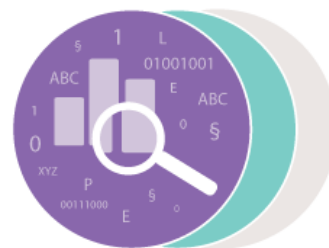




FutureTDM

Explore . Analyse . Improve



REDUCING BARRIERS AND INCREASING UPTAKE OF TEXT AND DATA MINING FOR RESEARCH ENVIRONMENTS USING A COLLABORATIVE KNOWLEDGE AND OPEN INFORMATION APPROACH

Deliverable D7.4

FutureTDM dissemination summary report

Project

Acronym: **FutureTDM**

Title: Reducing Barriers and Increasing Uptake of Text and Data Mining for Research Environments using a Collaborative Knowledge and Open Information Approach

Coordinator: SYNYO GmbH

Reference: 665940

Type: Collaborative project

Programme: HORIZON 2020

Theme: GARRI-3-2014 - Scientific Information in the Digital Age: Text and Data Mining (TDM)

Start: 01. September, 2015

Duration: 24 months

Website: <http://www.futuretdm.eu/>

E-Mail: office@futuretdm.eu

Consortium: **SYNYO GmbH**, Research & Development Department, Austria, (SYNYO)
Stichting LIBER, The Netherlands, (LIBER)
Open Knowledge, UK, (OK/CM)
Radboud University, Centre for Language Studies The Netherlands, (RU)
The British Library, UK, (BL)
Universiteit van Amsterdam, Inst. for Information Law, The Netherlands, (UVA)
Athena Research and Innovation Centre in Information, Communication and Knowledge Technologies, Inst. for Language and Speech Processing, Greece, (ARC)
Ubiquity Press Limited, UK, (UP)
Fundacja Projekt: Polska, Poland, (FPP)

Deliverable

Number:	D7.4
Title:	FutureTDM dissemination summary report
Lead beneficiary:	Open Knowledge International
Work package:	WP7: Disseminate: Project communication, publications, mobilisation and networking
Dissemination level:	Public (PU)
Nature:	Report (RE)
Due date:	31.08.2017
Submission date:	31.08.2017
Authors:	Lieke Ploeger , OK/CM
Contributors:	Burcu Akinci , SYNYO Bernhard Jäger , SYNYO Stefan Kasberger , ContentMine
Review:	Simone Sacchi , LIBER

Acknowledgement: This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 665940.

Disclaimer: The content of this publication is the sole responsibility of the authors, and does not in any way represent the view of the European Commission or its services.

This report by FutureTDM Consortium members can be reused under the CC-BY 4.0 licence (<https://creativecommons.org/licenses/by/4.0/>).

Table of Contents

Executive Summary	7
1. Introduction.....	8
2. Objectives of the Dissemination Strategy	10
3. A Strong Web Presence	11
3.1 Project Website	11
3.2 FutureTDM Platform	12
4. Events and Publications.....	15
4.1 External Events.....	15
4.2 Publications	15
4.3 Presentations.....	16
4.4 Project Poster Presentations	17
4.5 Knowledge Cafés	19
4.6 Workshops.....	20
4.7 Symposium	21
5. Ensuring Engagement.....	24
5.1 FutureTDM Blog	25
5.2 FutureTDM Videos.....	25
5.3 Surveys	27
5.4 Twitter	29
5.5 Branding and Print Material	32
5.6 Newsletter and Mailing List.....	33
6. Press	36
7. Promoting Best Practice and Expertise	38
7.1 Awareness Sheet Collection	38
7.2 Knowledge Base	39
7.3 Tutorials.....	40
8. Annex I: Metrics.....	41
9. Annex II: Selection of Print Material Produced During the Project.....	42

