Project acronym:	**EDSA**
Project full name:	**European Data Science Academy**
Grant agreement no:	**643937**

D4.4 Interim Community Engagement and Networking Report

Deliverable Editor:	**Gareth Beeston (SOTON)**
Other contributors:	**Elena Simperl (SOTON)**
	Christopher Phethean (SOTON)
Deliverable Reviewers:	**Alexander Mikroyannidis (OU)**
	Shatha Jaradat (KTH)
Deliverable due date:	**31/07/2016**
Submission date:	**29/07/2016**
Distribution level:	**Public**
Version:	**1.0**

This document is part of a research project funded
by the Horizon 2020 Framework Programme of the European Union

Change Log

Version	Date	Amended by	Changes
0.1	17/05/16	G Beeston	Website, Twitter and Process Mining MOOC details
0.2	17/05/16	G Beeston	Press coverage, real-world events
0.3	17/05/16	G Beeston	Ambassadors and real-world events
0.4	17/05/16	G Beeston	Dashboard, Review, Conclusion,
0.5	29/06/16	G Beeston	Minor additions before submission for review
0.6	18/07/16	G Beeston	Final revisions
0.7	19/07/2016	Elena Simperl	Scientific Review
1.0	27/07/2016	Aneta Tumilowicz	Final QA

Table of Contents

List of Tables

List of Figures

1. Executive Summary

The deliverable detailed in this document presents a summary of the communications and community engagement activities with the public by the respective partners of the EDSA project. These activities include the website, news pieces, newsletter, Twitter account, MOOCs, Search Engine Optimisation (SEO) and other learning activities, events and the ambassadors who have formally agreed to support the promotion and engagement of EDSA activities. The details within this deliverable document are written following the plans detailed in the community engagement plan outlined in D4.3 at M3 of the project to provide details and data on how the channels and engagement opportunities have been developed since and, based upon revision of the data, show how the plans have been adapted for optimum effectiveness in community engagement.

The aim of the D4.3 plan was to target a variety of channels in order to solicit feedback from stakeholders including researchers, industry-specialists, policy-makers and decision makers from a range of sectors. This deliverable therefore reports on the activities completed to date, and provides further insights into our online community, offline event attendance and the series of online courses managed by EDSA. By doing this we demonstrate the variety of events in which the EDSA consortium is participating in order to promote the project, and to source feedback and input on the data science skills gap. It also puts in to place the services and tools required in order to track the strength of the supporting community, and to monitor the dissemination rate of key EDSA learning materials, which will help us to continuously adapt to the changing needs and demands of the community as the project progresses. To conclude we also present a qualitative assessment of these results and what this means for our community-building work going forward with the remainder of the project.

2. Community Engagement Plan Recap

In D4.3 a real-world and online community engagement plan was presented, detailing the initial data regarding the communications and dissemination channels current at the time of writing the report. The aim of the plan was to outline a set of engagement tools and practices that the project could utilize in order to engage the community. These included a combination of both real-world events (such as offline, face-to-face meetings) and online communication and dissemination activities. Online activities were separated based on those that were broadcast-oriented and were primarily aimed to use to disseminate information, and those that were responsive and conversation-oriented and targeted at increasing communication and collaboration amongst the community.

Since this plan, a number of adaptations have been made to the community engagement and material dissemination plans that were detailed in D4.3. The adaptations have been developed due to a number of reasons including the fact that many courses, learning resources and course materials have since been launched and been engaged with by a large number of learners across the EU. Furthermore, LinkedIn activity has been reduced, due to the limited effectiveness of the channel, and others such as the Newsletter have been introduced to better target the respective audiences for the EDSA project. Additionally, SEO is now frequently used to ensure that the content, courses, website and other materials are more visible through search engine results. Utilising these new channels enables better targeting to be built into the plan, and discarding ineffective channels means that focus can be maintained on the channels that are effective for the project and the wider data science initiative. Ultimately this will allow us to maximize uptake of EDSA materials, and put in place a recognizable brand presence that will contribute to the sustainability and longevity of EDSA's legacy. As the course materials, learning resources and courses are increasingly launched and released throughout the project further adaptations will be made to the community engagement and dissemination plans. This is so as to ensure that appropriate target audiences are identified and are encouraged to engage with the content and enrol on the courses facilitated and produced by EDSA. The adaptations to the original community engagement plan are discussed in and throughout this report.

3. Channels and Activities

3.1 Website

The EDSA website (http://edsa-project.eu/) provides a centralized location for all of the project's activities. It offers a comprehensive overview of the main activities taken by project partners, along with details about the project and any activities planned for the future (Figure 1). Its design and structure was discussed in more detail in D2.1, when the website was setup for the project. Since then, we have added Google Analytics to the website, which allows us to measure statistics associated with the website usage, providing an insight into specific topics, features and functions that the community is particularly interested in.

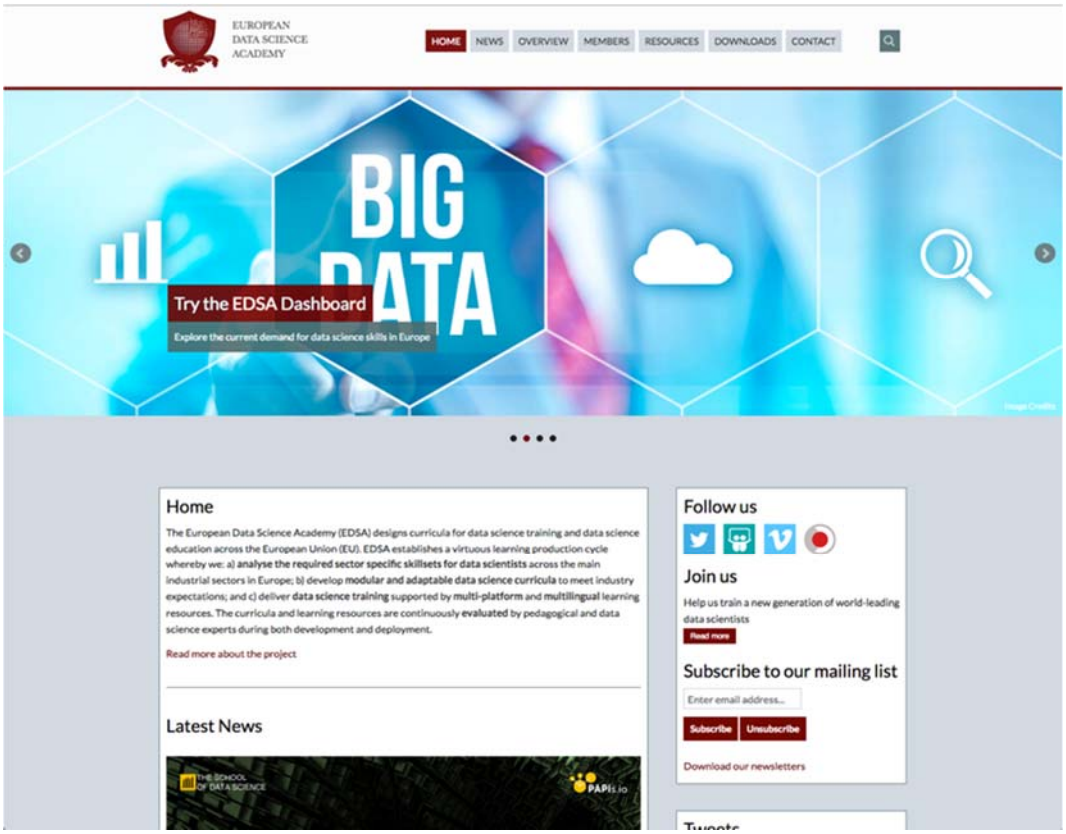

Figure 1 - EDSA project website homepage

3.1.1 Google Analytics and Traffic Data

Following a post-hoc analysis of the previous website traffic data sourced from server logs, it is evident that there were anomalies in these data. The anomalies in these data, supplied by the server logs, were then translated into the previous report, which presented a set of statistics that were unusually high for a website of this size and coverage. However, having identified these set of anomalies that skewed the figures from the previous report, and since using Google Analytics in this phase of the deliverable, it is evident that the figures shown in this report are a true representation of the traffic sourced for the website.

From the installation of Google Analytics, the following statistics have been gathered:

Table 1 - Website Visitor Statistics (Reporting Period: 16th April 2016 – 16th May 2016)

Metric	*Total*	*Avg./Day*
Sessions	1,335	45
Visitors	946	31
Page Views	3,317	110
Bounce Rate	56.63%	N/A
New Sessions	63.30%	N/A
Avg. Session Duration	2:22	N/A
New Visitors	63.4%	N/A
Returning Visitors	36.6%	N/A

Table 1 lists a number of general statistics for the website. The most revealing statistic is that the website received 946 visitors in the reporting period detailed above, averaging at 31 new visitors a day, with new visitors representing 63.4% of the traffic to the website. This shows that the website is encouraging a lot of new visitors to the website, from general promotion of the website and the project from a number of different channels detailed further in this document. It is also notable that the average session duration is 2 minutes and 22 seconds, representing an extended time on the website. This shows that the website has a wealth of 'sticky' content that retains the visitors' attention for a substantial amount of time. This statistic coincides with the relatively low bounce rate of 56.63%.

