Horizon 2020

Understanding Europe's Fashion Data Universe

Showcase Specification and Dissemination Summary

Deliverable number: D7.6

Version 1.0

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	Public (PU)
Contractual Delivery Date	31/12/2019
Actual Delivery Date	30/12/2019
Number of pages	48, the last one being no. 40
Authors	Alan Akbik, Josip Krapac - ZALANDO Alessandro Checco - USFD
Peer review	Ying Zhang - MDBS Benjamin Winter, Tom Oberhauser - BEUTH

Change Log

Version	Date	Status	Partner	Remarks
1.0	30/12/2019	Final	Zalando	

Deliverable Description

This deliverable will present the produced promotion and dissemination material, demonstration workflows, and the fully functional data integration infrastructure ready to be demo-able to the public also including screencasts (as indicated in T7.2) . We will grant the Commission the right to use the Showcase for its own dissemination and awareness activities (including Web based and electronic publications) after the completion of the project. The Showcase will feature a meaningful subset (software, data, etc.) of the functionality characterizing the project demonstrator(s) arrived at, along with relevant copyright notices and contact information, and suitable installation aids and run-time interfaces. We will also report about project activities undertaken to support standardisation of project results and collaboration with other projects and relevant initiatives as well as the results of reaching-out by means of press, social media, open-source communities using demos, use cases, and benchmark results realized during the project. As planned in T7.1, we will report on our contribution to the Big Data Value PPP activities.

Abstract

This deliverable presents the produced promotion and dissemination material, demonstration workflows, and the fully functional data integration infrastructure ready to be demo-able to the public. We grant the Commission the right to use the Showcase for its own dissemination and awareness activities (including Web based and electronic publications) after the completion of the project. The Showcase features a meaningful subset of the functionality characterizing the project demonstrators, along with relevant copyright notices and contact information, and information on the run-time procedures. We also report about project activities undertaken to support standardisation of project results and collaboration with other projects and relevant initiatives as well as the results of reaching-out by means of press, social media, open-source communities using demos, use cases, and benchmark results realized during the project. We also report on our contribution to the Big Data Value PPP activities.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

BDV	Big Data Value
BDVA	Big Data Value Association
CNN	Convolutional Neural
FaBIAM	FashionBrain Integrated Architecture
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
ML	Machine Learning
NLP	Natural Language Processing

1 Introduction

The FashionBrain project provides solutions for data integration in the fashion industry. This deliverable presents the set of demonstrations that can be used to showcase the project. Moreover, we report in this deliverable the dissemination results for the project.

A set of demonstrators and promotional documents has been packaged. These can be used by the Commission for its own dissemination and promotional activities (including web based and electronic publications) after the completion of the project.

The showcase is an evolution of the demos presented during the project, with the notable difference that in this form, the solutions presented are ready to be distributed in an interactive or static form (e.g. web sites, videos or documents) and consumed by a layman, without the need of a technical demonstrator nor a deep understanding of the input required.

1.1 Scope of This Deliverable

This deliverable is related to D7.4, that contains extended information about the specifications and architectural designs of the showcase.

For more information about the details of each part of the showcase, we refer to: (i) D4.2, D4.4 and D2.1 for entity recognition and relation extraction; (ii) D6.3 and D6.5 for end-to-end search; (iii) D5.2 for image recognition; (iv) D2.3 and D2.4 for FashionBrain Integrated Architecture (FaBIAM) and time series operators in MonetDB.

The rest of the document is structured as follows: Section 2 presents the showcases, Section 3 reports the dissemination efforts in terms of KPIs, project achievements, sponsorship and BDVA participation, and Section 4 concludes the document.

2 Showcases

The main solutions we showcase are:

Entity Recognition and Relation Extraction. We demonstrate how complex fashion corpora (e.g., fashion blogs and news) can be automatically analysed and annotated by our tools.

Shop the Look. We demonstrate the core functionalities that allow FashionBrain to detect and recognize fashion items like t-shirts, bags, etc., from images, and return the most similar products from a specific data set.

End-to-end Search. We demonstrate the FashionBrain ability to match textual search queries to a ranked list of fashion products, by using a mixture of computer vision methods and Natural Language Processing (NLP) to create a smart search index, able to cope with complex queries and multilinguality.

FashionBrain Integrated Architecture (FaBIAM). We provide a showcase and documentation of FaBIAM, a MonetDB-based architecture for storing, managing and analysing of both structured and unstructured data. This architecture allows the implementation of in-database analytical solutions with Machine Learning (ML).

For details regarding the architecture and technologies used, we refer to D7.4.

2.1 Showcases Structure

The showcases are available in the FashionBrain website. All of the datasets needed to run the demos are hosted in the cloud and can downloaded. Moreover, the solutions showcased are be presented in an interactive way.

Reliability and Maintenance

The FashionBrain data integration solutions require access to high performance computers and complex interaction between multiple services. Providing a long lasting offline solution to the Commission would be impractical, as it would require the duplication of a very expensive infrastructure.

For this reason, and to guarantee a reliable access to the promotional material, we provide an interactive version of the solutions (as explained in detail in D7.4), pre-applied on a meaningful subset of the data used during the project, as well as a set of static screencasts and webpages that can be reused by the Commission to promote FashionBrain solutions offline.

This solution guarantees the maximisation of reliability and the minimisation of the effort needed to maintain the information provided after the end of the project.

Figure 2.1: Showcases homepage.

Software Requirements

The only software required is a modern browser and a PDF reader.

Hosting

The showcase can be downloaded and deployed using this link: https://github.com/AlessandroChecco/FashionBrain-Showcase.

2.2 Showcases Homepage

The homepage of the showcase website is shown in Figure 2.1. From this page it will be possible to reach the 4 showcases: Relation Extraction, Shop the Look, End-to-End search, and Integrated Architecture.

2.3 Relation Extraction

Given a full text query such as "Tom is obsessed with Kimono jackets", our approach extracts the type of relation depicted in that sentence, "Wants/Likes" in this case.

Additionally it extracts the entities that make up this relation.

The underlying model employs a Hierarchical Reinforcement Learning approach in which the outer policy predicts the relation type sequentially for each token, while the inner policy works similarly for the entities. The embedding network for full text queries bidirectional Long-Short-Term-Memory (LSTM) unit that works with GloVe embeddings. For each relation type a separate Linear network is learned, that predicts the entities. For more information, we refer to D4.4.

The entire architecture is trained “end-to-end” to minimize a similarity cost function with supervised learning over the crowd-sources fashion corpus created by University of Sheffield. The model itself is implemented in Pytorch and Python.

A screenshot of the showcase is shown in Figure 2.2: for each text fragment, the recognised entities are highlighted, and the relation is indicated in the bottom section.

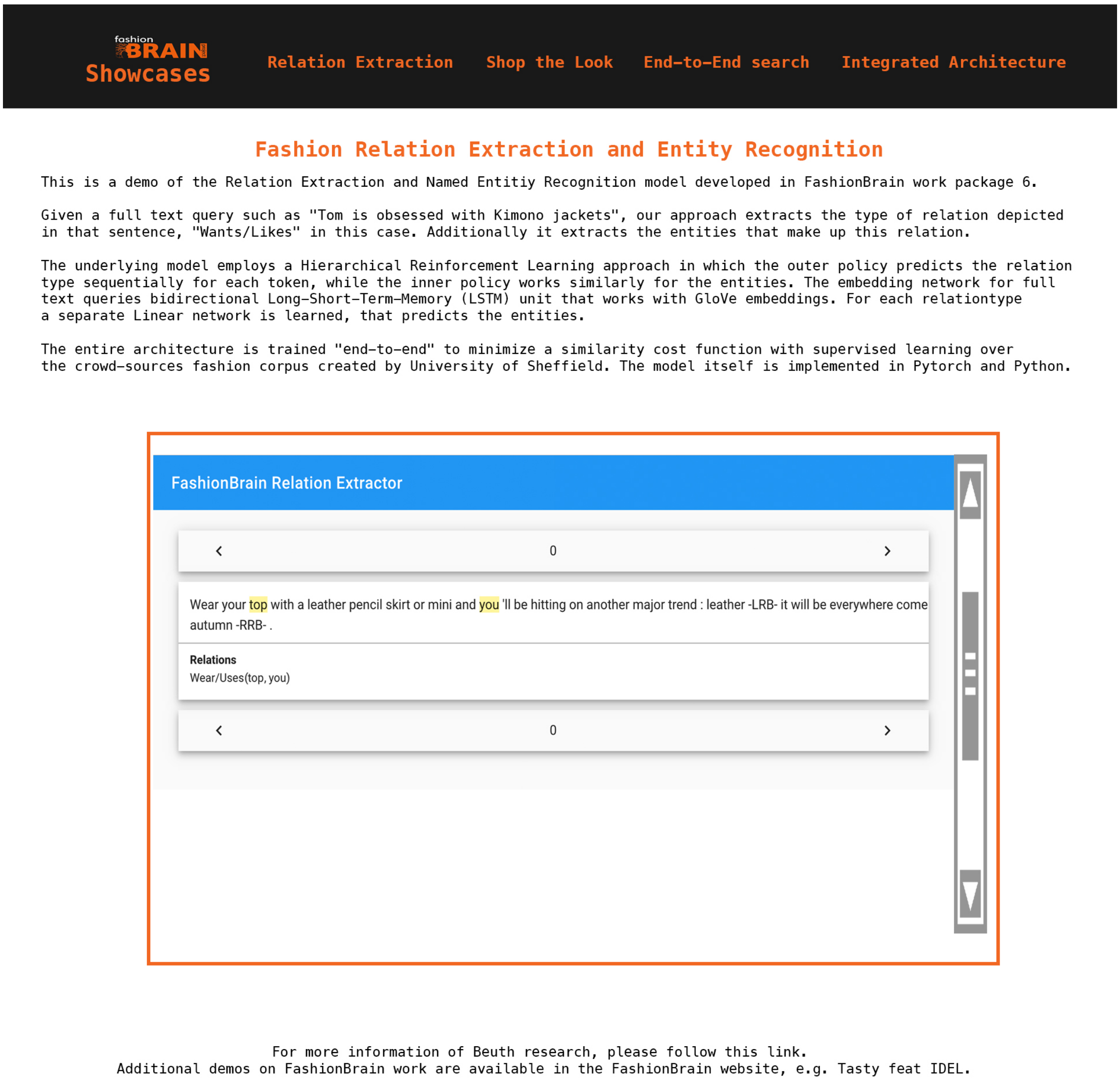


Figure 2.2: Relation Extraction.

2.4 Shop the Look

Social media and influencers have transformed the fashion landscape by becoming an advertising and marketing source for brands and retailers. In this showcase, we delved a step deeper into the products we see in images to extract attribute information about every product in the form of text labels. It may be relevant for a shopper to see a similar sneaker to that they see in social media image, however it may be even more interesting for a retailer to understand deep levels of product information in the form of text that could be aggregated and understood for trend scouting, seasonal buys or understanding from a much more concrete perspective

- what products are popular in images on social media. The product tagging technology does just this, it analyzes an image to assign keywords and phrases to products. We considered a group of publicly available images and matched it with Zalando products.

The extracted attributes are either general for all fashion objects or specific for one or several fashion product categories. For every category there are different attribute tags that are relevant. Take for example tops: tops have silhouettes and length whereas a shoe is described by a heel height, fastener and ankle height. On the other hand, color, pattern and occasion are general attributes that are relevant for any fashion product. We have built a classifier which is able to automatically extract such tags from images. We will get into how we collect data and train the classifier to perform such task but for now take a look at the attributes grouped by category that are made available for tagging:

General attributes: color, pattern, occasion, aesthetic

Tops&Blouses: neck type, sleeve length, silhouette, type, length

Pants&Trousers: type, length, cut, denim detail

Skirts: silhouette, length

Dresses: silhouette, length, sleeve length, neck type

Jackets&Coats: type, length, sleeve length, neck type

Bags: type, strap type, size, closer, leather type

Eyewear: style, frame

Shoes: type, heel type, heel height, toe type, ankle height, fastener

A screenshot of the showcase is shown in Figure 2.3: for each text image, the recognised products are highlighted, with the corresponding category and links to the Zalando webpage of the most similar product.

