketing; decreased demand for services; small concentration of revenue sources; channel conflict due to white licensing of technologies; lack of resources / loss of key personnel; information / cyber security; and inaccurate budget assumptions, have also been considered. The process of identifying, mitigating, sharing or transferring go-to-market risks will be continually assessed during the project's lifetime. Specific details with respect to each risk and strategies to manage those risks can be found in Deliverable 7.5 - Draft Business Plan.

4.7. Critical Risks

4.7.1. Foreseen Risks

1 Required recruiting may delay effective start-up of project.
Impact: Medium
Probability: Medium
Links to WP: 1 - 8
Risk Mitigation Measure: Organisations are planning many key activities with permanent staff so that conceptual work can begin immediately.
State of Play: Risk has matierialised resulting in redistribution of person months (as reported in Midterm Review).

2 Technical interdependencies may lead to duplication or incompatibility of work.
Impact: High
Probability: Medium
Links to WP: 1 - 8
Risk Mitigation Measure: Mitigated by the definition of the Project Steering and appointment of a Technical Coordinator.
State of Play: Risk has not materialized.

3 Computer Vision Accuracy. The computer vision algorithm developed in WP2 are not accurate or fine grained enough.
Impact: Medium
Probability: Medium
Links to WP: 5
Risk Mitigation Measure: Fashwell and Zalando have a strong experience in that field. Both companies are defining clear goals together and are planning to have regular meetings.
State of Play: Risk has not materialized.

4 No funding for Fashwell. Fashwell is a young startup, the company is financed by VCs and there is always a risk that the company doesn't find new capital.
Impact: Medium
Probability: Low
Links to WP: 2, 3 & 5
Risk Mitigation Measure: Fashwell is going to update the consortium about potential liquidation problems on a regular basis.
State of Play: Risk has not materialized.

5 Instagram blocks API. Instagram is the biggest content creation site. We are planing to use their API to download the fashion content.
Impact: Medium
Probability: Low
Links to WP: 2, 3 & 5
Risk Mitigation Measure: The partners are downloading only the publicly

available images. If Instagram is closing the API, the partner could write a crawler to download the public available data. A budget for stock photos (included in Zalando other direct costs) shall remedy the critical risk that Instagram blocks their API (cf. table "Critical Risks for Implementation"). In that case stock photos serve as a backup source for fashion photos. Also in case that with photos from social media streams sufficient diversity cannot be achieved, data can be enriched with stock photos. If necessary, this budget would allow, for example, for a subscription at Getty Images for 18 month (which should cover the image data consolidation phase of the project).
State of Play: Risk has not materialized.

6 Difficulties in integrating components with applications. With regard to a number of partners involved, different partner development experiences and distributed development environment one or more of the following may occur: delays in components and modules delivery, low quality of the components.
Impact: High
Probability: Medium
Links to WP: 2 - 6
Risk Mitigation Measure: The partner in charge of technological integration as well as other industry partners has successful experience in large-scale collaborative projects. The application of agile method, promoting rapid development, deployment and continuous integration and evaluation will allow to detect and solve early any issue with integration. A set of guidelines will be released as soon as possible to guide researchers and developers through the agreed development process.
State of Play: Risk has not materialized.

4.7.2. Other Risks

1 In addition to the technical risks above, the project will face various other risks including:

- Partner problems (e.g. a partner is under-performing or a key partner is leaving the project)
- Expertise risks (e.g. a key person with a specific expertise is leaving the project)
- Project execution risks (e.g. key milestones or critical deliverables are delayed)
- Agreement risks (e.g. Consortium partners cannot agree because of different interests)

Links to WP: 1 - 8
Risk Mitigation Measure: In large, complex projects, problems will inevitably arise and therefore, risks need to be clearly identified and assessed, and the project must be prepared for remedial actions. This forms part of our project

management approach (Section 2. The project management team will create and monitor a risk plan, identify threats and take necessary countermeasures. Contingency planning is built into the structure of the project both at high level as well as in individual tasks and work packages. At the overall project management level we have two main mechanisms:

- The governance structure of the project.
- The Consortium Agreement, which defines the responsibilities of each party and the rules for taking action against deviations.

State of Play: Risk has not materialized.

4.7.3. Unforeseen Risks

1 The recent changes in the EU data privacy regulations and the API changes from Instagram and Facebook do not allow us to perform Task 5.2 (Social Media Post Linking). We are not permitted to permanently store the data we collect and especially we do not have the permission to link one data subject from one platform to the same data subject on another platform.
Date Risk Added: 20 August 2018
Links to WP: 5
Risk Mitigation Measure: Fashwell and UNIFR would like to propose a new task, which would completely replace Task 5.2, by defining an attribute taxonomy (e.g. color:red, neck-type:v-neck) of fashion-products and extracting these attributes automatically using computer vision and deep learning techniques. This would have two significantly advantages: (1) we would not need to store the actual images of products, but can represent the images using attributes; and (2) the extracted attributes have the potential to significantly improve Tasks 5.3 and 5.4.