Figure 2 - Timeline of Sessions for 16th April 2016 – 16th May 2016

Table 2 - Website Visitor Country Statistics (1st Feb 2015 - 10th Jun 2015)

Country	*Hits*	*Percentage of Total (Hits)*
United Kingdom	317	23.75
Spain	115	8.61
Germany	87	6.52
Netherlands	73	5.47
Belgium	57	4.27
Brazil	53	3.97
United States of America	51	3.82
Italy	50	3.75
France	48	3.6
Portugal	47	3.52

Table 2 helps to reveal the international audience that the website is receiving. While at the moment the UK provides the highest number of hits to the website, there are also visits being attained from other EU member states including Germany, The Netherlands and Spain.

For much of the traffic, the primary landing page is the homepage, with other traffic landing at the news section of the website (Figure 3). Nearly 50% (524/1,335) of the total sessions have at least two interactions with the website. This shows that the traffic being sourced is interested in the content of the website, and of the project as a whole. The news sections of the website, that are frequently visited, are utilized primarily to disseminate information about the project, courses, events and reports that have resulted from the project, but also as a means to increase search engine rankings.

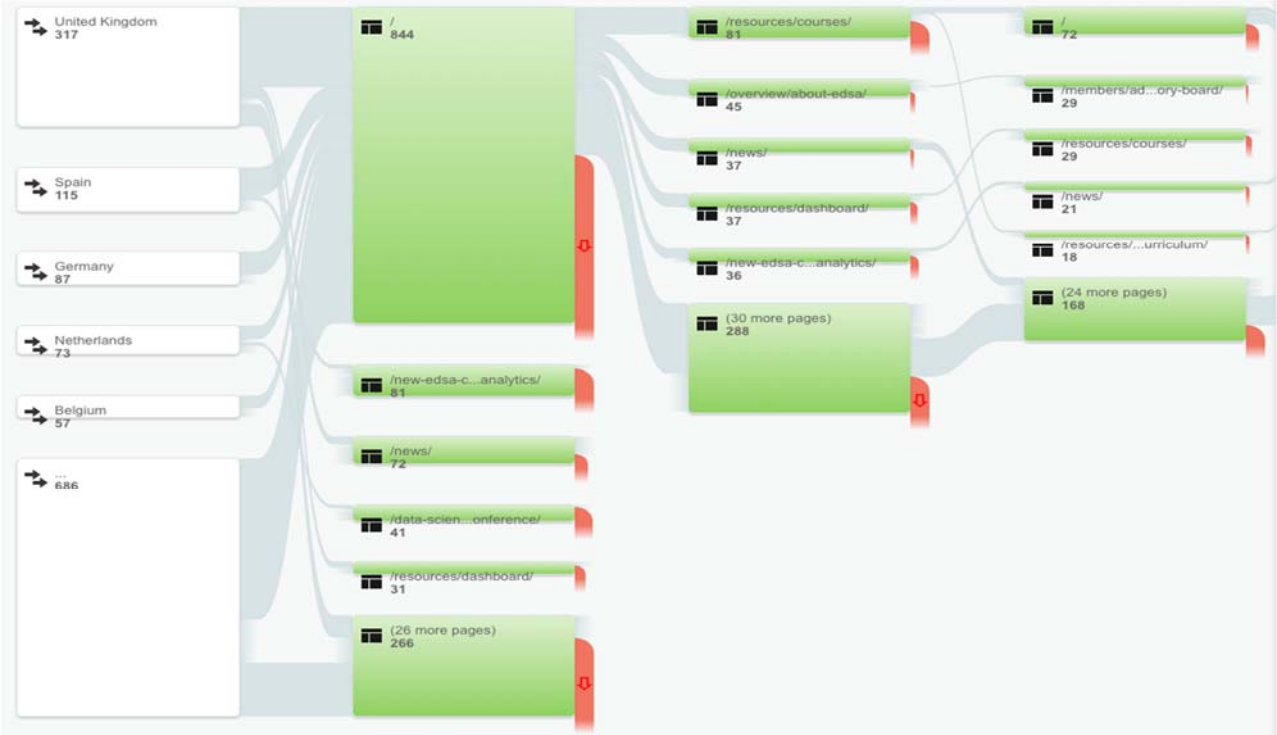

Figure 3 - Landing Pages and User Flow

3.1.2 Search Engine Optimisation for Google and Bing

As part of this deliverable, extensive Search Engine Optimisation (SEO) has been conducted to optimize the search rankings of the website and obtain a higher level of organic traffic from Google and Bing. A plugin installed into the WordPress website named All In One SEO, which enables the simpler implementation of on-page SEO, rich snippets, increased load speeds, keyword and keyphrase targeting, and content metadata. As a result of this optimization strategy, and the respective SEO methods, 36.3% of the traffic is now obtained from Organic Search traffic (Figure 4). SEO changes on a daily basis, and the rankings detailed below are subject to change due to the competitive nature of the keywords and keyphrases aimed for. With many organisations being founded that offer data science training, education and educational material, it is becoming increasingly competitive within the SEO space. The aims of the SEO are to continue to adapt, update and create content that enables the EDSA project to sustain a high ranking for the respective target keywords and keyphrases.

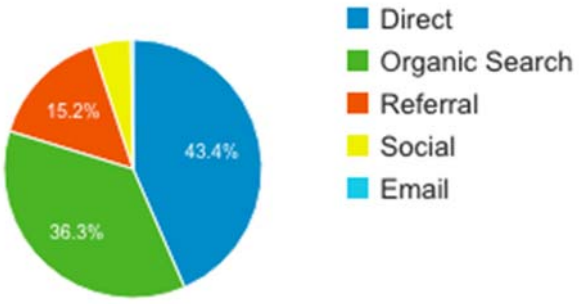

Figure 4 – Sources of Website Traffic

Table 3 – Search Engine Optimisation Keyphrase Ranking

Keyphrase	*Google Ranking*	*Bing Ranking*
Data Science Education	2nd	5th
Data Science Training	4th	-
Data Science Training Europe	-	1st
Data Science Education Europe	1st	-
Data Science Europe	3rd	5th

The keyphrases shown in Table 3 are the keywords optimized for within the Google and Bing searches. These keyphrases are focused upon due to extensive keyphrase research and planning that is aligned with the goals of the EDSA project. Due to the variability in the algorithms between the two search engine sources, it is evident that each would provide a different ranking. The optimization aspect of this deliverable will continue throughout the project in order to continue to optimize for the same set of keyphrases and a wider spectrum of keyphrases.

3.1.3 Videolectures

Lectures from videolectures.net have been implemented into the http://edsa-project.eu/video-lectures/ page of the website, which enables visitors to watch extended video lectures of experts in the field of data science and other related fields. These will be added to on a monthly basis, and currently consists of 8 video lectures selected by partners of the EDSA project. All credit for the recording and creation of the video lectures goes to videolectures.net.

3.2 News and Newsletter

The website has been the central channel through which to disseminate and distribute news associated with the EDSA project. The news feature on the website notes, on average, 2/3 news articles a month associated with the developments of the project, stories written by the EDSA ambassadors, new courses and other related news in the data science industry. This primary source supports the Search Engine Optimisation of the website, through utilizing the planned keyphrases and keywords.

Figure 5 - Operational Machine Learning Workshop

The news stories are distributed to all subscribers on the website mailing list (currently at 147) through an email newsletter, and also via the respective Twitter channel detailed below. The email newsletter is sent out every quarter starting January 1st.

Table 4 shows the number of subscribers to the email newsletter and the relative newsletter open and click-through percentages. As is evident, between the initiation of the newsletter in January, subscription has grown by 47 subscribers and continues to grow.

Table 4 – Newsletter Subscriptions

Newsletter Distribution Date	*Subscribers*	*Opens*	*Clicks*
1st January 2016	100	66%	18%
1st April 2016	119	50.8%	15.3%
1st July 2016	147 (TBC)	TBC	TBC

3.3 Twitter

The EDSA Twitter account (@edsa_project) has been actively used throughout the project up until this point. Using Twitter's built-in analytics service (http://analytics.twitter.com) it can be seen that the audience of the account has grown steadily to now stand at 840 followers as of 17th May 2016 (Figure 6). Using the same tool, we can also see the resulting engagement with tweets from the EDSA account (Figure 7) to get an idea about the size of the audience they reach, and how many people go on to interact with the message further.