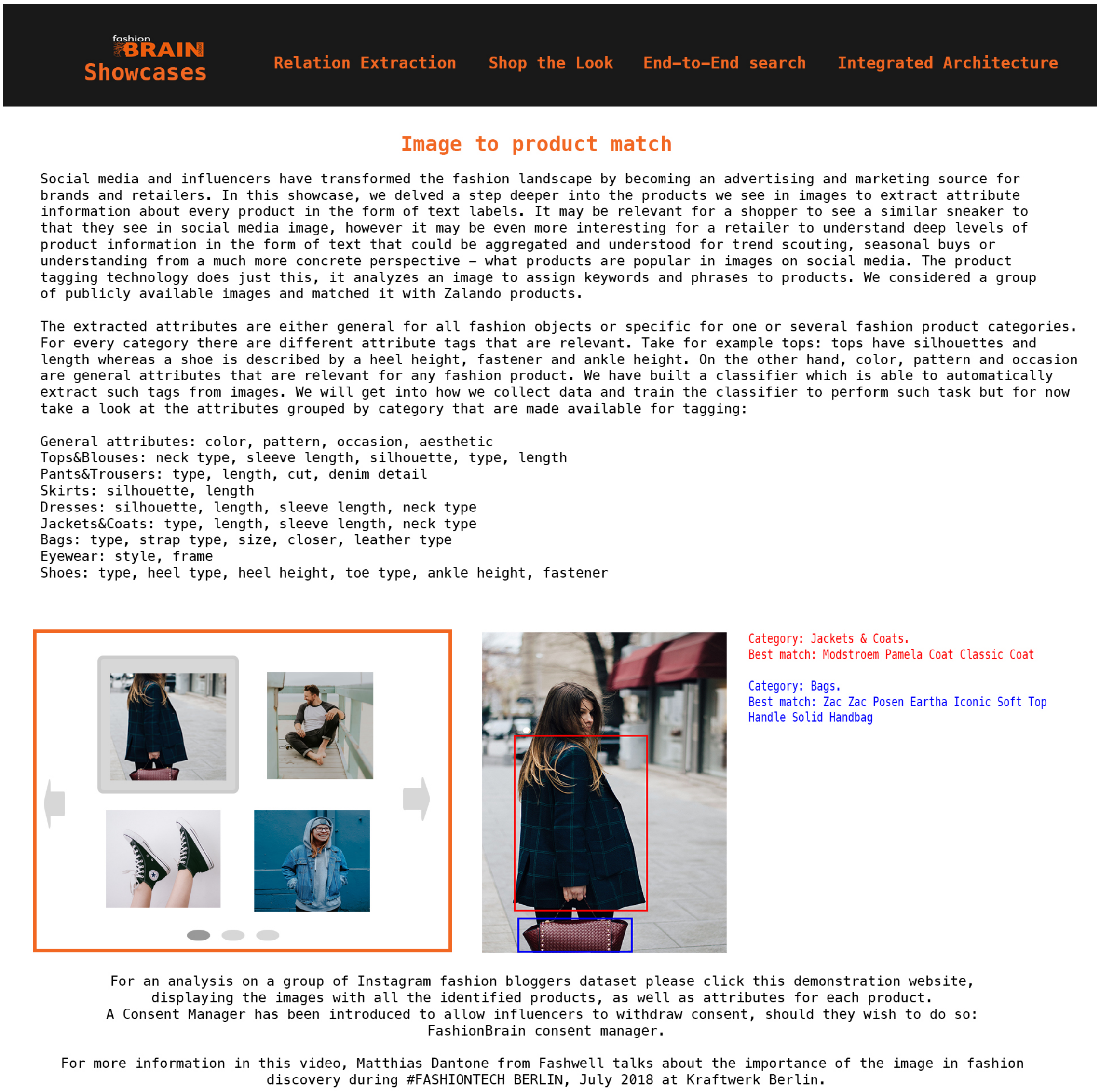


Figure 2.3: Shop the Look.

2.5 End-to-End Search

Given a full text query such das "rotes Kleid" (eng. "red dress"), our approach retrieves matching product images. Ddifferent queries can be tested in the demo and visually inspect retrieved images. This demo retrieves images from the test split of the Feidegger dataset (10% of data, 879 images) and is trained specifically for queries in German language, but the model is scalable to multilingual queries using the crosslingual embeddings supplied in Flair and multilingual datasets.

The underlying model employs a "two-tower" architecture in which each tower embeds one modality (i.e. full text queries and product images) into a shared embedding space. The embedding network for full text queries in the demo model is a bidirectional Gated Recurrent Unit (GRU) on top of pre-trained character-based LSTM (Long Short-Term Memory) embeddings for words. The embedding network for images is a shallow (3-layer) Convolutional Neural Net (CNN).

The entire architecture is trained "end-to-end" to minimize a similarity cost function with supervised learning over paired image-text datasets. For the demo model, we used a training split of Feidegger (80% of all data, 7034 images) which is open source and thus publicly available, allowing reproduction of our results. The model is implemented and trained in Flair framework.

A screenshot of the showcase is shown in Figure 2.4: for each text query, the best matches are shown. Moreover, the queries can be refined and the search can be updated.

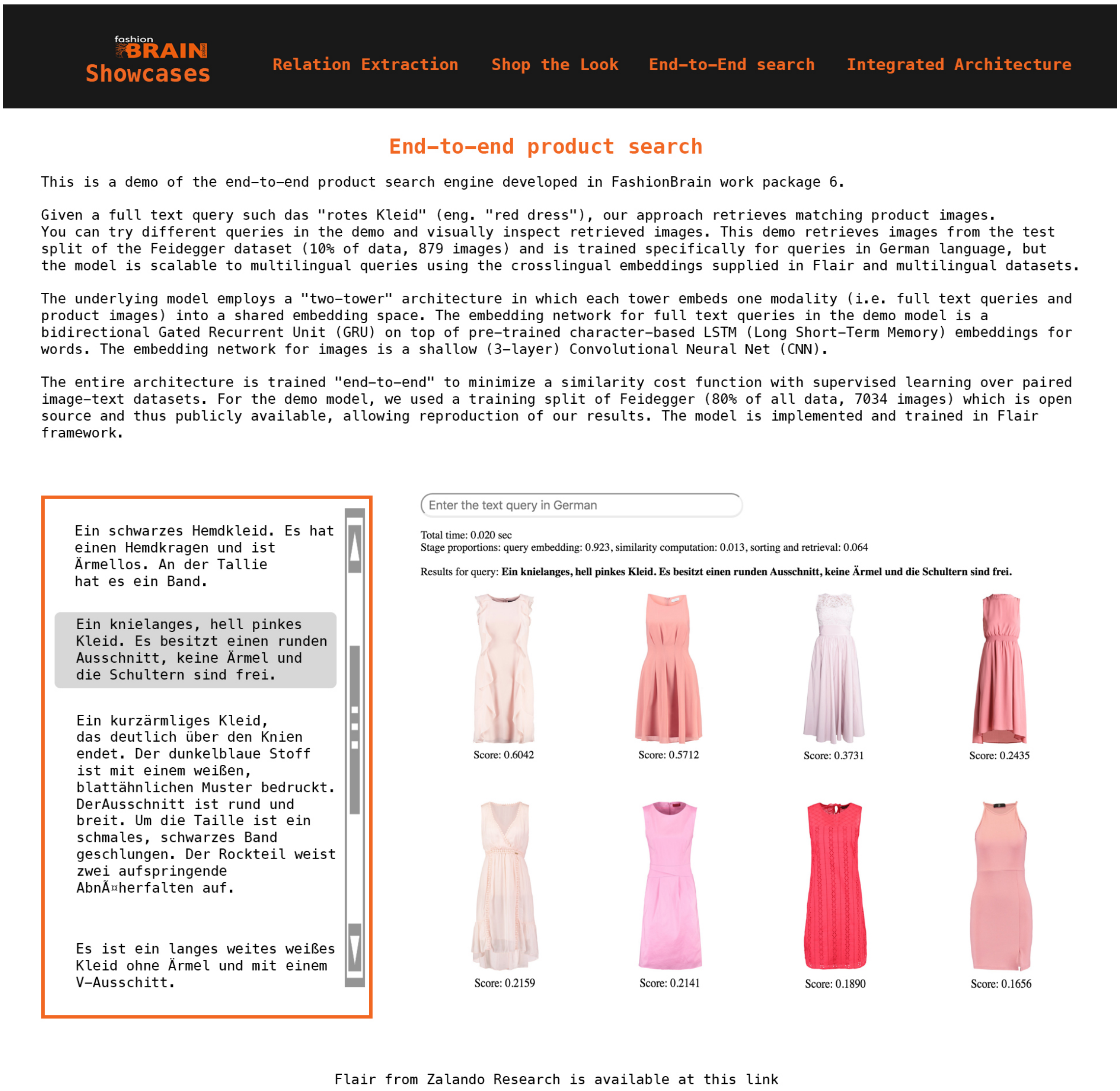


Figure 2.4: End-to-End Search.

2.6 FashionBrain Integrated Architecture

In this showcase, we report the design and implementation of the MonetDB-based FashionBrain integrated architecture for storing, managing and processing heterogeneous fashion data (i.e. structured relational and unstructured text data). The FashionBrain project targets at consolidating and extending existing European technologies in the area of database management, data mining, machine learning, image processing, information retrieval, and crowd sourcing to strengthen the positions of European (fashion) retailers among their world-wide competitors. For fashion retailers, the ability to efficiently extract, store, manage and analyse information from heterogeneous data sources, including structured data (e.g. product catalogues and sales information), unstructured data (e.g. Twitter, blogs and customer reviews), and binary (multimedia) data (e.g. YouTube and Instagram), is business critical. Therefore, one of the main objectives is to " Design and deploy novel big data infrastructure to support scalable multi-source data federation, and implement efficient analysis primitives at the core of the data management solution".

This showcase presents three main solutions: **FaBIAM Overview**, **MonetDB Window Function Extensions**, and **Trend Detection in Twitter text**.

A screenshot of the showcase is shown in Figure 2.5.