2 A crawl of 1.5 billion URLs of the .DE Web Dataset cannot be published for the following reasons. Firstly, we foresee changes in the European copy right policy in text and data mining. Enclosed is the text of the JURI committee for the text and data mining exception (Article 3) that was confirmed first by the subcommittee and later rejected by the European parliament. It will be decided at the end of September 2018. Basically, even a part of a URL published for the task of text and data mining can represent a copyright infringement.[3] [4] Therefore, a corpus of 1.5 billion URLs will very likely contain copyright protected material owned by the holders of the domain names and the creators of the URLs. Hence, we would have to ask hundreds of millions domain holders, which is infeasible. Secondly, the corpus was gathered in a

[3] https://www.communia-association.org/wp-content/uploads/2018/05/24052018-Draft-CA-on-Art.-3-v1.pdf.

[4] https://en.wikipedia.org/wiki/Directive_on_Copyright_in_the_Digital_Single_Market.

project with six other national partners in 2013. We have not obtained the rights from these partners to share this dataset publicly. Because of these two reasons we cannot publicly distribute this dataset.
Date Risk Added: 20 August 2018
Links to WP: 1 & 4
Risk Mitigation Measure: We do not want to carry out work using data sets whilst the political situation is unclear. Work within the FashionBrain project is not biased by this decision since we can train our statistical models on alternative data sets, as explained in more detail for the relevant deliverables below.

- D4.2 & D4.1 (Entity Linkage Text2Table)
  We use the WebNLG data set.
  Justification: There exists no data set for Text Joins in the fashion domain yet. Therefore we searched for data sets with similar characteristics. One Data set is WEBNLG (see D4.1 and D4.2) that uses idiosyncratic entities (as used in the fashion domain), such as products, comic figures or buildings, or companies, persons and locations (used as well in the fashion domain).
- D4.3 & D4.4 (Event and Relation Extraction)
  We will measure our work on existing event / relation extraction data sets (i.e. data sets that entail Organization/Brand and Celebrity/Person information). Example corpora are:
  - ACE 2005
  - Semeval 2010 including ProductProducer, WorksIn etc.
  - Ontotext 5
  - KBP 2010

Relation extraction is a very hot area. Therefore we currently will investigate these and other datasets for training a neural relation extractor wrt. amount of labeled data, variance, density etc. All of the above mentioned datasets contain some relevant relationship types for FashionBrain, such as information about celebrities, about vendors and brands and also which vendor produces which brands. Our work in D4.3 will include a justification for one of the above or/and additional datasets not mention here depending on the match to the task in D4.3 and D4.4.

A. Appendix

A PDF version of the Deliverables overleaf template as described in Section 3.3.2 can be found below.

Horizon 2020

Understanding Europe's Fashion Data Universe

Title of This Deliverable (Title Case)

Deliverable number: DX.X

Version X.X

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	Public (PU)
Contractual Delivery Date	XXX
Actual / Resubmitted Delivery Date	XXX
Number of pages	15, the last one being no. 9
Authors	Author1, Author2 . . . - Institution1 Author1, Author2 . . . - Institution2
Peer review	Authors - Institution

Change Log

Version	Date	Status	Partner	Remarks
X.1	DD/MM/YYYY	Draft	USFD	
X.0	DD/MM/YYYY	Final	USFD	

Deliverable Description

Insert the description as written in the DoA.

Abstract

Provide a short description, highlight the topics considered in the deliverable. The abstract must be self-contained, without abbreviations, footnotes, or references. It should be a microcosm of the full deliverable. The abstract must be between 150-300 words. The abstract should not contain displayed mathematical equations or tabular material, nor abbreviations or acronyms without expansion.

Ensure that your abstract reads well and is grammatically correct. It should

- be written in language appropriate for the target audience (simple)
- consist of short, concise paragraphs
- only include material present in the main report

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AMT	Amazon Mechanical Turk (www.mturk.com), micro-task crowdsourcing platform
CF	Crowdflower (www.crowdflower.com), micro-task crowdsourcing platform

1. Introduction

Insert an introduction that explains

- The background of this deliverable (why the problem is important, how it has been tackled in the past).
- How the problem is tackled.
- The structure of the report (in Section 2 we present the related work, in Section... etc).

If the deliverable is very small (dataset etc), you might not need to write a lot here but please do not omit next subsection 1.1 (Scope of this Deliverable).

1.1. Scope of This Deliverable

In this section you should explain how this deliverable is related with the other deliverables. **This section should not be omitted!** It should identify the objectives of the corresponding WP and Task(s) as defined in the DoW, and how are those objectives addressed by the work that will be reported in this deliverable.

It is important to be exhaustive and please try to explain the reasons of redundancies with other partners/deliverables (if any) and how this fits w.r.t. to the core technologies presented in the website. This is an example of the structure of this section:

Scope This deliverable continues the work of DX.X and provides bla bla.