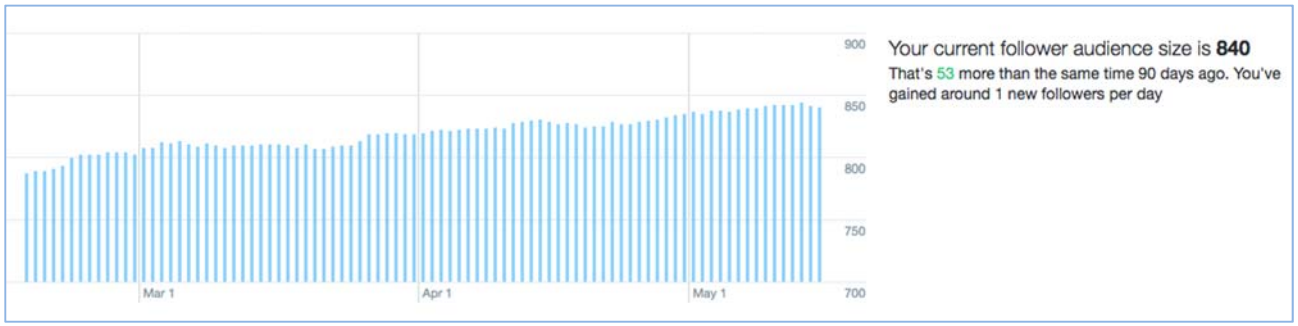

Figure 6 - Twitter Followers' Growth

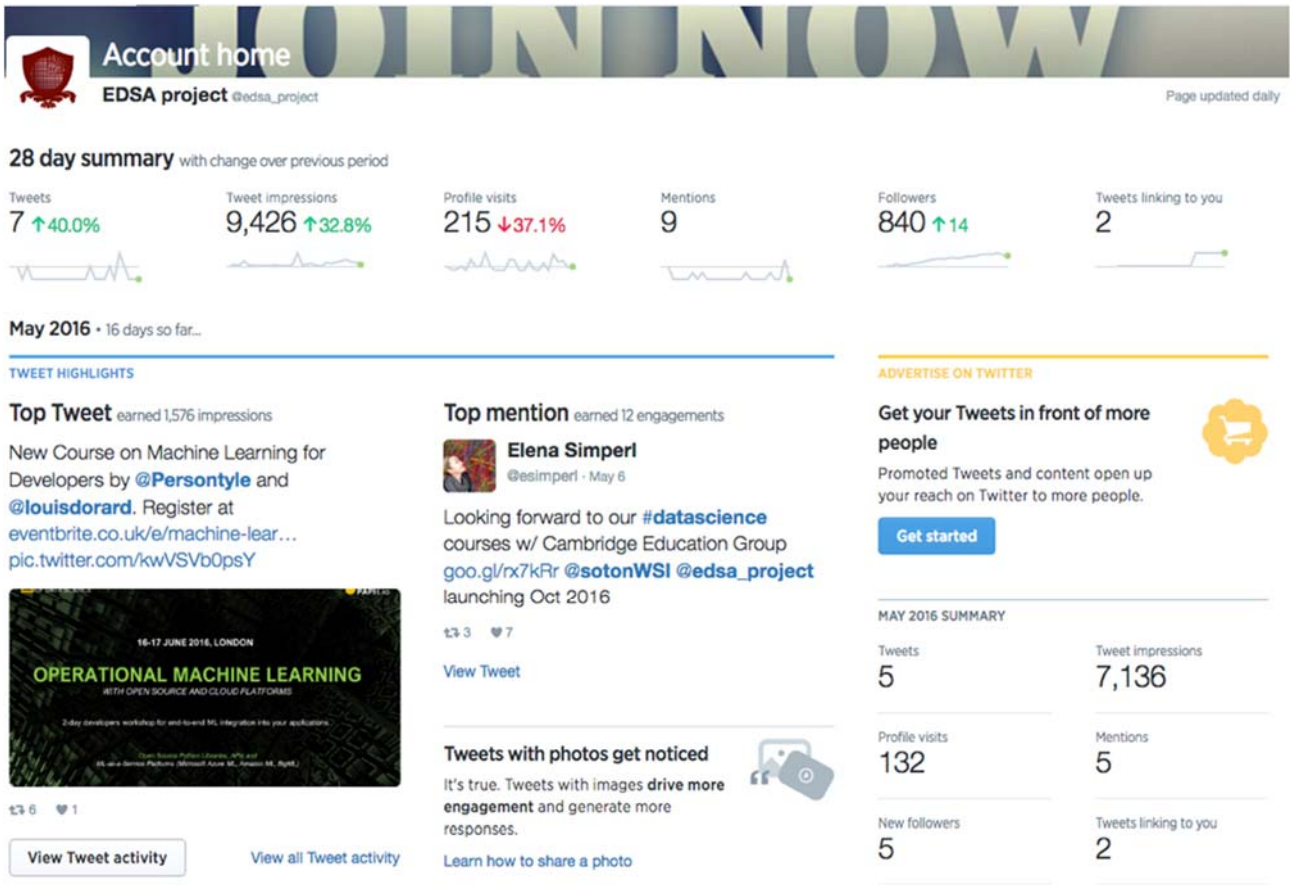

Figure 7 - Recent growth of EDSA Twitter audience

The top tweet since the set-up of the Twitter account earned a total of 1,576 impressions, which promoted a course in machine learning due to be hosted by one of EDSAs key partners, Persontyle. This also received a total of 6 retweets and 1 like. Figure 8 shows that the overall impressions for a month period has grown from the previously reported 2,800 impressions to 9,400 impressions, with engagements ranging from 2 to 20 engagements. The engagements are detailed within this statistic as retweets, likes, link clicks and media engagements. The increase in engagements is largely due to the increase in course and educational materials being made available through the various channels, which in turn, are being promoted through Twitter and the respective news page on the website. As identified using the "Top Tweets" tab of the Twitter analytics, 3 out of the top 5 tweets discuss the launch of new courses, managed and implemented by EDSA specifically. This shows a wider engagement involved with activities relating to course materials and learning in data science. As such, it is perhaps justified to primarily focus on these topics in subsequent tweets distributed through this channel.

Figure 8 - Tweet Engagement Statistics for May 2016

3.4 Courses Portal

The EDSA courses portal (courses.edsa-project.eu) provides a centralized location for all of the courses offered by the EDSA project. This includes courses developed by the EDSA project and courses owned and managed by affiliated institutions. It enables website visitors to enrol on courses and for the EDSA partners to check the progress of the learners. We utilise Google Analytics to track the visitors to the website and utilise the statistics from the Moodle platform to gather learner progression data for learning analytics.

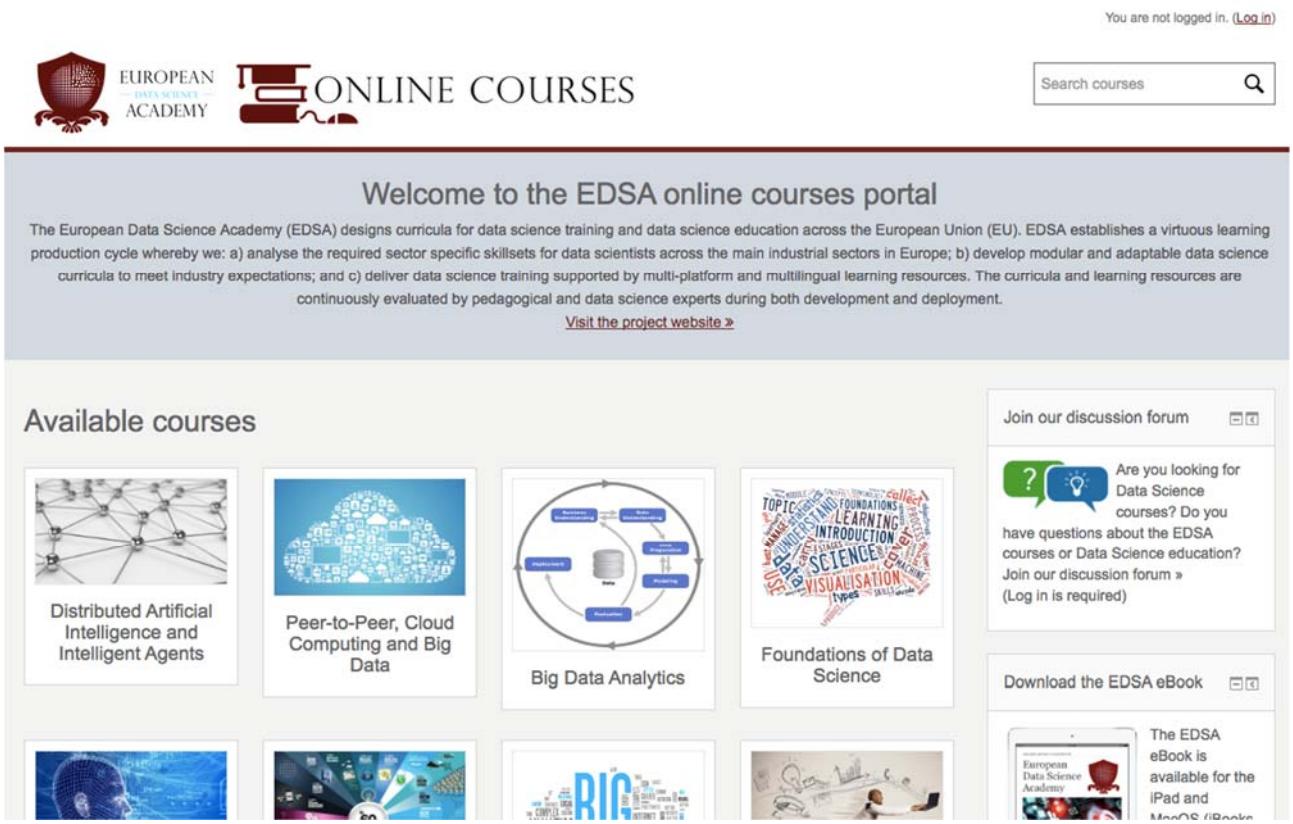

Figure 9 - EDSA Courses Portal Homepage

3.4.1 Courses Portal Google Analytics and Traffic Data

Table 5 - EDSA Courses Portal Analytics Traffic Data

Metric	*Total*	*Avg./Day*
Sessions	893	29
Visitors	599	20
Page Views	4,382	146
Bounce Rate	63.49%	N/A
New Sessions	63.30%	N/A
Avg. Session Duration	3:26	N/A
New Visitors	61.59%	N/A
Returning Visitors	38.4%	N/A