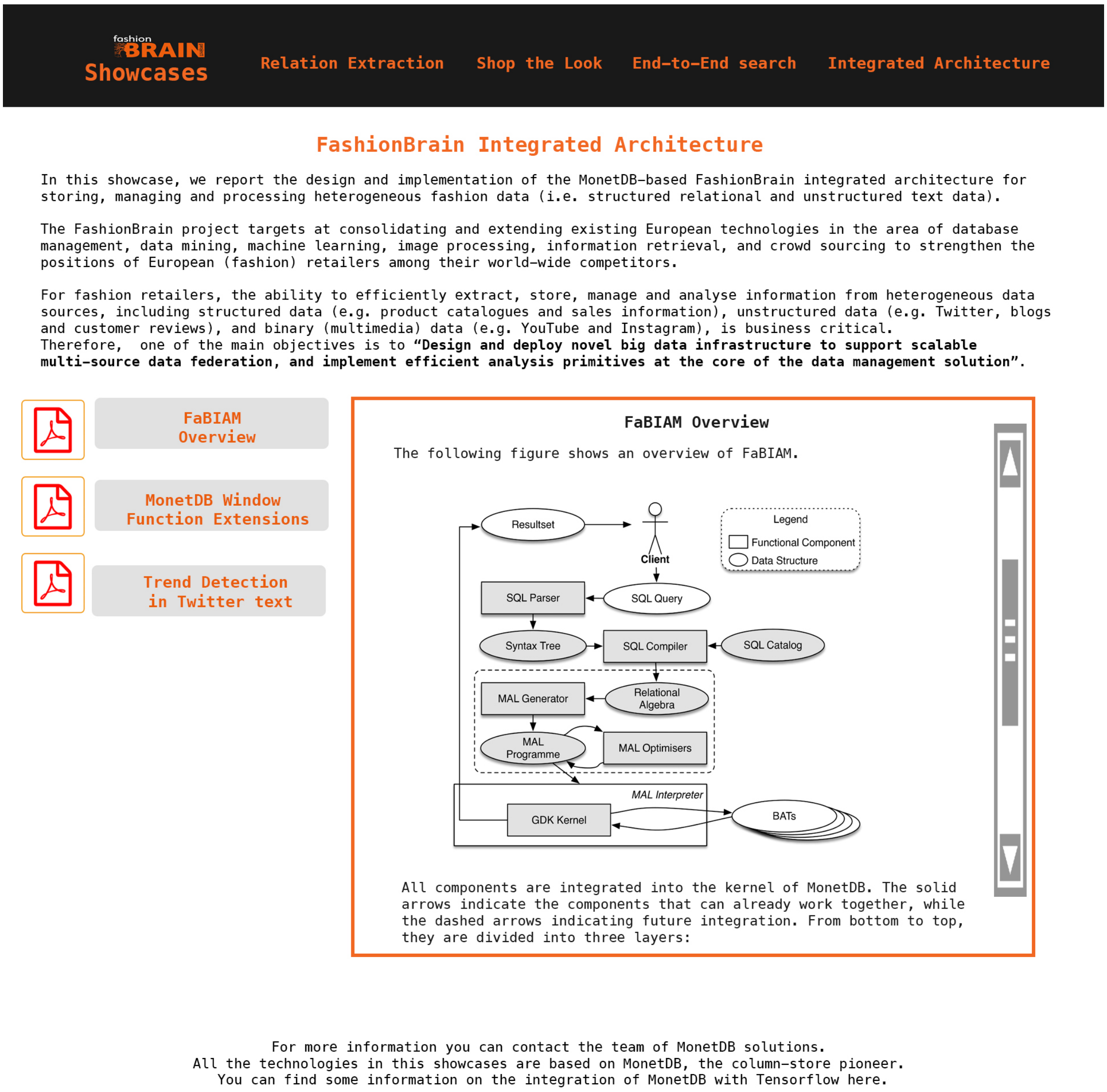


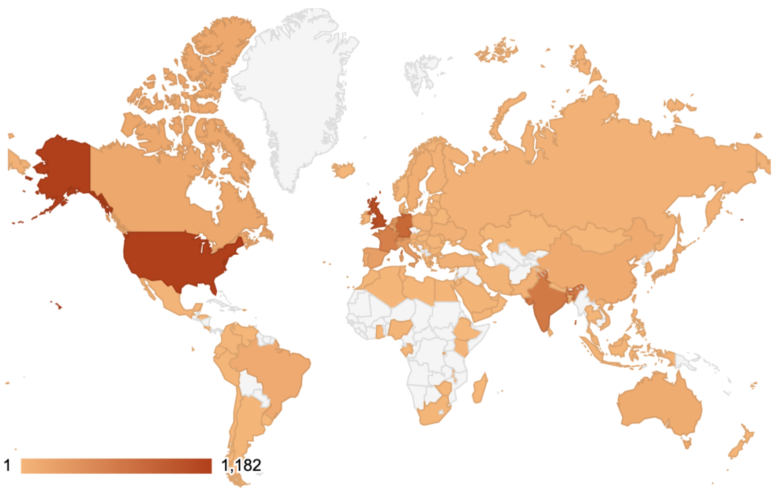
Figure 2.5: FashionBrain Integrated Architecture.

3 Communication and Dissemination

Website Traffic and Statistics

Analyzing the project's website traffic and statistics is essential in measuring and managing its efficiency in disseminating project information.

Website metrics (at time of publication) indicate that the project is targeting and reaching a wide audience across the globe (Figure 4.1)[1].



Countries	Sessions
United States	1,182
United Kingdom	1,017
Germany	768
India	591
France	518
Switzerland	418
Netherlands	301
Italy	248
Spain	224
Belgium	197

Figure 3.1: FashionBrain website visits by country (with top 10 in tabular form).

Figure 4.2 shows the growth the website has achieved since its inception in February 2017 to November 2019.

[1] data collected from February 2017 to October 2019

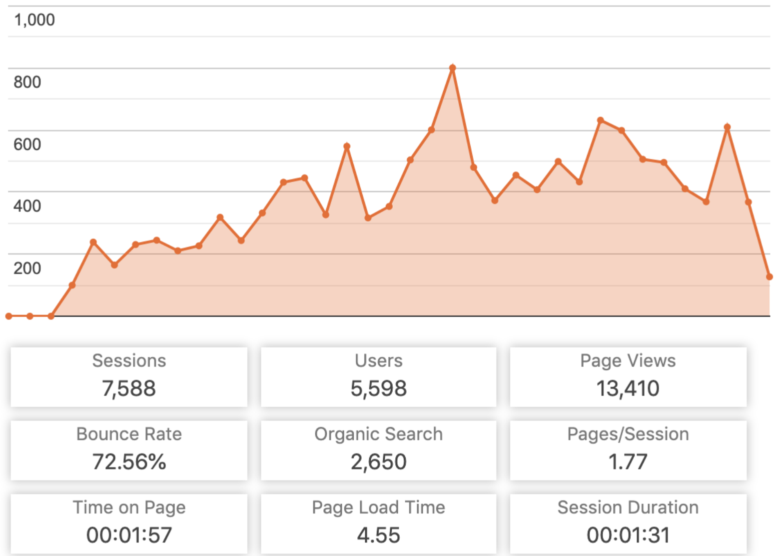


Figure 3.2: Webpage views and traffic statistics (Oct 2019).

Twitter

The FashionBrain project also has a visual presence on social media. The Twitter account, FashionBrain Project and associated handle, `@FashionBrain1`, is used to engage and communicate with already existing collaborators and interested parties (57 followers) as well as to attract potential new ones.

Figure 4.3 shows an overview of the project's Twitter visitors provenance, traffic medium and Social Network traffic: Twitter is confirmed to be the main source of traffic.

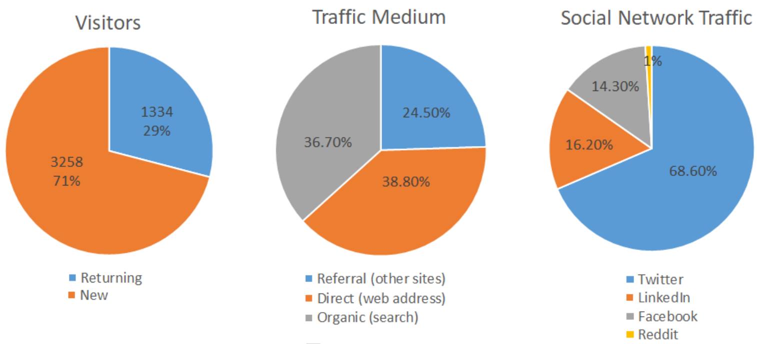


Figure 3.3: Visitors provenance, traffic medium and social network traffic.

Project partners can also update their own personal or institutional twitter accounts with relevant project communications and by simply mentioning the FashionBrain handle, the information is immediately shared with all of our followers without delay. The associated tweet can then be retweeted directly from the project's account (at a later time) without risking relevant news not being shared when opportunities for public engagement are at their highest, no matter the time or day (e.g. immediately following a high profile event).

4 Dissemination Report

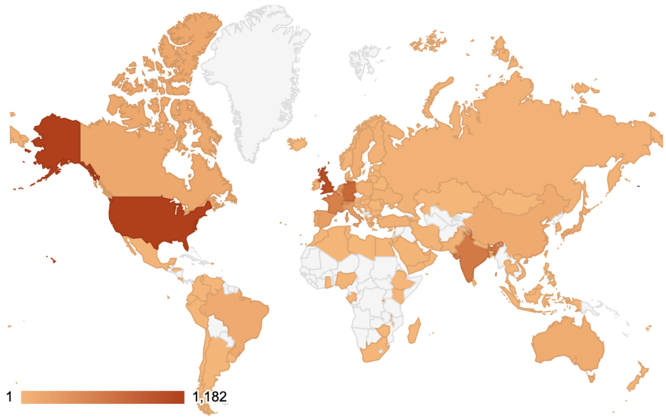
4.1 Dissemination Strategy, Channels, Promotion

We refer to D7.4 for details on the dissemination strategy, channels and promotional tools. We will update here the figures and details on the achievements and outputs.

4.2 Website Traffic and Statistics

Analyzing the project's website traffic and statistics is essential in measuring and managing its efficiency in disseminating project information.

Website metrics (at time of publication) indicate that the project is targeting and reaching a wide audience across the globe (Figure 4.1)[1].



Countries	Sessions
United States	1,182
United Kingdom	1,017
Germany	768
India	591
France	518
Switzerland	418
Netherlands	301
Italy	248
Spain	224
Belgium	197

Figure 4.1: FashionBrain website visits by country (with top 10 in tabular form).

Figure 4.2 shows the growth the website has achieved since its inception in February 2017 to November 2019.

[1]data collected from February 2017 to October 2019

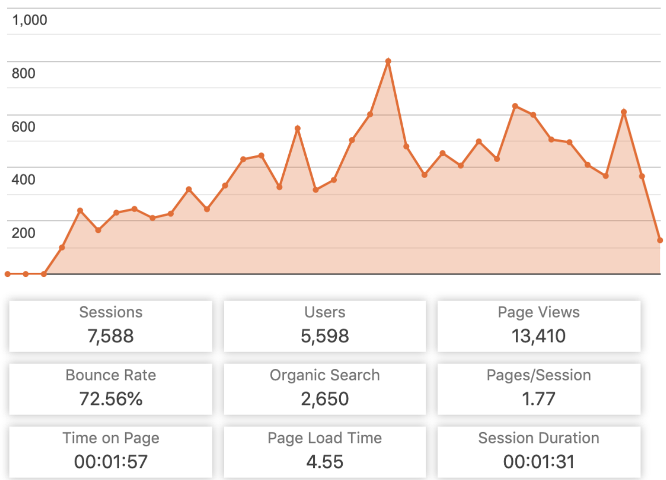


Figure 4.2: Webpage views and traffic statistics (Oct 2019).

Twitter

The FashionBrain project also has a visual presence on social media. The Twitter account, FashionBrain Project and associated handle, `@FashionBrain1`, is used to engage and communicate with already existing collaborators and interested parties (57 followers) as well as to attract potential new ones.

Figure 4.3 shows an overview of the project's Twitter visitors provenance, traffic medium and Social Network traffic: Twitter is confirmed to be the main source of traffic.

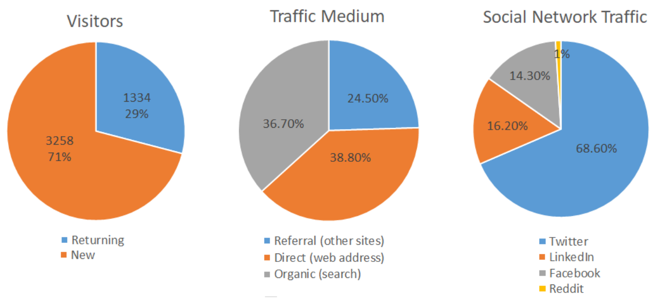


Figure 4.3: Visitors provenance, traffic medium and social network traffic.

Project partners can also update their own personal or institutional twitter accounts with relevant project communications and by simply mentioning the FashionBrain handle, the information is immediately shared with all of our followers without delay. The associated tweet can then be retweeted directly from the project's account (at a later time) without risking relevant news not being shared when opportunities for public engagement are at their highest, no matter the time or day (e.g. immediately following a high profile event).

4.3 KPIs

Key Performance Indicators (KPIs) are directly related to the performance of BDV SRIA activities that will deliver solutions, architectures, technologies and standards for the data value chain over the next decade.