Dependencies This deliverable is using the results of deliverables DX.X, and in particular bla bla. It should link the work reported here to other WPs/Tasks.

Contribution This deliverable will contribute to tasks TX.X.

Collaborations show the collaborations both within this Task/WP and with other WP(s)

Datasets Talk about datasets w.r.t. the FashionBrain data lake, benchmark datasets (if any); refer to D2.2. **important!**

See D4.2 for a short example.

2. Related Work

If appropriate, please insert the related work here.

3. Instructions - How to Use This Template

If you want to use Overleaf, make a copy of this template by clicking on Menu/Copy Project. If you don't have an Overleaf account you can still copy the source and use it locally (Menu/Source), but you will not be able to update the glossary to the common shared one. In that case please let the programme coordinator know about this so we can update the file with your glossary.

Please modify the commands that starts with `% EDIT THIS` in `main.tex` and `title.tex`.

Start a main section with the `\chapter{}` command and the subsequent subsections with `\section{}`, `\subsection{}` and `\subsection*{}` if no number is needed. **Please Capitalize the Title!** (see https://www.bkacontent.com/how-to-correctly-use-apa-style-title-case/).

3.1. Abbreviations and Acronyms

Please **make an effort** to define all abbreviations and acronyms! Remember to keep updated the glossary file (https://www.overleaf.com/8548773748ymqchwgtzxts)[1] and to click refresh here before compiling.

From the external file, only the acronyms you actually use will be shown in the table of contents of this report.

Simple Acronyms

To make a defined acronym appear in the list, use the command `\acronymused{AMT}` in `main.tex` before `\begin{document}`.

Advanced Acronyms (usually not needed, can be skipped)

You can use acronyms obtained from the external glossary file like this (automatic expansion first time that is used in the document): Amazon Mechanical Turk (www.mturk.com) (AMT), Crowdflower (www.crowdflower.com) (CF), this is the second time: AMT, you can also have plurals AMTs and the long version if needed again Amazon Mechanical Turk (www.mturk.com) (AMT), and the long version without the acronym in parentheses Amazon Mechanical Turk (www.mturk.com).

[1] have a look at the glossary source for more details on how to use acronyms https://www.overleaf.com/project/5be8a444db30c7318938db4c.

For more information see http://mirrors.ctan.org/macros/latex/contrib/acronym/acronym.pdf.

3.2. Reviews and Track Changes

If you are using overleaf you can click on "Review" on the top right toolbar and add comments. Clicking on history you can access to the recent changes. If track changes functionality is needed, please contact the programme coordinator. If you don't use any acronym, the list of acronym will be not generated.

3.3. Citations

Use the bib file `references.bib` to cite like this:[3], [2], Bellalta et al., [1], Bellalta et al. [1]. You can use google scholar export to bibtex to get the snippet. If you have citation that you want to use in the future, you can add them to the file `global.bib` here: https://www.overleaf.com/project/5be8a444db30c7318938db4c. If you put your references in the global file you should not put them in the `references.bib` file (duplicates will throw an error).

3.3.1. Subsection Example

3.3.1.1. Sub-subsection Example

4. Instructions - Tables and Figures

4.1. Tables

You can import tables from spreadsheets or google docs using https://www.tablesgenerator.com/latex_tables. See Table 4.1 for an example of a table.

	Item	
Animal	Description	Price ($)
Gnat	per gram	13.65
	each	0.01
Gnu	stuffed	92.50
Emu	stuffed	33.33
Armadillo	frozen	8.99

Table 4.1: Table Example.

If you use tables, they will automatically compile and a list of tables will appear in the Table of Contents. If you don't use tables, the list of tables will not appear.

4.2. Images

Upload images in the images folder, see the following for an example. You can see examples of figures in Figure 4.1 and 4.2.

Figure 4.1: Example of figure, logo.

If you use figures, they will automatically compile and a list of figures will appear in the Table of Contents. If you don't use images, the list of images will not appear.

Figure 4.2: Example of figure, flag.

5. Conclusions

Summarise and conclude your work including future directions.

Bibliography

[1] Boris Bellalta, Alessandro Checco, Alessandro Zocca, and Jaume Barcelo. On the interactions between multiple overlapping wlans using channel bonding. *IEEE Transactions on Vehicular Technology*, 65(2):796–812, 2016.

[2] OLSRd. OLSRd – a routing protocol for IP-based mesh networks, September 2015. http://www.olsr.org.

[3] David Palma and Marilia Curado. Onto scalable ad-hoc networks: Deferred routing. *Computer Communications*, 35(13):1574 – 1589, 2012. ISSN 0140-3664. doi: http://dx.doi.org/10.1016/j.comcom.2012.04.026. URL http://www.sciencedirect.com/science/article/pii/S014036641200151X.

A. Appendix

Include documents, datasets, etc. that are relevant to the report.