Table 5 lists a number of general statistics for the courses portal. Although, this portal comparatively has a lower amount of users and a higher bounce rate than that of the main website, the average session duration is longer by 1 minute and 4 seconds. This shows that the learners are having an extended amount of interaction with the portal itself, and the respective course material. It is projected that the amount of users will increase as more courses are implemented on the portal.

Figure 10 - Timeline of Sessions between 17th June 2016 – 17th July 2016

Table 6 - Website Visitor Country Statistics (1st Feb 2015 - 10th Jun 2015)

Country	*Hits*	*Percentage of Total (Hits)*
Russia	216	24.19
United Kingdom	198	17.45
Netherlands	63	7.05
Spain	57	6.38
Germany	39	4.37
Sweden	35	3.92
Kyrgyzstan	31	3.47
Brazil	25	2.80
Croatia	19	2.13
United States of America	13	1.46

Table 6 details the geographic location of the users and learners of the website, with the 4 our of the top 5 being from EU member states. While at the moment the United Kingdom provides the highest number of visitors to the portal, there are also many visits from The Netherlands, Spain, Germany, Sweden and Croatia.

3.5 Course Material

3.5.1 Process Mining MOOC (TU/e)

While the majority of course content is in development following the curricula design in D2.1, the first EDSA-marketed course has been run in the form of a MOOC. TU/e ran the second iteration of their Process Mining MOOC on Coursera from April – June 2015, which was purposely branded as an EDSA MOOC. This course is due to run again from the 11th July 2016 for a period of 8 weeks on the FutureLearn

platform (https://www.futurelearn.com/courses/process-mining), due to the evident success of the previous versions of the course (Table 7 (see also D4.2)).

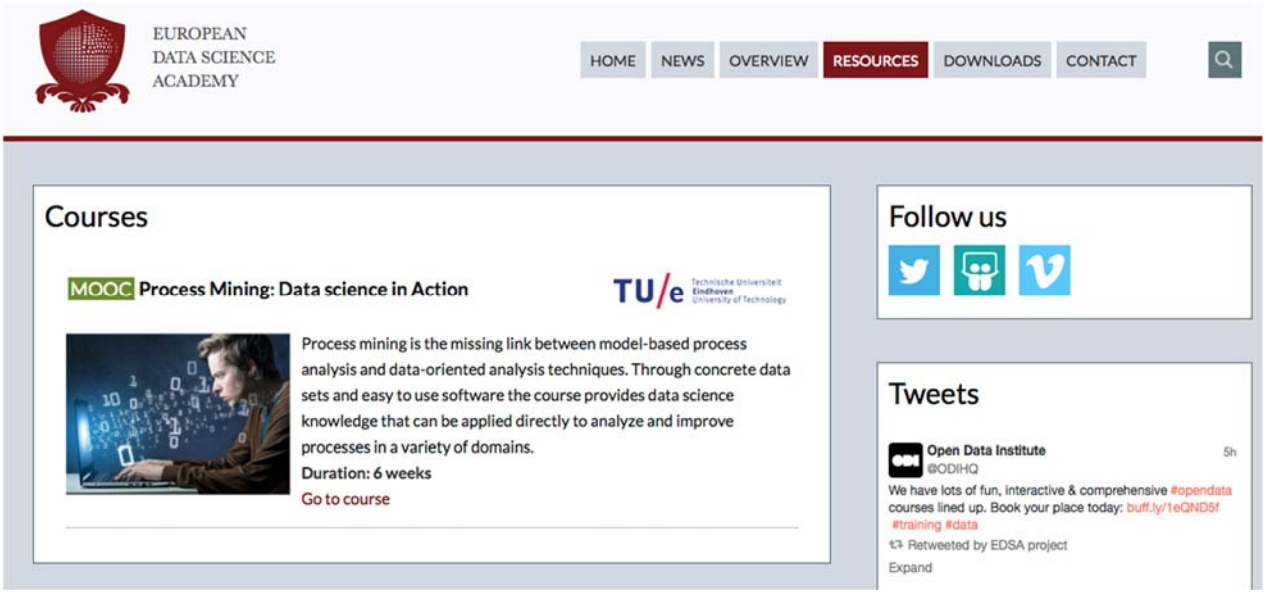

Figure 11 - EDSA website listing of Process Mining MOOC

Table 7 - Audience Statistics for Process Mining MOOC

Metric	Session 2
Start/end dates	April 1 '15 – June 3, '15
# Registered	24558
# Visited course page	17089
# Watched a lecture	11631
# Browsed forums	2899
# submitted an exercise	3212
# Certificates (normal/distinction)	916 (572 / 344)

3.5.2 Introduction to Linked Data and the Semantic Web

Subsequently a course has been run titled: "Introduction to Linked Data and the Semantic Web" (Figure 12) as part of the program of courses as a result of the work of the EDSA project. The University of Southampton ran the Massive Open Online Course (MOOC) on the FutureLearn platform for a duration of 3 weeks, between the 11th of April and the 2nd of May. Although run by the University of Southampton, this course was branded as part of the EDSA project as was promoted as such. Audience statistic shown in Table 8 shows that the course received over 4,700 learners, many of which continued to complete the course.

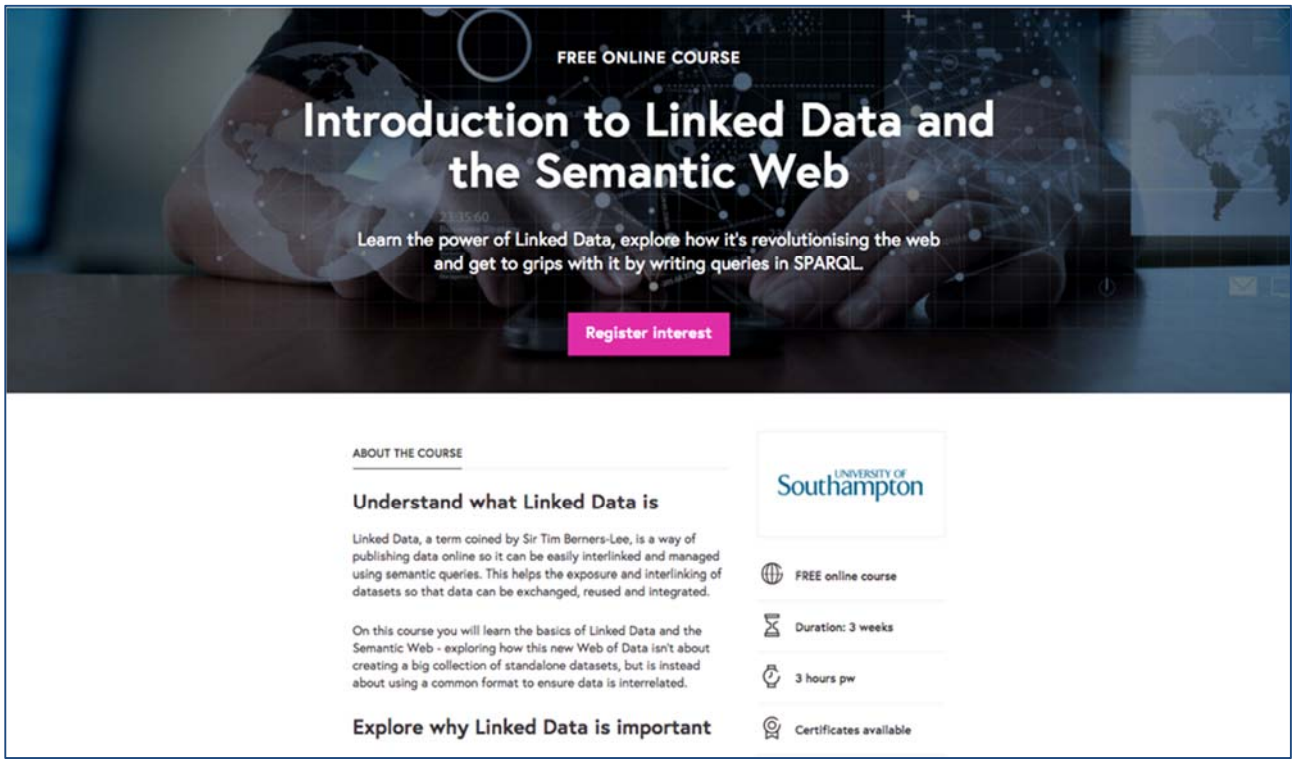

Figure 12 - Introduction to Linked Data and the Semantic Web FutureLearn MOOC

Table 8 - Audience Statistics for Introduction to Linked Data and the Semantic Web

Metric	Session 1
Start/end dates	April 11 '16 – May 2 '16
# Registered	4,715
# Visited course page	28,004
# Completed steps	22,455
# Average completed steps per stage per user	10.76
# Comments	2,361

Other courses have been added to the EDSA courses portal (http://courses.edsa-project.eu/) such as the self-study course in Big Data Analytics, Foundations of Big Data and Big Data Architecture courses amongst many others. At the time of writing this deliverable (June 2016), the EDSA courses portal contains 9 self-study courses, 3 MOOCs, 3 blended courses and 15 face-to-face courses. The current number of registered users is 345. Table 9 shows the statistics of enrolment on all other courses on the courses portal. Insights and learning analytics from the EDSA self-study courses and MOOCs can be found in D3.3.