The following KPIs for FashionBrain are an update on the ones reported in the BDV PPP Survey 2018 (March).

Job creation, job profiles and skills development (Contributes to KPI I-2, II-5 and II-8[2])

1. Does your project contribute to job creation? (Job creation forecasted within a reasonable timeframe in the future, 3 years after H2020 ends) **Yes.**
2. If yes, can you provide an estimation of the impact of your project in job creation forecasted within a reasonable timeframe in the future (or at least until 2022)? **Our solutions will increase the revenue of the European stakeholders involved in the project, which in turn will create more jobs. Moreover, our solution require the use of European crowdsourcing.**
3. Does your project contribute to increase the number of data workers in Europe? **Yes.**
4. If yes, please provide some qualitative statements about how your project is contributing to increase the number of data workers in Europe? **The number of employees of our partners increased (MDBS size doubled) since the inception of FashionBrain.**
5. Does your project contribute to the creation of new job profiles? **No.**
6. Does your project contribute to the creation/development of new skills? **Yes.**
7. If yes, explain how your project has contributed to the development / creation of new skills, and list the new skills developed. **Expertise of data analytics in the fashion industry.**
8. Number of Master students involved in your project. **2.**
9. Number of PhD students involved in your project. **3.**
10. Number of training activities and programmes (such as tutorials, webinars, etc.) organized by your project in 2018 and number of people benefited by those activities
    - **1 First Symposium on Biases in Human Computation and Crowdsourcing - 30 people.**
    - **1 Tutorial on crowdsourcing - 30 people.**

[2] BDV PPP KPI catalog

- **1 Social computing module - 30 people.**

Innovations and technical results (Contributes to KPI I-4, II-4, II-7, II-10, II-13, II-15, II-16, II-17)

1. Please list and describe all the Innovations with marketable or exploitable value developed by your project during 2018
   - We refer to D7.7 for details on this point.
2. Number of systems and technologies developed in the relevant sector in the project. **3.**
3. List the sectors and major domains supported by Big Data technology and applications developed in your project. **Commerce, Fashion**
4. Number of contributions to the technical priorities of the BDV SRIA (beyond state of the art)?
   - Contributions to Data Analytics **2.**
   - Contributions to Data Visualisation and User Interaction / Experience **1.**
   - Contributions to Engineering and DevOps for Big Data **1**
5. Are you assessing quality, diversity and value of data assets? **Yes.**
6. If yes, what metrics are you using to quantify them? **Evaluation and crowdsourcing to detect entity linking error rate, evaluation of inter-rater agreement.**

Experiments in Big Data (Contributes to KPI II-11, II-12 and II-14)

1. Please describe your criteria for an experiment to qualify as "large-scale" **Investment and number of users**
2. Number of data experiments/use cases of any kind or size conducted **8**
3. Please provide some qualitative data to support your answer **Various experiments for quality control and labeling in crowdsourced data**

Contribution to Macro-Economics KPIs (Contributes to KPI II-1, II-2 and II-3)

1. Does your project contribute (has contributed or plan to contribute) to increase revenue share of EU companies against total of revenue of EU, US, Japan, Brazil? **Yes.**
2. If yes, please provide some qualitative statements about how your project is contributing to this. **Increase the market position of European fashion retailers.**
3. Does your project contribute (has contributed or plan to contribute) to increase the number of European Companies offering data technology, applications? **Yes.**

4. If yes, please provide some qualitative statements about how your project is contributing to this. **A data analytics startup focusing on fashion will open at the end of the project.**

Mobilisation of stakeholders, outreach, success stories (Contributes to section 2.2 of the monitoring report of the contractual PPPs)

1. Number of dissemination events, seminars, conferences organised by your project. **11**
2. List and describe all the activities performed to mobilise and outreach to stakeholders in your project
   - **We refer to the next sections for more detail.**
3. List your project main stakeholders (the ones you are addressing and need to outreach) and briefly indicate how did you address outreach to them. **Fashion experts and fashion companies interested in big data analytics and trend detection.** We refer to D7.7 for more details.

4.4 Project Achievements

We now report the project achievements, first with a summary of the different type of outputs, and then with a more detailed report, containing the individual outputs.

4.4.1 Summary

We refer to the updated PROJECT ACHIEVEMENTS page (https://fashionbrain-project.eu/project-achievements/).

A summary of external dissemination is provided in Figures 4.4-4.7.

Figure 4.4 summarises the peer-reviewed and scholarly periodicals output, aimed at specialist experts and researchers. Such publications contain original research and developments or conclusions based on collected data, which are written using technical language.

Press releases (e.g. printed, online, video or radio) which are sent to the media by FashionBrain team members, serve to inform a wider-range of collaborators, as well as the general public, about newsworthy and exciting project achievements, and are summarised in Figure 4.8.

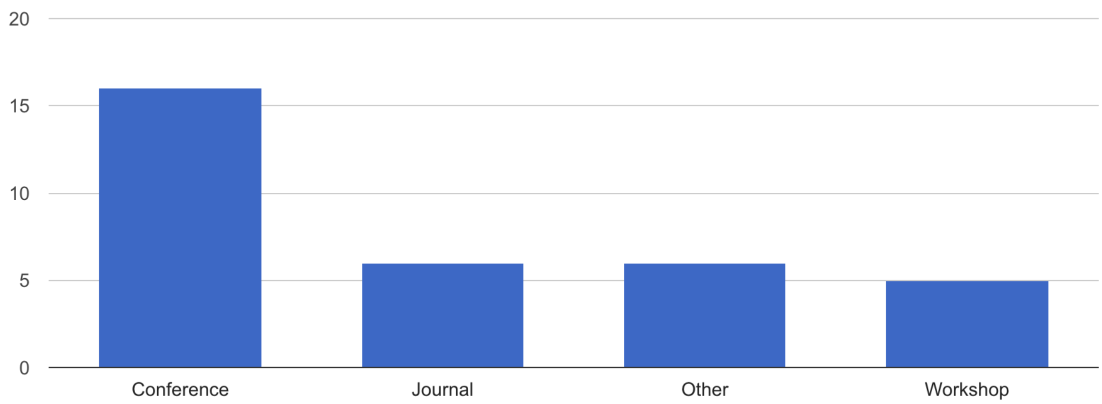


Figure 4.4: Published articles summary.

Attending conferences is integral to the success of the FashionBrain project as it provides an opportunity to gain new knowledge in the field and present the project's technical research. Also, by meeting with like-minded peers, in both academia and industry, the project is exposed to invaluable referrals, products and best-practices (Figure 4.6).

It has also proven useful to attend other events (outside of academic and industrial conferences) that bring together people who share a common interest in the field, as it increases the project's professional network and resources (Figure 4.7).

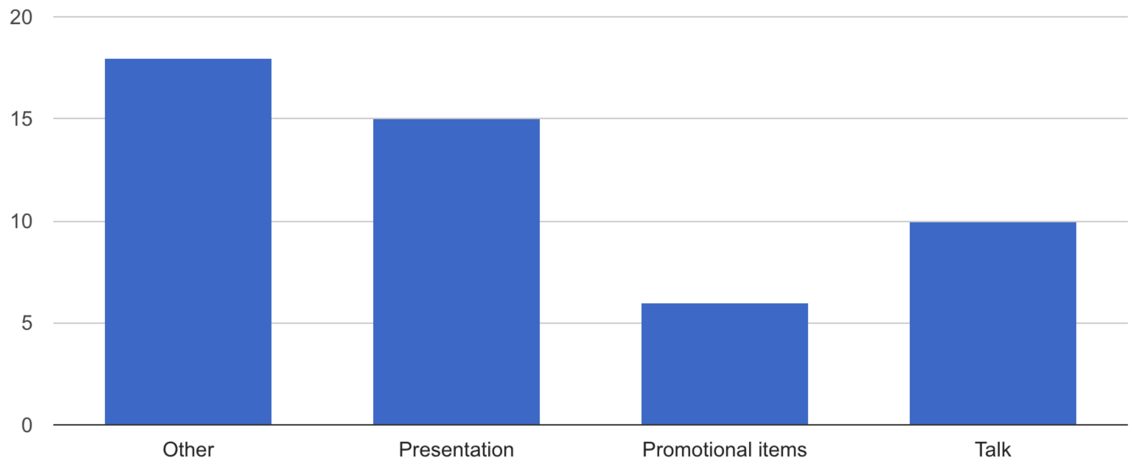


Figure 4.5: Industrial events summary.

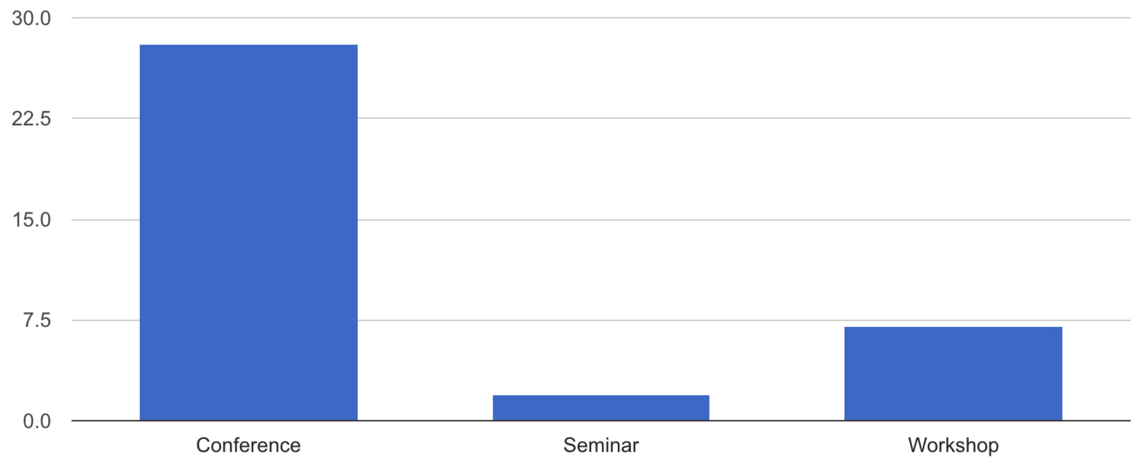


Figure 4.6: Academic events participation summary.

4.4.2 Beuth Notable Exploitations

- In 2017 the startup Qualifiction.com (aka Qualifiction GmbH) has been established as spin off from the Beuth group. The startup uses deep learning and text mining technologies from the group (and FashionBrain) to predict next best sellers. So far, the startup has been acquired more than 250K funding and has been received several awards at the Leipzig book fair for digital products. Too, the startup has been published in major newspapers, including die Zeit, Süddeutsche Zeitung, Handelsblatt as well as in major TV channels, such as WDR. See more at https://www.qualifiction.info/aktuelles/.
- The in-database entity linkage of Beuth and MonetDB received a best paper award in 2019 at IEEE BigComp. Torsten Kilias, the main author has been joining in 2019 the German database manufacturer Exasol SE (110 persons)

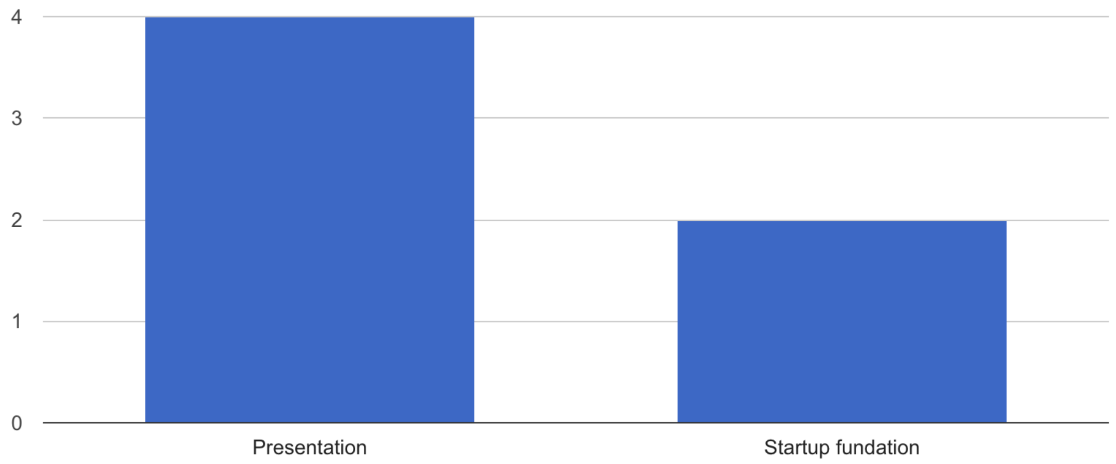


Figure 4.7: Other dissemination summary.