Table 9 - Course Enrolment Statistics

Course	Enrolled users	Average Learner Progress
Foundations of Big Data	116	60%
Essentials of Data Analytics and Machine Learning	108	5%
Big Data Architecture	67	30%
Big Data Analytics	65	0%
Process Mining	33	2%
Foundations of Data Science	24	14%
Distributed Computing	19	2%
Distributed Artificial Intelligence and Intelligent Agents	5	0%
Peer-to-Peer, Cloud Computing and Big Data	2	0%

3.5.3 SlideShare Materials

In addition to the online courses, presentation files are placed within the EDSA SlideShare account (http://www.slideshare.net/edsa-project/) in order for presentation materials to be accessed publicly. Currently the account has a total of 7 Slide shares, 22 followers, 12 downloads with a total of 1808 views and 14 sharable actions in the past year (Figure 13).

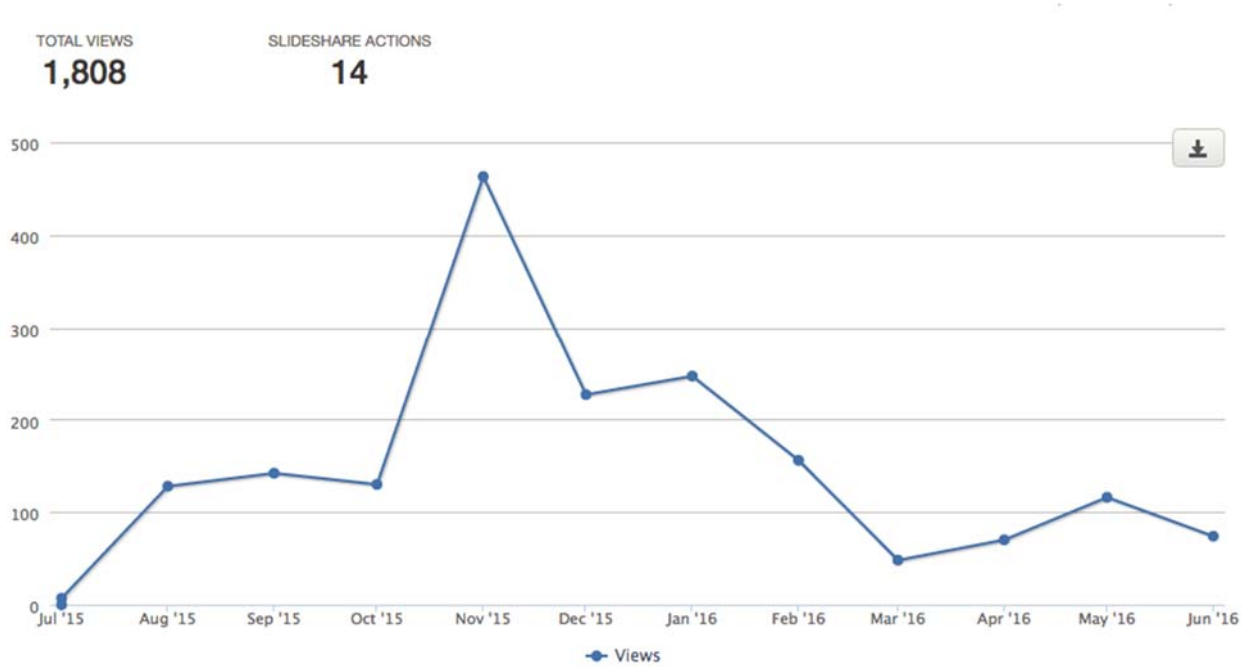

Figure 13 - SlideShare View Analytics

Figure 13 shows the visits to the SlideShare account over the previous year, with one month (November 2015) receiving 463 visits. This was largely as a result of the upload of the Top Content SlideShare that received 202 views within that month.

Table 10 - SlideShare Statistics

Metric	Statistic
Total Views	1808 Views
Top Content	
Dissemination and Engagement	560 Views
Downloads	12
Top Countries	United Kingdom (370 Views), United States (356 Views), France (122 Views)

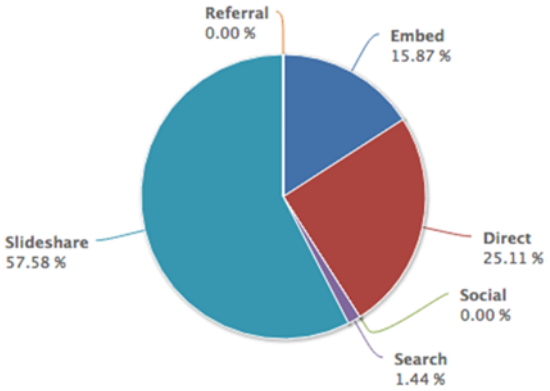

Figure 14 - SlideShare Traffic Sources

From Figure 14 57.58% of traffic stems from users coming from other sources on SlideShare, representing 1,041 of the total views. Alternatively, users are finding the contact through Embedded links (15.87%) and Direct links (25.11%). It is evident that more sharing of SlideShare needs to be performed through social media as there are currently no views from any social links.

3.5.4 Vimeo

The EDSA project has developed a Vimeo account (https://vimeo.com/edsa) through which to disseminate video content to users and learners. As of July 2016, the Vimeo account has 18 videos about data science-related topics including Agent Coordination and Agent-Orientated Software Engineering. Since the implementation of Vimeo, the videos have had a total of 68 plays and 28 finishes (Figure 15).