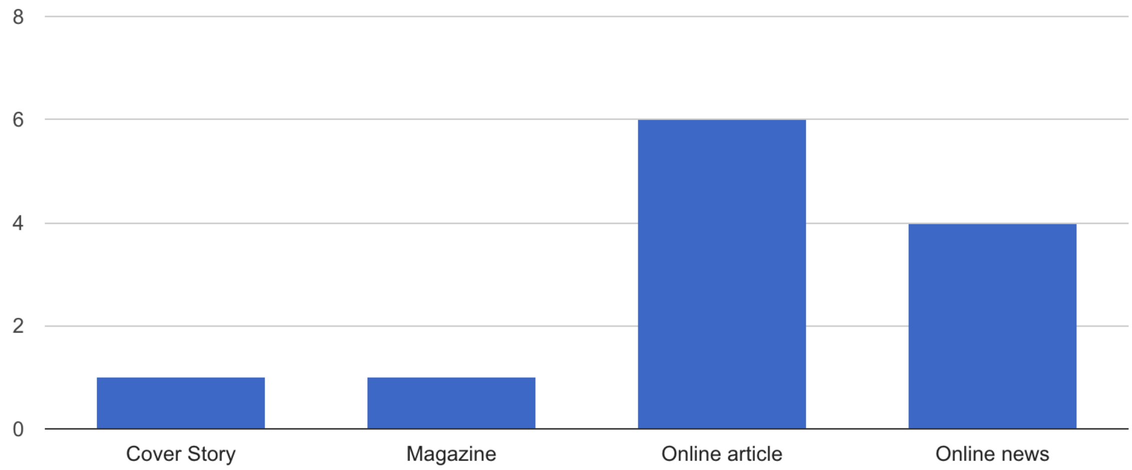


Figure 4.8: Press releases section.

and is currently implementing the principles of handling neural networks in a main memory database. So far, this is the very first commercial database that implements these principles. FashionBrain was here a pioneer in research but also transferring these results to a commercial product.

- Further outstanding results had been in research dissemination, including presentation at top-level academic venues, including ACL 2019 or CIKM 2019. More recently, Alexander Löser gave talks at Peking University or Lenovo AI labs to disseminate results and to establish relations also to China. Peking University is the most prestigious university in China. Most relevant, Beuth could unravel, how neural networks learn. This helps to explain results for tasks such as entity linkage and event extraction (D4.3 and D4.4). A Demonstrator for explaining how BERT and other transformer networks extract relations is available at https://prof.beuth-hochschule.de/loeser/what-we-do/explaining-and-benchmarking/explaining-transformer-networks/?L=220 Beuth was again pioneering this research direction of explaining very deep neural networks and for tasks crucial in fashion brain, such as entity linkage or relation extraction.

- Based on the profound knowledge of industry gained from Fashionbrain and the outstanding research success of Beuth, the Beuth team could acquire ¿ 2.2 M€ from federal funds for next research projects. That includes PLASS a project with Siemens and others to extract textual data for improving supply chain management, project Smart-MD for extracting textual data from doctor letters with the largest clinics in Germany (Charité and Helios Kliniken) as well as the health startup number one in Europe Ada health. Both projects have been funded with ¿ 650K starting mid 2019. The last project Service Meister has been funded with 1 Mio Euro and utilizes neural text mining technologies invented in fashionbrain for extracting helpful information from service tickets in engineering. Project partners in this AI flagship project of the German national AI program include large scale machine engineering companies such as Trumpf, Wirth or Heidelberger Druck. The project starts 2020.

4.4.3 UNIFR Notable Exploitations

Teaching: UNIFR introduced fashion related topics to their teaching activities. For example, the main topic of 2019 year's data science seminar given to MSc students is neutral networks and their applications to fashion data (https://diuf.unifr.ch/main/xi/ds-xi/index.html). The attendance to the seminar has increased by 20% as compared to last year.

Student Supervision: UNIFR has supervised four MSc/BSc students working on fashion related topics since the beginning of the project. The list of topics is as follows: (i) Ahana Mallik: Trend Prediction using Fashion Datasets, 2018; (ii) Leutrim Kaleci: Modeling the Evolution of Fashion Trends using Matrix Factorization Techniques, 2018; (iii) Mili Biswas: Incremental Update of FBTaxonomy, 2019; (iv) Daniel Sanz: Experimental Evaluation of Time Series Prediction Techniques, 2019.

Presentation: The FashionBrain project has been selected to be presented during the university open day event https://events.unifr.ch/explora/fr/fashionbrain.html.

USFD Notable Exploitation

The main exploitations for USFD have been the organisation of BHCC (as shown in detail in Section 4), the teaching of Social Computing with topics from FashionBrain, the realisation of dissertation topics on fashion taxonomy, and the production of over

20 peer-reviewed publications.
We refer to the next sections for a more detailed report on these outputs.

4.4.4 FashionBrain Glossary

The FashionBrain glossary has been updated and is available at https://fashionbrain-project.eu/wp-content/uploads/2019/12/FashionBrain_Glossary.pdf.

4.4.5 Academic Publications

We report in Table 4.1 the academic publications produced during the duration of the project.

Title	Authors	Venue	Type
Let's Agree to Disagree: Fixing Agreement Measures for Crowdsourcing	Alessandro Checco, Kevin Roitero, Eddy Maddalena, Stefano Mizzaro and Gianluca Demartini	HCOMP 2017	Conference
Understanding Engagement through Searching Behaviour	Mengdie Zhuang, Gianluca Demartini and Elaine Toms	CIKM 2017	Conference
Considering Assessor Agreement in IR Evaluation	Eddy Maddalena, Kevin Roitero, Gianluca Demartini and Stefano Mizzaro	ICTIR 2017	Conference
FashionBrain Project: A Vision for Understanding Europe's Fashion Data Universe	Alessandro Checco, Gianluca Demartini, Alexander Löser, Ines Arous, Matthias Dantone, Richard Koopmanschap, Svetlin Stalinov, Martin Kersten, Ying Zhang	KDD Fashion 2017	Workshop
The Projector: An Interactive Annotation Projection Visualization Tool	Alan Akbik and Roland Vollgraf	EMNLP 2017	Conference
ZAP: An Open-Source Multilingual Annotation Projection Framework	Alan Akbik and Roland Vollgraf	LREC 2018	Conference
FEIDEGGER: A Multi-modal Corpus of Fashion Images and Descriptions in German	Leonidas Lefakis, Alan Akbik and Roland Vollgraf	LREC 2018	Conference
Love at First Sight: MonetDB/TensorFlow	Richard Koopmanschap, Ying Zhang and Martin Kersten	ICDE 2018	Other
Love at First Sight: MonetDB/TensorFlow	Richard Koopmanschap, Ying Zhang and Martin Kersten	XLDB2018	Other
In-Database Machine Learning with MonetDB/TensorFlow	Richard Koopmanschap, Ying Zhang, Martin Kersten	XLDB2018	Other
On Fine-Grained Relevance Scales	Kevin Roitero, Eddy Maddalena, Gianluca Demartini, and Stefano Mizzaro	SIGIR2018	Other
Investigating User Perception of Gender Bias in Image Search: The Role of Sexism	Jahna Otterbacher, Alessandro Checco, Gianluca Demartini, and Paul Clough	SIGIR2018	Conference
On the Volatility of Commercial Search Engines and its Impact on Information Retrieval Research	Jimmy, Guido Zuccon, and Gianluca Demartini	SIGIR2018	Other
The Evolution of Power and Standard Wikidata Editors: Comparing Editing Behavior over Time to Predict Lifespan and Volume of Edits	Cristina Sarasua, Alessandro Checco, Gianluca Demartini, Djellel Difallah, Michael Feldman, and Lydia Pintscher	CSCW	Journal
An Introduction to Hybrid Human-Machine Information Systems	Gianluca Demartini, Djellel Eddine Difallah, Ujwal Gadiraju, and Michele Catasta	Foundation and Trends in Web Science	Other
All That Glitters is Gold - An Attack Scheme on Gold Questions in Crowdsourcing	Alessandro Checco, Jo Bates, and Gianluca Demartini	HCOMP 2018	Conference
Investigating Stability and Reliability of Crowdsourcing Output	Rehab K. Qarout, Alessandro Checco, Kalina Bontcheva	CrowdBias 2018	Workshop

RelVis: Benchmarking OpenIE Systems	Rudolf Schneider, Tom Oberhauser, Tobias Klatt, Felix A. Gers, Alexander Löser	ISWC 2017	Conference
Analysing Errors of Open Information Extraction Systems	Rudolf Schneider, Tom Oberhauser, Tobias Klatt, Felix A. Gers, Alexander Löser	EMNLP 2017 Workshop	Workshop
Contextual String Embeddings for Sequence Labeling	Alan Akbik, Duncan Blythe and Roland Vollgrad	COLING 2018	Conference
All Those Wasted Hours: On Task Abandonment in Crowdsourcing	Lei Han, Kevin Roitero, Ujwal Gadiraju, Cristina Sarasua, Alessandro Checco, Eddy Maddalena and Gianluca Demartini	WSDM 2019	Conference
IDEL: In-Database Neural Entity Linking	Torsten Kilias, Alexander Löser, Felix A. Gers, Richard Koopmanschap, Ying Zhang and Martin Kersten	IEEE BigComp2019	Conference
RecovDB: accurate and efficient missing values recovery for large time series	Ines Arous, Mourad Khayati, Philippe Cudré-Mauroux, Ying Zhang, Martin Kersten and Svetlin Stalinlov	ICDE 2019	Conference
Deadline-Aware Fair Scheduling for Multi-Tenant Crowd-Powered Systems	Djellel Difallah, Alessandro Checco, Gianluca Demartini and Philippe Cudré-Mauroux	TSC	Journal
Implicit Bias in Crowdsourced Knowledge Graphs	Gianluca Demartini	HumBL-WWW2019	Workshop
The Impact of Task Abandonment in Crowdsourcing	Lei Han, Kevin Roitero, Ujwal Gadiraju, Cristina Sarasua, Alessandro Checco, Eddy Maddalena and Gianluca Demartini	TKDE	Journal
Platform-related Factors in Repeatability and Reproducibility of Crowdsourcing Tasks	Rehab Qarout, Alessandro Checco, Gianluca Demartini and Kalina Bontcheva	HCOMP 2019	Conference
Scalable recovery of missing blocks in time series with high and low cross-correlations	Mourad Khayati, Philippe Cudré-Mauroux and Michael H. Böhlen	KAIS 2019	Journal
Mind the Gap: An Experimental Evaluation of Imputation of Missing Values Techniques in Time Series	Mourad Khayati, Alberto Lerner, Zakhar Tymchenko and Philippe Cudré-Mauroux	VLDB 2020	Conference
SECTOR: A Neural Model for Coherent Topic Segmentation and Classification	Sebastian Arnold, Rudolf Schneider, Philippe Cudré-Mauroux, Felix A. Gers, Alexander Löser	TACL 2019	Journal
FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP	Alan Akbik, Tanja Bergmann, Duncan Blythe, Kashif Rasul, Stefan Schweter and Roland Vollgraf	NAACL-HLT 2019	Conference
Multilingual Sequence Labeling With One Model	Alan Akbik, Tanja Bergmann and Roland Vollgraf	NLDL 2019	Workshop
Adversarial Attacks on Crowdsourcing Quality Control	Alessandro Checco, Jo Bates, Gianluca Demartini	JAIR	Journal

Table 4.1: FashionBrain publications.

4.4.6 Academic Events Participation

We report in Table 4.2 the academic events the member of the consortium attended under the FashionBrain sponsorship.

Event	Type	Date	Attending representative
Digital Transformation & Global Society (DTGS 2018)	Conference	Jun-17	Gianluca Demartini
Machine learning meets fashion' workshop at KDD 2017	Workshop	Aug-17	Alessandro Checco
2017 Workshop on Hybrid Human-Machine Computing (HHMC 2017). Guildford, UK	Workshop	Sep-17	Alessandro Checco
The 28th edition of the Australasian Database Conference, ADC 2017	Conference	Sep-17	Gianluca Demartini
2017 Conference on Empirical Methods on Natural Language Processing (EMNLP 2017)	Conference	Sep-17	Alan Akbik, Duncan Blythe
The fifth AAAI Conference on Human Computation and Crowdsourcing	Conference	Oct-17	Alessandro Checco
ReWork Machine Learning Summit 2017	Seminar	Oct-17	Roland Vollgraf
Australasian Document Computing Symposium	Conference	Dec-17	Gianluca Demartini
Thirty-first Conference on Neural Information Processing Systems (NIPS 2017)	Conference	Dec-17	Roland Vollgraf
The 3rd Strategic Workshop on Information Retrieval in Lorne (SWIRL)	Workshop	Feb-18	Gianluca Demartini
iConference 2018	Conference	Mar-18	Alessandro Checco
34th IEEE International Conference on Data Engineering	Conference	Apr-18	Ying Zhang, Martin Kersten
11TH Extremely Large Databases Conference	Conference	May-18	Ying Zhang, Sjoerd Mullender
11th Edition of the Language Resources and Evaluation Conference (LREC 2018)	Conference	May-18	Alan Akbik
ACM SIGMOD/PODS International Conference on Management of Data	Conference	Jun-18	Martin Kersten
the 7th International workshop on Testing Database Systems (DBTest)	Workshop	Jun-18	Martin Kersten
The sixth AAAI Conference on Human Computation and Crowdsourcing	Conference	Jul-18	Alessandro Checco, Gianluca Demartini
The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval	Conference	Jul-18	Gianluca Demartini
LIBER Annual Conference (LIBER 2018)	Conference	Jul-18	Alan Akbik
CrowdBias 2018	Workshop	Jul-18	Alessandro Checo, Gianluca Demartini
The 27th International Conference on Computational Linguistics (COLING 2018)	Conference	Aug-18	Alan Akbik
International Conference on Computer Vision (ICCV 2017)	Conference	Oct-18	Marko Jocic, Matthias Dantone
The 9th biennial Conference on Innovative Data Systems Research (CIDR)	Conference	Jan-19	Ying Zhang, Martin Kersten
Second Workshop on Software Foundations for Data Interoperability	Workshop	Feb-19	Alexander Löser (Chair)
Northern Lights Deep Learning Workshop, NLDL 2019	Workshop	Feb-19	Alan Akbik

IEEE BigComp2019	Conference	Mar-19	Alexander Löser (Tutorial Chair)
35th IEEE International Conference on Data Engineering (ICDE 2019)	Conference	Apr-19	Svetlin Stalinov
35th IEEE International Conference on Data Engineering (ICDE 2019)	Conference	Apr-19	Svetlin Stalinov (demo)
Dagstuhl Seminar on Multi-Document Information Consolidation	Seminar	Apr-19	Sebastian Arnold (speaker)
the 2019 ACM SIGMOD/PODS Conference	Conference	Jun-19	Ying Zhang, Martin Kersten
Annual Conference of the North American Chapter of the Association for Computational Linguistics, NAACL 2019	Conference	Jun-19	Alan Akbik
ACM SIGMOD/PODS International Conference on Management of Data (SIGMOD/PODS 2019)	Conference	Jul-19	Ying Zhang (Industrial track chair), Martin Kersten (General chair)
ACL 2019	Conference	Jul-19	Sebastian Arnold (speaker)
45th International Conference on Very Large Data Bases	Conference	Aug-19	Ying Zhang, Martin Kersten
Data Power 2019	Conference	Sep-19	Alessandro Checco (speaker)
HCOMP 2019	Conference	Oct-19	Alessandro Checco (speaker)
First symposium on Biases in Human Computation and Crowdsourcing	Workshop	Oct-19	Alessandro Checco (speaker)
ACM CIKM 2019	Conference	Nov-19	Benjamin Winter (Speaker), Alexander Löser (Speaker)

Table 4.2: Academic events participation.

4.4.7 Press Releases

We report in Table 4.3 the press releases produced during the project.

Name	Partner	Type	Link
datanami	MDBS	Online news	https://www.datanami.com/this-just-in/monetdb-solutions-appoints-niels-nes-as-cto/
EEnterpriseAI news	MDBS	Online news	https://www.enterpriseai.news/2019/10/07/monetdb-solutions-secures-an-investment-from-servicenow-to-help-large-enterprises-drive-digital-transformation-at-scale/
HiPEAC news	MDBS	Online news	https://www.hipeac.net/press/6829/ten-winners-selected-for-the-2017-hipeac-tech-transfer-awards/
HiPEAC info 51	MDBS	Magazine	https://www.hipeac.net/assets/public/publications/newsletter/hipeacinfo51_final_corrected.pdf
Computer Weekly	MDBS	Online article	http://www.computerweekly.com/news/450422330/Dutch-database-design-drives-practical-innovation
Handelsblatt	Beuth-HS	Online article	http://veranstaltungen.handelsblatt.com/kuenstliche-intelligenz/2018/03/03/ki-als-enabler/
Beuth-Magazin	Beuth-HS	Cover Story	http://www.beuth-hochschule.de/fileadmin/oe/pressestelle/beuth-magazin/2017-1_beuth-magazin.pdf
The University of Sheffield	USFD	Online article	https://www.sheffield.ac.uk/faculty/social-sciences/news/fashion-algorithm-future-trends-project-1.671380
The University of Sheffield	USFD	Online article	https://www.sheffield.ac.uk/is/research/projects/fashionbrain
Tagesspiegel	Beuth-HS	Online news	https://science-match.tagesspiegel.de/digital-future-2018/speakers/alexander-loser
Exasol Magazine	Beuth-HS	Online article	https://www.exasol.com/en/blog/interactive-text-mining-exasol-indrex-mm/
KI-Berlin	Beuth-HS	Online article	https://ki-berlin.de/en/blog/article/prof-dr-alexander-loeser-beuth-university-of-applied-sciences/

Table 4.3: Press releases.

4.4.8 Prototypes and Technologies

We summarise in Table 4.4 the software produced. Regarding the demonstrators, we refer to D4.2, D4.4, D5.1, D5.2, D5.3, D5.5, D6.3, D6.5 for more details.

Name	Partners	Type	Link	Description
MonetDB with extended windowing functions	MDBS	software	https://www.monetdb.org/Downloads	MonetDB Apr2019 feature release including the extended SQL windowing functions
In-Database Machine Learning	MDBS	software	https://github.com/MonetDB	MonetDB-Tensorflow integration through SQL Python UDFs which allows executing machine learning tasks inside the kernel of the MonetDB RDBMS
MonetDB continuous query extension	MDBS	software	https://dev.monetdb.org/hg/MonetDB/shortlog/trails	MonetDB extended with a continuous query processing engine for IoT/Streaming data
MonetDB JSON (renewed)	MDBS	software	https://dev.monetdb.org/hg/MonetDB/shortlog/json	Renewed support for JSON data loading and processing in MonetDB

In-Database Entity Linking	Beuth, MDBS	software	not public	Integration of Beuth Entity Linking technology with MonetDB for advanced text analysis
RecovDB	UNIFR, MDBS	software	http://revival.exascale.info	Integration of UNIFR's CD-based technology with MonetDB for missing value recovery in time series
Agreement Phi	USFD	software	http://agreement-measure.sheffield.ac.uk	Source code and live demo of a novel agreement measure for crowdsourcing
Crowdsourcing logging interface	USFD	API	https://github.com/AlessandroChecco/herokulogging	Append-only, ephemeral in-memory logging REST interface. https://fast-logging.herokuapp.com
Gender bias dataset	USFD	dataset	https://github.com/AlessandroChecco/gender_bias	Dataset used in "Investigating User Bias in Image Search: A Cross-Regional Study". It contains 2,811 query-description comparisons for 281 different users.
Tasty Entity Linkage	Beuth	API	http://demo.datexis.com/tasty	Entity Linkage against Wikipedia
Scalable Crowdsourced Social Media Annotation Demo	Fashwell	API	https://fashionbrain-project.eu/scalable-crowdsourced-social-media-annotation-demo	
Product Taxonomy Linking	Fashwell	API	https://fashionbrain-project.eu/product-taxonomy-linking	
Demo on Zalando deep learning powered search engines	Zalando	software	https://fashionbrain-project.eu/demo-on-zalando-deep-learning-powered-search-engines	
Tasty feat. IDEL Demonstration	Beuth	software	https://fashionbrain-project.eu/beuth-tasty-feat-idel-demonstration	
BERT Layerwise Analysis	Beuth	software	https://demo.datexis.com/visbert	We visualize the most important representation BERT for text mining in FashionBrain
Flair release 0.4.4	Zalando	software	https://github.com/zalandoresearch/flair/releases/tag/v0.4.4	Release 0.4.4 of the popular Flair library

Table 4.4: Software output.

4.4.9 Event Organisation

We report in Table 4.5 the event organised by the FashionBrain consortium during the duration of the project.

Event Name	Organisers	Venue	Type	Date	Link	Description
CrowdBias	USFD	Zurich	Workshop	Jul-18	https://sites.google.com/view/crowdbias	The goal of this workshop is to analyze both existing biases in crowdsourcing, and methods to manage bias via crowdsourcing. We will discuss different types of biases, measures and methods to track bias, as well as methodologies to prevent and solve bias.
IEEE BigComp2019	Beuth+Uni Kyoto+ Uni Osaka+ Uni Tokyo	Kyoto	Conference	Mar-19	http://www.bigcomputing.org	The goal of the International Conference on Big Data and Smart Computing (BigComp), initiated by KIISE (Korean Institute of Information Scientists and Engineers), is to provide an international forum for exchanging ideas and information on current studies, challenges, research results, system developments, and practical experiences in these emerging fields.
Second Workshop on Software Foundations for Data Interoperability	Beuth+Uni Kyoto	Kyoto	Workshop	Feb-19	http://www.prg.nii.ac.jp/projects/biscuits/second-workshop	This workshop aims at fostering discussion, exchange, and innovation in research and development in software foundations for data interoperability based on bidirectional transformation. Researchers and professionals from software, database, big data processing, and distributed and parallel processing are invited to share their knowledge and experience.
HumBL 2019: Augmenting Intelligence with Bias-aware Humans-in-the-Loop	USFD	S. Francisco	Workshop	May-19	https://humlworkshop.github.io/HumBL-WWW2019	The goal of this workshop is to bring together researchers and practitioners in various areas of AI (i.e., Machine Learning, NLP, Computational Advertising, etc.) to explore new pathways of the human-in-the-loop paradigm. We aim to analyze both existing biases in crowdsourcing, and explore various methods to manage bias via crowdsourcing. We would like to discuss different types of biases, measures and methods to track bias, as well as methodologies to prevent and mitigate different types of bias. We will provide a framework for discussion among scholars, practitioners and other interested parties, including crowd workers, requesters and crowdsourcing platform managers.