Figure 15 - Vimeo Statistics

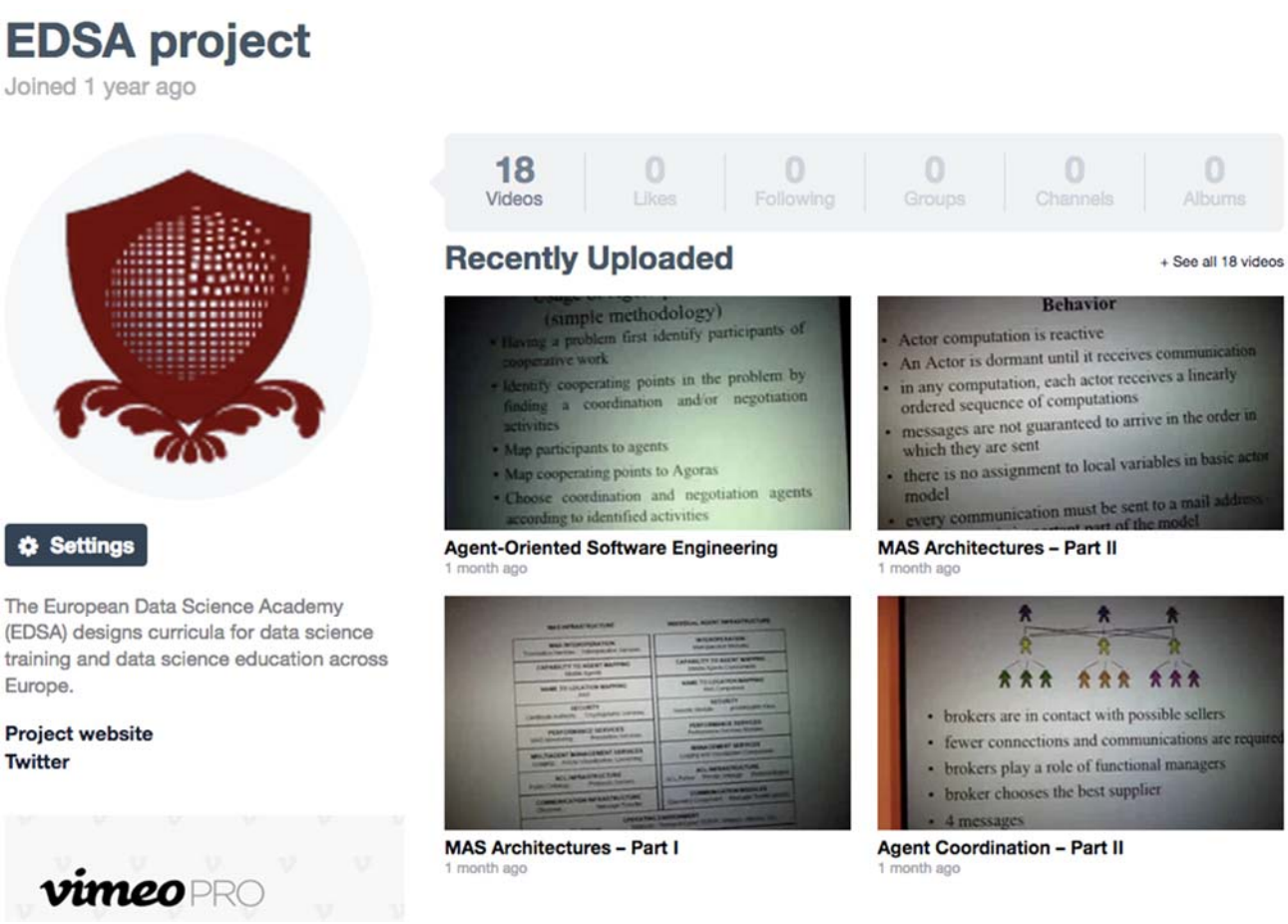

Figure 16 - EDSA Vimeo Homepage

3.6 Real-World Events

One of the main points of D4.3's real-world and online community engagement plan was an emphasis on promoting EDSA through a combination of online communication and real-world event attendance. Partners from across the consortium have been attending such events since the start of the project, speaking about and promoting the work that EDSA is carrying out. Details of these events are shown in

Table 11. As discussed in the real-world and online community engagement plan (D4.3), these events cover two important areas related to the project: data-science themed events, and e-Learning themed events. Therefore this categorization is continued within

Table 11 to demonstrate the variety of events that the consortium has been participating in.

Table 11 – EDSA Real-World Event Participation

Event	*Date*	*Location*	*Event Type*	*Attended By*	*Interaction With Event*	*Number of Attendees*	*Link*
Data science focus group, UK government organisation	January, 2015	London, UK	Focus Group	ODI	Ran focus group		N/A
LAK'15	March, 2015	Poughkeepsie, USA	Conference (e-Learning)	Fraunhofer IAIS	Paper presentation		http://lak15.solaresearch.org
EIT ICT Labs Partner Event	April, 2015	Trento, Italy	Meeting (data science)	KTH	Presentation		https://www.eitdigital.eu/news-events/events/article/eit-ict-labs-partner-event-2015/
Smart Data - Deutschland und Europa auf dem Weg zu einer digitalen Datenökonomie	April, 2015	Berlin, Germany	Conference (data science)	Fraunhofer IAIS	Ran booth on data science training and promotional material sharing		http://www.bmwi.de/DE/Service/veranstaltungen,did=693428.html
PAPIS Connect	May, 2015	Paris, France	Conference (data science)	Persontyle	Project details disseminated		http://www.papis.io/connect
VOR Divsion ICT and HE Symposium Research on MOOCs	May, 2015	Wageningen, Netherlands	Research symposium (e-Learning)	TU/e	Presentation		http://www.vorsite.nl/nl/divisies-en-themagroepen/vor-divsion-ict-and-he-symposium-research-on-moocs-26-mei-2015.html
ODI Open Data Meetup	Feb 2015, May 2015	London, UK	Meet-up (data science)	ODI	Ran meet up		http://www.meetup.com/Open-Data-London/
Open University visit by UK HE Commission	June, 2015	London, UK	Meeting	OU	Presentation		N/A
Data Literacy Workshop at WebSci'15	June-July 2015	Oxford, UK	Conference (e-Learning, data science)	UoS	Presentation and flyer dissemination		http://www.dataliteracy.eita.org.br/

Data science focus group, UK health organisation	July, 2015	London, UK	Focus Group	ODI	Ran focus group		N/A
ESWC Summer School	September, 2015	Kalamaki, Greece	Summer School	UoS, OU	Ran summer school		http://summerschool2015.eswc-conferences.org/
ODI Network Event	September, 2015	London, UK	Networking event	ODI	Organised event		
ICT 2015	October, 2015	Lisbon, Portugal	Conference (Data, IT, Innovation)	Fraunhofer, UoS, OU	Conducted interviews, promoted dashboard and ran workshop. Showing insights of the European data market that we have acquired.	6,000	https://ec.europa.eu/digital-single-market/en/ict2015
Interdisciplinary Mathematics	October, 2015	Uppsala, Sweden	Meetup	Persontyle	Attended		N/A
VOILA, 2015	October, 2015	Bethlehem, USA	Workshop	OU	Presented paper		http://voila2015.visualdataweb.org/
Deep Learning London	November, 2015	London, UK	Meetup	Persontyle	Organised meetup		http://www.meetup.com/Deep-Learning-London/
European Data Forum	November, 2015	Luxembourg, Luxembourg	Conference (data science)	UoS, OU, Fraunhofer	Moderated session, ran session and presented dashboard. Showing insights of the European data market that we have acquired.		http://2015.data-forum.eu/
ODI Pre-Summit Training Day	November, 2015	London, UK	Training	ODI	Ran training		N/A
ODI Summit	November, 2015	London, UK	Summit – open data	ODI	Organised event		N/A
EDSA PMB	December, 2015	London, UK	Meeting	All	Organised event	30	N/A
E-Skills 2015	December, 2015	Luxembourg, Luxembourg	Conference	OU	Panel member		http://eskills4jobs.lu/programme

SIMPDA 2015	December, 2015	Vienna, Austria	Symposium	TU/e			http://simpda2015.di.unimi.it/
e-Leadership Skills Expert Workshop	January, 2016	Brussels, Belgium	Workshop (data science)	UoS	Attended workshop		N/A
EDISON Workshop	February, 2016	Brussels, Belgium	Workshop (data science)	UoS, OU	Attended workshop		http://edison-project.eu/data-infrastructure-competences-and-skills-framework-european-and-global-challenge
Big Data Summit	February, 2016	Hanau, Germany	Conference (big data)	Fraunhofer	Attended	500	http://www.bitkom-bigdata.de/
Fraunhofer's Big Data Alliance	February, 2016	Sankt Augustin, Germnay	Annual meeting (data science)	Fraunhofer	Attended	40	N/A
PAPI's Connect	March, 2016	Valencia, Spain	Conference (IT)	Persontyle	Attended		http://www.papis.io/connect/
ODI Network Event	March, 2016	London, UK	Networking event	ODI	Ran event		N/A
Open Data Science	March, 2016	Lancaster, UK	Workshop	ODI	Attended		N/A
Data science focus group	March, 2016	London, UK	Focus group	ODI	Ran focus group		N/A
BigDataValue Association	March, 2016	The Hague, Netherlands	Summit	Fraunhofer	Attended		http://www.bdva.eu/
WWW Conference	April, 2016	Montreal, Canada	Conference (Web)	OU, UoS	Ran workshop		http://microposts2016.seas.upenn.edu/
Web Science 2016, Data-Driven Innovation Workshop	May, 2016	Hannover, Germany	Workshop (innovation, data science)	UoS	Ran workshop and distributed flyers	20	http://websci16.org/
ODI Node Gathering	May, 2016	London, UK	Networking event	ODI	Organised event		N/A

EDSA PMB	June, 2016	Dubrovnik, Croatia	Meeting	All	Organised event	30	N/A
European Data Forum	June, 2016	Eindhoven, Netherlands	Conference	UoS, OU, Persontyle, TU/e	Organised event, presented project		http://2016.data-forum.eu
EDISON Data Science Champions Conference	July, 2016	Brockenhurst, UK	Conference (data science)	UoS	Keynote	50	N/A
Essex Big Data and Analytics Summer School	September, 2016	Essex, UK	Summer School	ODI	Organised event		N/A
ODI Pre-Summit Training	October, 2016	London, UK	Training	ODI	Ran event		N/A
ODI Summit	November, 2016	London, UK	Summit – open data	ODI	Organised event		N/A
ODI Pre-Summit Training	October, 2017	London, UK	Training	ODI	Ran event		N/A
ODI Summit	November, 2017	London, UK	Summit – open data	ODI	Organised event		N/A

A key event that was promoted extensively following the event was the ICT 2015 conference in Lisbon. This event was covered extensively in Twitter (e.g. Figure 17) and through the news section of the website, both prior to the event (http://edsa-project.eu/ict-2015-networking-session-addressing-the-big-data-and-data-science-skills-gap/) and after the event (http://edsa-project.eu/addressing-the-big-data-and-data-science-skills-gap-at-ict2015/) in order to encourage registration to the conference and as a form of press release, respectively. This event was specifically focused upon as it was an event where the dashboard, and the project as a whole, was promoted extensively. Furthermore, interview data was collected at this event for the respective WP.