First symposium on Biases in Human Computation and Crowdsourcing	USFD+Qrowd	Sheffield	Symposium	Oct-19	https://sites.google.com/sheffield.ac.uk/bhcc2019	The goal of this symposium is to analyse both existing human biases in hybrid systems, and methods to manage bias via crowdsourcing and human computation. We will discuss different types of biases, measures and methods to track bias, as well as methodologies to prevent and solve bias.

Table 4.5: Events Organisation.

4.4.10 Education

In Table 4.6 we report the educational events carried out by the member of the Consortium during the duration of the project.

Event Name	Organisers	Venue	Type	Description
Online MOOC Machine Learning	Beuth	CEBIT	Acatech Online MOOC	Module LSTM & Text Mining
Master Data Science (Engl Language)	Beuth	Beuth	Master (4 Semesters)	22 places (2017: >180 applicants, 2018 >330 applicants)
Social Computing Master internship project	USFD MDBS	Zurich Amsterdam	Bachelor Master	Crowdsourcing design and implement trend-mining demo using MonetDB and FLAIR
Master graduation project	MDBS	Braga (PT) & Amsterdam	Master	continuous query processing, streaming time series data

Table 4.6: Educational events.

4.4.11 Industrial Events

In Table 4.7 we report the industrial events attended by the members of the consortium during the duration of the project.

Name	Venue	Date	Attendees	Type
Big Data, Amsterdam v 6.0	Funda, Amsterdam, NL	January 2017	Richard Koopmanschap	Other
Deep Learning for Text Mining Tasks , inovex GmbH	Hamburg, DE	February 2017	Alexander Löser	Presentation
Mooc Artificial Intelligence, acatech, CEBIT	Hannover, DE	March 2017	Alexander Löser	Presentation
ACM Distinguished Speaker talk	Accenture Latvia	April 2017	Gianluca Demartini	Talk
Amsterdam Artificial Intelligence & Deep Learning (H2O & Booking.com)	Booking.com, Amsterdam, NL	April 2017	Richard Koopmanschap	Other

Artificial Intelligence Day, Springer Nature	Berlin, DE	May 2017	Alexander Löser	Talk
Panel debate Let's talk about Data Products, inovex GmbH	Cologne, DE	May 2017	Alexander Löser	Talk
Panel Debate Data Products Whats Next, inovex GmbH	Hamburg, DE	May 2017	Alexander Löser	Presentation
"Deep Learning & AI" by Scyfer #1	Impact Hub Amsterdam, NL	May 2017	Richard Koopmanschap	Other
CWI in Bedrijf	CWI Amsterdam, NL	May 2017	Ying Zhang	Promotional items
amst-R-dam Simple Imputation and Date Padding	CWI Amsterdam, NL	May 2017	Richard Koopmanschap	Other
Data Products and Exchange	Hasso Plattner Insititute, Potsdam, DE	June 2017	Alexander Löser	Presentation
Text and data mining(TDM) workshop in European Parliament	Brussels, BE	June 2017	Alan Akbik	Presentation
PyData Amsterdam: Data Science week edition @ Flow Traders	Flow Traders, Amsterdam, NL	June 2017	Richard Koopmanschap	Other
Smart Cities 2.0 congres	Figi Zeist, NL	June 2017	Ying Zhang	Other
ADS Drinks & Pizza Summer Startup	UvA, Amsterdam, NL	June 2017	Richard Koopmanschap	Other
ADS Coffee & Data: Visual Analytics	UvA, Amsterdam, NL	July 2017	Richard Koopmanschap, Ying Zhang	Other
"So how does Tensorflow work?", guest star Siraj Raval	Google Netherlands, Amsterdam	August 2017	Richard Koopmanschap, Ying Zhang	Other
New challenges in Reinforcement Learning: Dr. O. Vinyals (Google DeepMind)	Amsterdam Science Park, NL	September 2017	Richard Koopmanschap	Other
Shoptalk Europe	Copenhagen, DK	October 2017	Matthias Dantone	Promotional items
European Big Data Value Forum 2017	Paris, FR	November 2017	Ying Zhang, Martin Kersten	Other
Data Datives 2017	Berlin, DE	November 2017	Alan Akbik	Talk
20e editie Data Donderdag - ING, NS, Growth Tribe, Valuemaat	GoDataDriven, Amsterdam, NL	November 2017	Richard Koopmanschap	Other
CWI Lectures on Machine Learning	CWI Amsterdam, NL	November 2017	Richard Koopmanschap, Joeri van Ruth	Promotional items
ADS Festive Drinks & Data: 2017 Highlights & Looking Forward to 2018	Amsterdam Business School, NL	December 2017	Ying Zhang, Pedro Ferreira	Other
Influx/Days	London, UK	January 2018	Ying Zhang, Pedro Ferreira	Other
Handelsblatt goes future: Artifical Intelligence" Conference	Munich, DE	February 2018	Alexander Löser	Presentation

ProductTank Berlin , Data Products	Mircosoft, Berlin, DE	March 2018	Alexander Löser	Presentation
SAP Conference on Machine Learning	Berlin, DE	March 2018	Roland Vollgraf	Talk
Shoptalk	Las Vegas, USA	March 2018	Matthias Dantone	Talk
Federal Minisitry of Economics: German Finish Information Exchange	Finish Embassy Berlin, DE	April 2018	Alexander Löser	Talk
Brussels TechSummit 2018	Brussels, BE	June 2018	Alan Akbik	Talk
K5	Berlin, DE	June 2018	Matthias Dantone	Talk
AI Expo Europe 2018	Amsterdam Rai, NL	June 2018	Martin Kersten, Svetlin Stalinov	Other
ADS Drinks & Data Summer Startup	Amsterdam Business School, NL	June 2018	Ying Zhang	Presentation
FashionTech	Berlin, DE	July 2018	Matthias Dantone	Presentation
Big Data Expo	Utrecht, NL	September 2018	Aris Koning, Svetlin Stalinov	Other
HiPEAC CSW Autumn 2018	Heraklion, GR	October 2018	Ying Zhang, Martin Kersten	Promotional items
EBDVF 2018	Vienna, AT	November 2018	Aris Koning, Martin Kersten, Ying Zhang	Promotional items
ACE startup meeting	Amsterdam, NL	November 2018	Ying Zhang	Other
Meeting organised by Dutch consulate to meet the local Big Data bureau and AI companies	ChongQing, CN	November 2018	Ying Zhang	Other
HiPEAC, European Network on High Performance and Embedded Architecture and Compilation	Valencia, SP	January, 2019	Ying Zhang, Martin Kersten, Aris Koning, Pedro Ferreira	Promotional items
H2020 Successful R&I in Europe 2019 - 10th European Networking Event	Düsseldorf, DE	February 2019	Ying Zhang	Presentation
Tagesspiegel 5th Digital Future Science 2019	Berlin, DE	March 2019	Alan Akbik	Presentation
Data Warehousing & Business Intelligence summit	Utrecht, NL	March 2019	Martin Kersten	Presentation
Data Warehousing & Business Intelligence summit	Utrecht, NL	March 2019	Martin Kersten	Presentation
FOX AI Summit	Köln, DE	May 2019	Alexander Löser	Talk
Swedish German Business Days (Swedish Embassy)	Berlin, DE	November 2019	Alexander Löser	Presentation
Austrian German Business Days (BMVIT and BMWi)	Berlin, DE	November 2019	Alexander Löser	Presentation

Table 4.7: Industrial events.

4.4.12 Other Dissemination Events

In Table 4.8 we report other dissemination events.

Event	Venue	Date/s	Attending representative	Type	Description
Startup Qualifiction.com	EXIST (BMWi)	Jul-17	Alexander Löser	Startup foundation	Text Mining for spotting bestsellers
Startup Beezdata.de	BerlinStartupGrant	Jan-18	Alexander Löser	Startup foundation	Matching NGOs and Trusts
Data Science at ASOS.com	ASOS.com HQ (London)	Aug-18	Paul Clough	Presentation	Presentation of FashionBrain and The University of Sheffield research in Data Science
FashionBrain with projectstarling.com	Online	Sep-18	Alessandro Checco	Presentation	Presentation of FashionBrain and collaboration plans
Crowdsourcing papers presentation	University of Queensland	Oct-18	Alessandro Checco	Presentation	Presentation of FashionBrain research in Crowdsourcing
IDEL Paper (D4.3)	IEEE BigComp2019	Feb-19	Alexander Löser	Presentation	Best Paper Award (145 submissions, 42 Accepted)

Table 4.8: Other events.

4.5 Sponsorships

We report here some events sponsored by FashionBrain, that have reached a vast audience.

4.5.1 Collective Intelligence 2018

The Collective Intelligence Conference Series is an interdisciplinary event that brings together researchers from academia, business, nonprofits, governments and the interested public to share insights and ideas relevant to understanding and designing and fostering the use of collective intelligence in its many forms.

The FashionBrain Project was a sponsor of Collective Intelligence 2018 and showcased a promotional booth as part of the exhibition (Figure 4.9).

Collective Intelligence 2018[3] was held in Zurich, Switzerland, on July 7-8, 2018 (its first time outside of the USA) along side HCOMP 2018, bringing together two interdisciplinary communities to foster new connections among collective intelligence and crowdsourcing researchers and developers.

[3] https://ci.acm.org/2018/

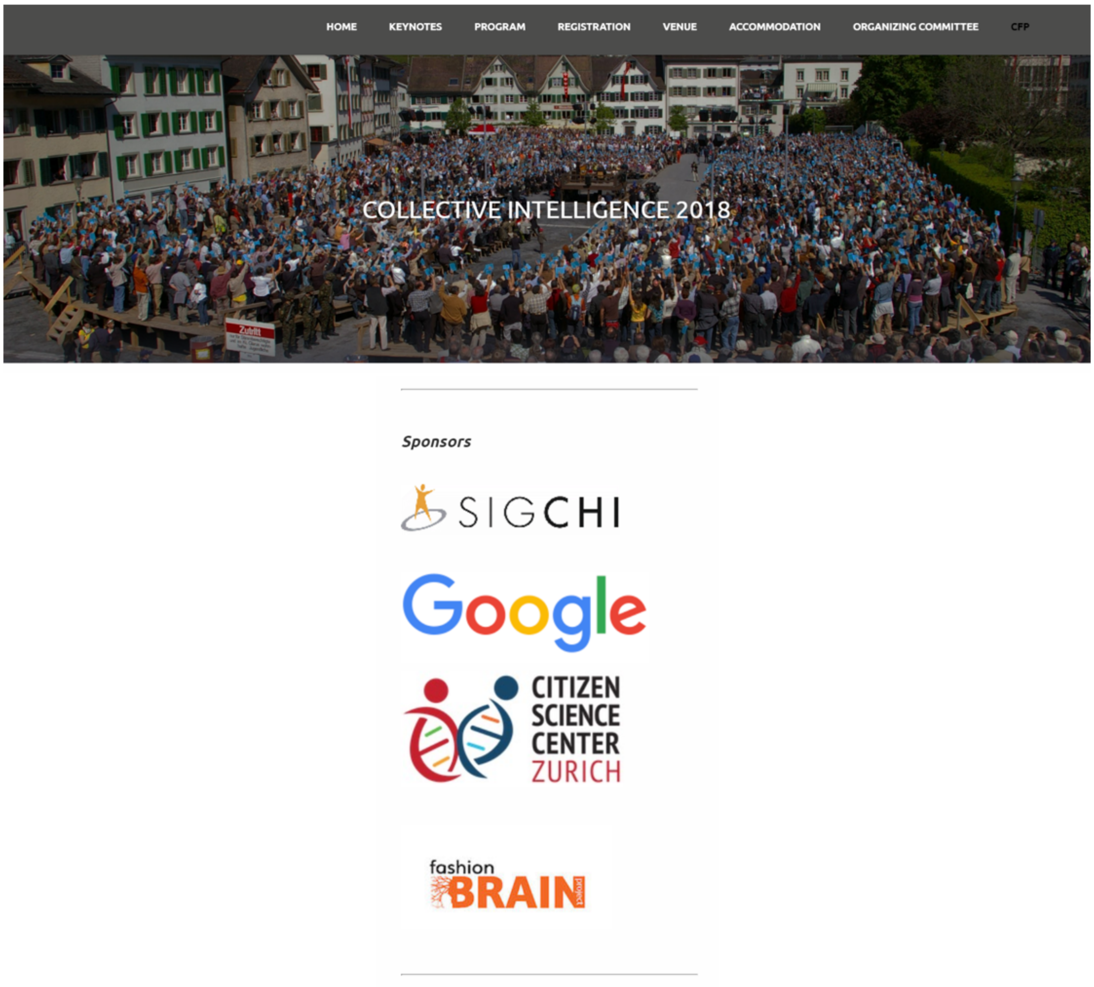


Figure 4.9: Sponsors at Collective Intelligence 2018.