Figure 17 - Twitter during and following ICT2015

3.1 Press Coverage

Since the start of the EDSA project, a number of articles have been released to the press detailing the project's scope, aims and ambitions (**Error! Not a valid bookmark self-reference.**). These initially began with press releases from partners regarding the start of the project, along with news articles explaining what the funding was going towards. More recently, our lead partner of EDSA, the Open University, contributed to a policy report on the use of data in higher education. D4.2 reported on the press coverage up until April 2015, whereas more recent additions to the press coverage are shown after this date in Table 11.

Table 12 – EDSA Press Coverage

Date	*Coverage Type*	*Link*
November 2014	Press Release (ODI)	http://theodi.org/news/odi-helps-unlock-11m-for-open-data-innovation-in-europe
November 2014	News Articles	For example: http://www.theguardian.com/technology/2014/nov/04/eu-commits-144m-to-support-open-data-across-europe
January 2015	Press Release (OU)	http://www3.open.ac.uk/media/fullstory.aspx?id=28449
February 2015	Press Release (Southampton)	http://www.southampton.ac.uk/mediacentre/news/2015/Feb/15_35.shtml#.VYPhL2BhhRo
March 2015	Course Advertisement – Process Mining MOOC (TU/e)	http://www.kdnuggets.com/2015/03/coursera-process-mining-data-science-action.html
April 2015	Online Article (Tech.eu)	http://tech.eu/features/4339/data-science-big-data-europe/
January 2016	Policy Report	http://www.policyconnect.org.uk/hec/research/report-bricks-clicks-potential-data-and-analytics-higher-education
January 2016	Press Release (Guardian)	http://www.theguardian.com/education/2016/jan/19/big-brother-universities-data-higher-education-students
February 2016	Online Article (ODI)	http://theodi.org/blog/what-is-a-data-scientist-odi-trainer-answers-our-questions

3.2 Ambassadors

Since the previous report associated with the dissemination and engagement activities in the EDSA project, the partners within the project have recruited a set of 'EDSA Ambassadors' whom are professionals, experts and academics within the field of data science in the EU. They are situated within a wide spectrum of EU countries, and enable better positioning and promotion of the EDSA project in that respective country. As such, encouraging a set of data science experts in different locations across the EU to become ambassadors enables the EDSA project to extend its reach and target untapped markets of data science professionals within that particular country. The list of current ambassadors can be seen in

Table 13 - Ambassadors The ambassadors are to be promoted on our website through the Ambassadors section which will detail their respective experience and expertise in the field of data science, and other fields. Although not all ambassadors are currently found on the website, the initial set of ambassadors can be found at http://edsa-project.eu/members/ambassadors/.

Table 13 - Ambassadors

Ambassador Name	*Role/Experience*	*Constituent Country*
Mariana Damova	CEO of Mozaika Ltd (Data Science SME)	Bulgaria
Joseph Dureau	Data Scientists at Snips and Co-Founder of Standard Analytics	France
Martin Lochman/Milena Dobreva	Researcher/Professor of Data Analysis	Malta
Hugo Silveira Da Cunha	Founder and CEO of BIN	Portugal
Rosa Maria Martin	IT Technical Director of inLab FIB	Spain
Stefano Pagnano	IT Project Manager in Business Intelligence and Analytics	Italy
Ali Khan	Data Analysist for American Express	United Kingdom
Steve Brewer	Lead Project Partner at EDISON	United Kingdom
Gyorgy Szabo	Founder Open Data Portal, Debrecen Hungary	Hungary

4. Dashboards

As part of the project, and in order to disseminate the work completed in WP1, we are developing a number of visualisations through the dashboard section of the EDSA website (http://edsa-project.eu/resources/dashboard/) to present and disseminate information about EDSA, and the current market associated with data science in the EU. The EDSA dashboard enables visual discovery using different perspectives on a dataset containing about 300,000 job postings in Data Science across the EU, highlighting skills required of each role. The data was generated based upon new job positions posted online between October 2014 and May 2016, enriched and annotated to aid visual browsing and querying.

Figure 18 enables users to view the locations of the data science jobs in the EU, and the respective data science skills required as part of the application process (Figure 18). This will enable data science professionals to better understand the demands of the industry and how it varies across the EU. The secondary visualization depicts a wordcloud associated with the data science jobs, showing that the size of the particular skill is proportional to the demand in the industry. The final visualization is a representation of the skills required per country within a specific job – for instance data science manager (Figure 19). This visualization is currently under development and has thus far not been made available to the general public. The launch of this visualization is planned by month 24 of the project.