4.5.2 First symposium on Biases in Human Computation and Crowdsourcing

The University of Sheffield, together with another H2020 PPP member (Qrowd, University of Southampton) organised the First symposium on Biases in Human Computation and Crowdsourcing 2019 https://sites.google.com/sheffield.ac.uk/bhcc2019 that has taken place on October 21-22 in Sheffield, UK. The goal of the symposium was to analyse both existing human biases in hybrid systems, and methods to manage bias via crowdsourcing and human computation. The attendees discussed different types of biases, measures and methods to track bias, as well as methodologies to prevent and solve bias. Some of the output of this symposium has been an important contribution for D3.4.

USFD and Qrowd provided a framework for discussion among scholars, practitioners and other interested parties, including industry, crowd workers and crowdsourcing platform managers.

The symposium had two keynotes: Michael Rovatsos, from the University of Edinburgh, and Ricardo Kawase from ebay. Moreover, the symposium had two invited speakers: Amrapali Zaveri, from Maastricht University, and Ujwal Gadiraju, from Leibniz University of Hannover.

During this symposium the FashionBrain project has been presented with an oral presentation and a poster. In Figures 4.10-4.15 we report some informations about the BHCC website and statistics on attendance.

Figure 4.10: Homepage of BHCC.

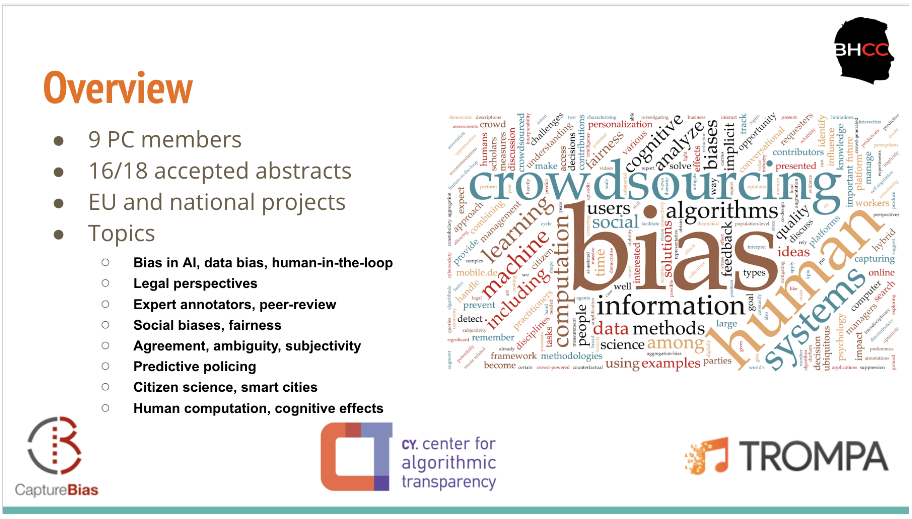


Figure 4.11: Summary of BHCC topics with wordcloud from the presented abstracts.

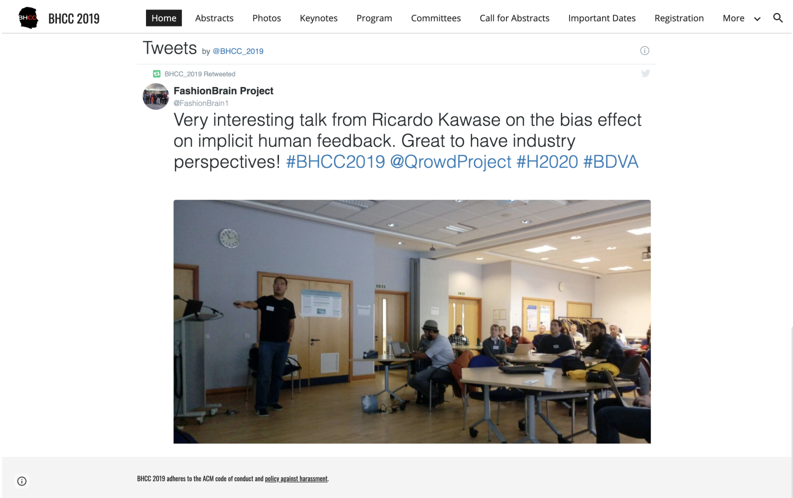


Figure 4.12: Twitter section of BHCC.

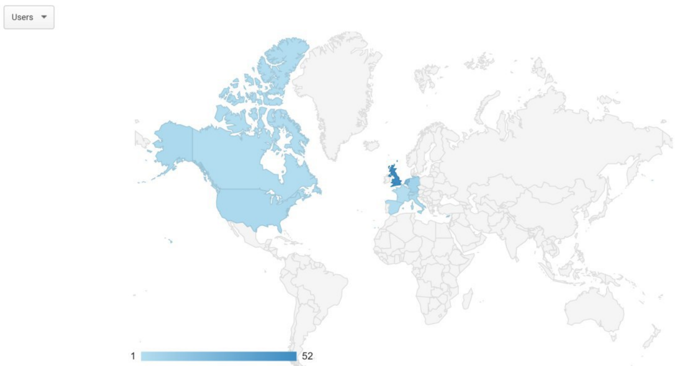


Figure 4.13: Provenance of BHCC participants.

	Country	Acquisition: Users	Acquisition: New Users	Acquisition: Sessions	Behaviour: Bounce Rate	Behaviour: Pages/Session	Behaviour: Avg. Session Duration	Conversions: Goal Conversion Rate	Conversions: Goal Completions	Conversions: Goal Value
		84 % of Total: 100.00% (84)	60 % of Total: 100.00% (60)	213 % of Total: 100.00% (213)	45.07% Avg for View: 45.07% (0.00%)	2.18 Avg for View: 2.18 (0.00%)	00:01:50 Avg for View: 00:01:50 (0.00%)	0.00% Avg for View: 0.00% (0.00%)	0 % of Total: 0.00% (0)	US$0.00 % of Total: 0.00% (US$0.00)
1.	United Kingdom	52 (54.17%)	29 (48.33%)	125 (58.69%)	44.00%	2.30	00:01:57	0.00%	0 (0.00%)	US$0.00 (0.00%)
2.	Netherlands	19 (19.79%)	14 (23.33%)	43 (20.19%)	53.49%	2.05	00:02:06	0.00%	0 (0.00%)	US$0.00 (0.00%)
3.	Germany	7 (7.29%)	4 (6.67%)	10 (4.69%)	40.00%	2.00	00:00:44	0.00%	0 (0.00%)	US$0.00 (0.00%)
4.	Cyprus	5 (5.21%)	4 (6.67%)	10 (4.69%)	50.00%	1.70	00:01:17	0.00%	0 (0.00%)	US$0.00 (0.00%)
5.	United States	4 (4.17%)	4 (6.67%)	4 (1.88%)	0.00%	2.25	00:00:09	0.00%	0 (0.00%)	US$0.00 (0.00%)
6.	Italy	3 (3.12%)	3 (5.00%)	7 (3.29%)	28.57%	2.57	00:00:45	0.00%	0 (0.00%)	US$0.00 (0.00%)
7.	Canada	2 (2.08%)	0 (0.00%)	6 (2.82%)	50.00%	1.67	00:02:11	0.00%	0 (0.00%)	US$0.00 (0.00%)
8.	Spain	2 (2.08%)	1 (1.67%)	4 (1.88%)	50.00%	2.00	00:04:04	0.00%	0 (0.00%)	US$0.00 (0.00%)
9.	Switzerland	1 (1.04%)	0 (0.00%)	2 (0.94%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	US$0.00 (0.00%)
10.	France	1 (1.04%)	1 (1.67%)	2 (0.94%)	0.00%	2.00	00:00:08	0.00%	0 (0.00%)	US$0.00 (0.00%)

Figure 4.14: BHCC website visits statistics.

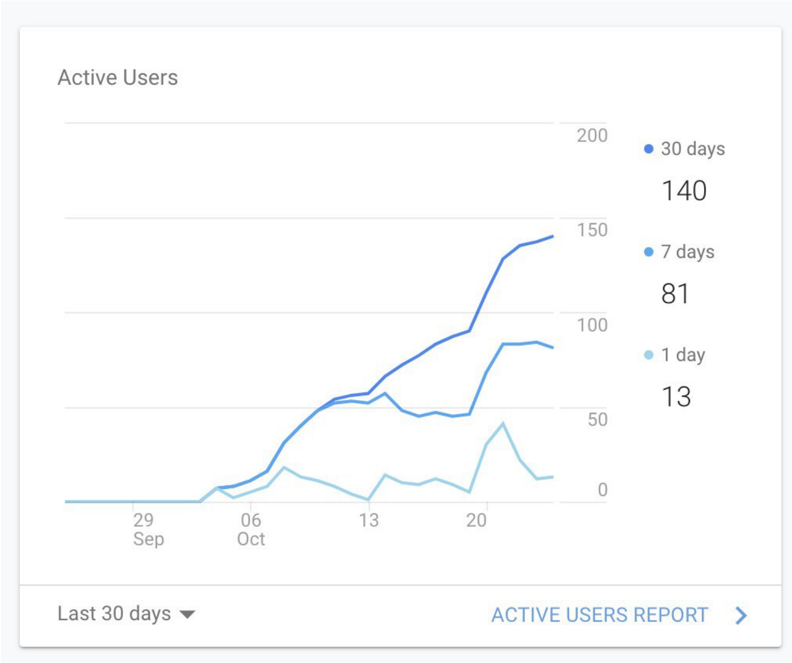


Figure 4.15: Active users on the BHCC Symposium website.

4.6 BDVA and H2020 Synergies

4.6.1 BDV Steering and technical committee participation

FashionBrain contributed substantially to the BDV PPP meetings:

- March 2017, Brussels: launch of BDV PPP Steering Committee.
- November 2017, Versailles: BDV PPP meeting.
- February 2018, Brussels: BDV PPP meeting.
- May 2018, Sofia: BDV PPP meeting.
- June 2019, Riga: BDV PPP meeting.

4.6.2 BDV Marketplace

Flair has been published in the BDV marketplace as a technical solution. This improves the visibility of the project and, at the same time, enhances the diversity of the marketplace.

4.6.3 Synergies - Qrowd

As shown in Figure 4.16, FashionBrain has started a synergy with Qrowd, another H2020 PPP partner, to share methodologies on crowdsourcing, and to join forces in the dissemination process.

Figure 4.16: Sponsor page of BHCC.

5 Conclusions

This deliverable presented the produced promotion and dissemination material, demonstration workflows, and the fully functional data integration infrastructure ready to be demo-able to the public. We grant the Commission the right to use the Showcase for its own dissemination and awareness activities (including Web based and electronic publications) after the completion of the project. The Showcase features a meaningful subset of the functionality characterizing the project demonstrators, along with relevant copyright notices and contact information, and information on the run-time procedures.

We also reported about project activities undertaken to support standardisation of project results and collaboration with other projects and relevant initiatives as well as the results of reaching-out by means of press, social media, open-source communities using demos, use cases, and benchmark results realized during the project. Finally, we reported on our contribution to the Big Data Value PPP activities.