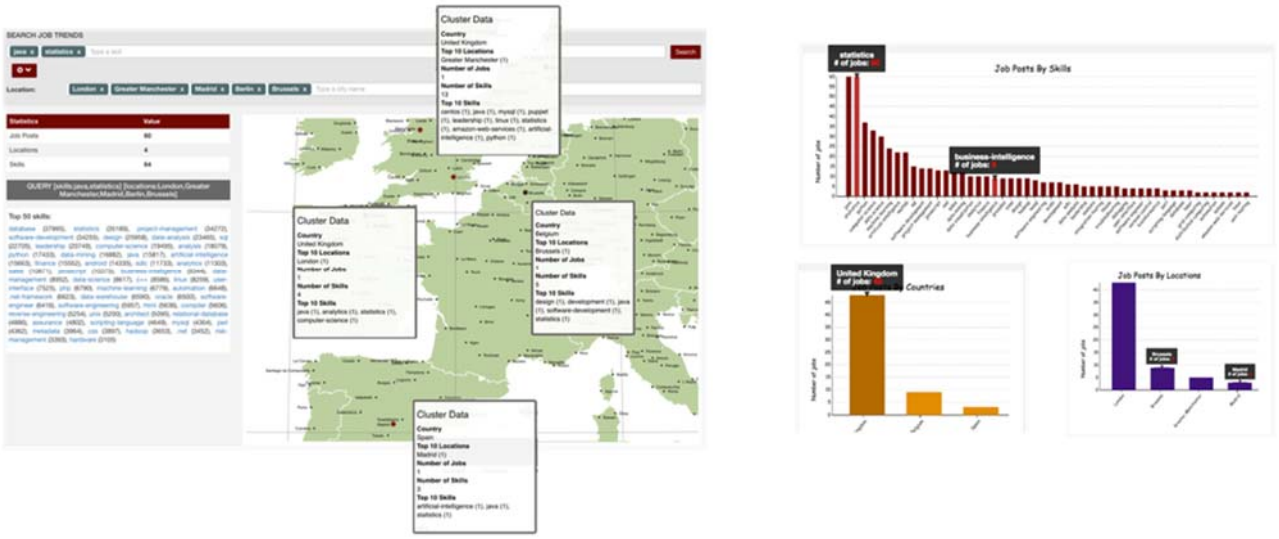

Figure 18 - Location and Job Skills Map

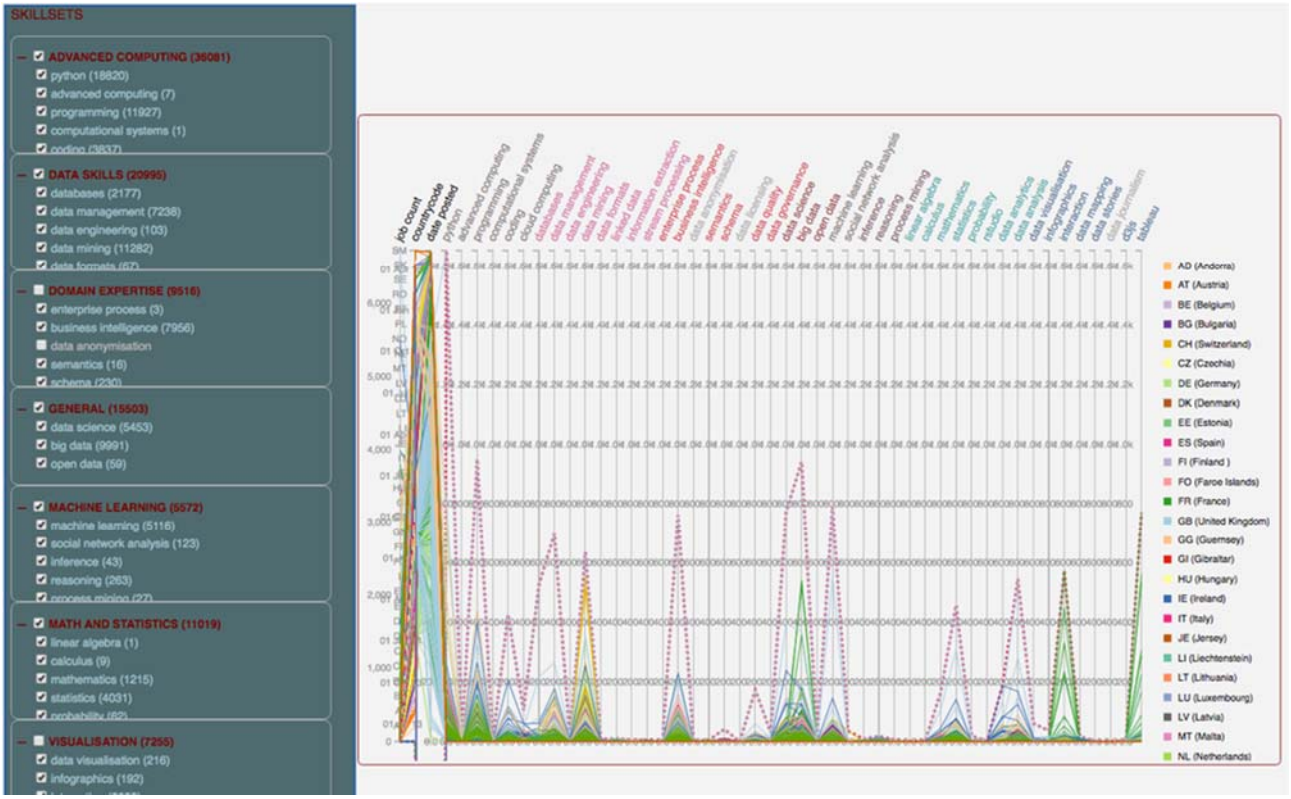

Figure 19 - Skills Viewer Visualisation

5. Insights, Reflections and Revisions to the Engagement Plan

The statistics, data and visualisations within this report show a summary of the current array of tools and channels that are being used in order to support the engagement and dissemination WP within the EDSA project. The statistics and data collected from each of the channels have been reviewed since the initial plan, and adapting accordingly. This section of the report details a summary of the insights obtained from this data collection and analysis, reflections on the levels of engagement with each of the aforementioned channels, and a summary of the revisions that have thus far been conducted, but also how the engagement channels may be adapted in order to further develop the engagement and dissemination activity.

As identified in 3.1.1, the website statistics in the previous report were not true representations of the website traffic. This was due to a discrepancy in the statistics that skewed the results reported from the server logs. Learning from this, the community management team have examined this discrepancy and utilized a clearer tool to obtain statistics relating to the website, namely Google analytics. The Google analytics reported above shows a userbase in line with the expectations of a website of this size, coverage and topic (Table 1, Table 2, Figure 2, Figure 3, and Figure 4), averaging 31 users a day, viewing 110 pages. It was identified that the bounce rate is relatively low and an average session duration of 2 minutes and 22 seconds, depicting an engaged audience. Following these data, the use of increasingly effective channels for promotion, and in-line with the better visibility amongst search engines (Table 3), it is expected that the traffic will continue to increase throughout the course of the project.

The high ranking for certain keywords depicted in Table 3, and the relative amount of traffic from organic search (36.3% of the total traffic), depicts the effectiveness of the SEO campaign that is being run in parallel to the planned dissemination activities. Firstly, it is imperative that we attract other EU member states to increase their visits in order to address the skills gaps across the European Union, and not just in the English-speaking markets. This is currently an issue as the keywords/keyphrases thus far that have been targeted have been in the English language and the website perhaps would benefit from multilingual keyword targeting. Additionally, from reviewing the keyword/keyphrase visibility on the primary search engines of Bing and Google, it is evident that the keyword visibility is subject to change very often. This is largely due to the competitive nature of the keywords within this particular topic. For instance, achieving high rankings for keyphrases such as “data science courses” is very difficult as longer-established websites (e.g. university websites) have the advantage of being established before the EDSA project. In order to overcome this issue, the EDSA community management team will increase the frequency of the creative and original content being distributed through the news section of the website. Incorporating this tactic, amongst others (e.g. further keyword research, metadata creation, web optimization etc.) will further develop the SEO of EDSA, and encourage further engagement with the website.

This report also sees the introduction of the videolectures and newsletter to the dissemination and engagement marketing mix. The videolectures will be added to on a monthly basis to ensure that there is a return engagement on this educational resource. The newsletter is largely formed from the most prominent news posts within the news section of the website, focusing primarily on the news items that have a distinct call-to-action. For example, a prominent news story, which was distributed through the April newsletter was encouraging newsletter subscribers to sign up to any one of the new courses available through the courses portal. Currently, there are 3 newsletters that have been sent out, each of which are available in a downloadable PDF format from the EDSA website. Statistics regarding the engagement with this channel can be seen at

Table 4.

As a channel that has been consistently growing throughout the project, Twitter is understood as being a powerful and engaging tool for a wide spectrum of data science professionals, researchers and educators. All news is distributed through this channel to encourage readership of the news stories and subscription to the newsletter. The level of growth of the Twitter channel is aligned with the expectations of this topic, and of the project itself, growing at a rate of 1 new follower a day. Although

this may be considered a slow rate, it is argued that the EDSA project account does not use 'black hat' techniques to boost the followers; instead followers are obtained by producing original, creative and valuable content that the audience wishes to keep abreast of. Thus, it is understood that the followers within Twitter are an engaged and interested audience appropriate for the aims of the engagement and dissemination plans.

The MOOCs ran by partner institutions under the EDSA brand, have since become very successful in terms of disseminating data science education and promoting the project as a whole (Table 7, and Table 8). The growth of online learning provides an opportune platform for the dissemination of EDSA educational material and with the increased spectrum of courses, course material and learning resources now available through the EDSA project's respective course portal and central website more focused and dedicated dissemination has occurred. Furthermore as the resources are made open through videolectures.net, the course portal, and SlideShare encourages a growing demand for data science education (Mikroyannidis *et al,* 2016). As more course material and courses are opened up to a wider audience, and disseminated appropriately and effectively, we see an increase in the amount of enrolments to our respective courses. As such, it has been identified that EDSA must continue to produce open course material to encourage the growth in enrolment. In support of this it is notable that prestige and reputation of an educational institution can support the uptake and engagement with course material (Mukala *et al,* 2015). As such, the partner institutions must continue to leverage their experience and reputation in order to increase engagement, and ultimately, enrolment.

The EDSA project is increasing in scope and reach through many of the channels, especially through the additional channel of our ambassadors – a series of expert data science professionals, researchers, academics and managers whom support the promotion and dissemination of EDSA content and further data science activity in the EU. It is planned that we are to obtain at least an ambassador per constituent country in the EU by the end of the project in order to ensure that we have maximum coverage in all EU member states. The ambassadors will be encouraged to run co-branded events, asked to promote the survey for data collection in WP1, and support the wider growth of the EDSA network through sharing contacts and course materials. Should the ambassador be associated with pedagogy, plans are in place to request their support in developing data science educational material for further EDSA courses. Aligned with the new channel of ambassadors, real-world events have been attended consistently throughout the project. Further to this, we also see an increase in the running of workshops, the distribution of flyers and the organization of events by the respective partners. It is planned throughout the rest of the project that the ambassadors will start to host and organize events on behalf of the EDSA project, to further the dissemination and engagement in their respective country.

6. Conclusion

This report provides an extensive overview of the channels, tools and data being leveraged for the development of the engagement and dissemination activities. In 3.1 summaries of the engagement and dissemination activity is provided, detailing the specific statistics collected based upon the EDSA project website, its related analytics and SEO visibility, videolectures, news and newsletter, and Twitter. Following this, an overview of the audiences for the EDSA-branded courses was presented, showcasing statistics around the engagement with the Process Mining and Introduction to Linked Data and the Semantic Web MOOCs, accompanied by statistics relating to the EDSA SlideShare account. Finally this report details an overview of the real-world events attended, organised and ran by the EDSA project partners, the press coverage that has been conducted both before and since the previous report, and an introduction to the data science ambassadors, their respective experience with data science and their constituent country.

Continually collecting data from all the aforementioned channels will enable the dissemination and engagement activity to be adapted accordingly. With reflection upon D4.2, the activity on LinkedIn has been reduced due to the relatively little engagement which was being received through this particular channel which ultimately could affect the reputation of the project. Instead more effort is being focused upon Twitter, the Newsletter and the news section of the website as they have shown to encourage an extensive amount of engagement and viewing from the appropriate target audience of the project. It is notable that the increase in subscribers for the newsletter shows an increased level of engagement and interest in the project, and of data science as a whole. Remaining flexible and adaptive to the appropriate channels enables the dissemination and engagement WP to ensure that the ideal audience is being targeted and the correct messages are being portrayed (details of which can be found in section 5 of this report). Additionally, SEO has been added as a channel since D4.3, which indeed has increased the engagement and visitation of the website, and ultimately the sign up to the courses.

7. References

Mikroyannidis, A., Domingue, J., Maleshkova, M., Norton, B. and Simperl, E. (2016) Teaching Linked Open Data Using Open Educational Resources, in eds. Mouromtsev, D., d'Aquin, M., Open Data for Education: Linked, Shared, and Reusable Data for Teaching and Learning, pp. 135-152, Springer International Publishing.

Mukala, P., Buijs, J.C.A.M., van der Aalst, W.M.P. (2015) Exploring students' learning behaviour in moocs using process mining techniques. Tech. rep., Eindhoven University of Technology, BPM Center Report BPM-15-10, BPMcenter.org