Table of Figures

Figure 1: FutureTDM logo and core message	7
Figure 2: Communication and exploitation plan	9
Figure 3: FutureTDM project website	11
Figure 4: Google analytics of the project website throughout the duration of the project	12
Figure 5: FutureTDM platform homepage	13
Figure 6: Google analytics of the FutureTDM platform throughout the duration of the project.....	13
Figure 7: Country statistics for the FutureTDM platform users	14
Figure 8: Gender statistics for the FutureTDM platform users	14
Figure 9: Discussion scene from Leiden Knowledge Café	19
Figure 10: Knowledge Cafes mapped in Europe.....	20
Figure 11: Word Cloud with Affiliations and Gender Diversity (%) of Attendees of FutureTDM workshop 1 (left) and workshop 2 (right)	21
Figure 12: FutureTDM symposium announcement banner	21
Figure 13: Bernhard Jäger opening the FutureTDM symposium with a keynote speech	22
Figure 14: FutureTDM Symposium Session.....	22
Figure 15: FutureTDM awareness sheets from FutureTDM Symposium	23
Figure 16: Attendance to the keynote speech	23
Figure 17: Tweet during the first FutureTDM workshop on 27 September 2016 by MEP Marietje Schaake.....	24
Figure 18: Screenshot of the video 'Join the FutureTDM community, tell your TDM story!' (June 2016), https://www.youtube.com/watch?v=hwWEsmicew0	26
Figure 19: Screenshot of the video interview 'Peter Knoth - FutureTDM Symposium' (June 2017), https://www.youtube.com/watch?v=Zt61D8Gh1Qk	26
Figure 20: Results of online poll 1	27
Figure 21: Results of online poll 2	27
Figure 22: Inviting participants to provide feedback on TDM barriers, Digital4EU Summit, Brussels, 25 Feb 2016	28
Figure 23: Feedback through stickers on the Knowledge Tree poster, FTDM workshop, Brussels, 29.3.2017.....	28
Figure 24: Number of tweets throughout project lifetime, with peaks for the Knowledge Cafes (April-June 2016), workshops (sept 2016 and March 2017) and symposium (June 2017).	30
Figure 25: @FutureTDM mentions and potential reach	30
Figure 26: Twitter growth, increase in followers leading up to the end of the project (May-July 2017)	31
Figure 27: Twitter followers: interests	31
Figure 28: Twitter followers: gender, age and country.....	32
Figure 29: MEP Lidia Geringer tweeting on @FutureTDM during the second workshop (March 2017)	32
Figure 30: FutureTDM dissemination pack (June 2016).....	33
Figure 31: FutureTDM banner next to Julia Reda at the first FutureTDM workshop, Sep 2015.....	33
Figure 32: Interview with Julia Reda at the European Parliament.....	34
Figure 33: FutureTDM newsletters, edition 3 and 4	35

Figure 34: FOSTER Open Science Facebook post, 13 July 2017, https://www.facebook.com/fosteropenscience/posts/1427939340576462	37
Figure 35: FutureTDM awareness sheets available at the FutureTDM workshop on 27 Sep 2016	39
Figure 36: ContentMine	40
Figure 38: FutureTDM print materials collection	42
Figure 39: FutureTDM Feedback Cards	43
Figure 40: FutureTDM Awareness Sheets Collection	43
Figure 41: FutureTDM Poster	44
Figure 42: FutureTDM poster for Practitioner Guidelines	45
Figure 43: Tree of Digital Knowledge	46

Table of Tables

Table 1: FutureTDM KPIs Metrics	41
---------------------------------------	----

EXECUTIVE SUMMARY

The FutureTDM dissemination report summarises the progress in achieving project Objective #6: INCREASE awareness of TDM to especially attract new target groups and science domains by creating a roadmap, run mobilisation and engagement activities and provide information material and modern TDM visualisations and infographics. It details how the project has raised awareness about FutureTDM and the European TDM landscape and communicated its findings to targeted stakeholder groups and end-users. It is part of Work Package 7 (WP7) Disseminate, which covers project communication, publications, mobilisation and networking.

The overarching goal of the dissemination work was to ensure the broadest possible awareness of the project and its results through clear and informative communications. Secondly, the aim of the communication and dissemination work was to build an invested and growing TDM community to serve as an ever-larger soundboard for the project. Following a brief introduction to these objectives, this report goes into further details on how and to which extent they have been achieved throughout the project's lifetime. Further details on key performance indicators have been included as an annex.



Figure 1: FutureTDM logo and core message

1. INTRODUCTION

The FutureTDM dissemination report summarises the progress in achieving project Objective #6: INCREASE awareness of TDM to especially attract new target groups and science domains by creating a roadmap, run mobilisation and engagement activities and provide information material and modern TDM visualisations and infographics. It details how the project has raised awareness about FutureTDM and communicated its findings to targeted stakeholder groups and end-users. It is part of Work Package 7 (WP7) Disseminate, which covers project communication, publications, mobilisation and networking.

WP7 set the following objectives (in order of production):

- To develop a creative project website to relay timely information about project activities and accomplishments to the public. In addition, social media channels will be exploited to increase project awareness and stimulate stakeholder participation.
- To produce the communication and exploitation plan in the early stages of the project to build sound approaches for networking and outreach.
- To create stimulating dissemination materials (leaflets, newsletters, etc.) to distribute over multiple channels using online and print media to promote the FutureTDM project.
- To organise a Symposium, a culmination of the project's events, to connect key actors and interest groups to promote open dialogue via discussion panels and informal workshops.

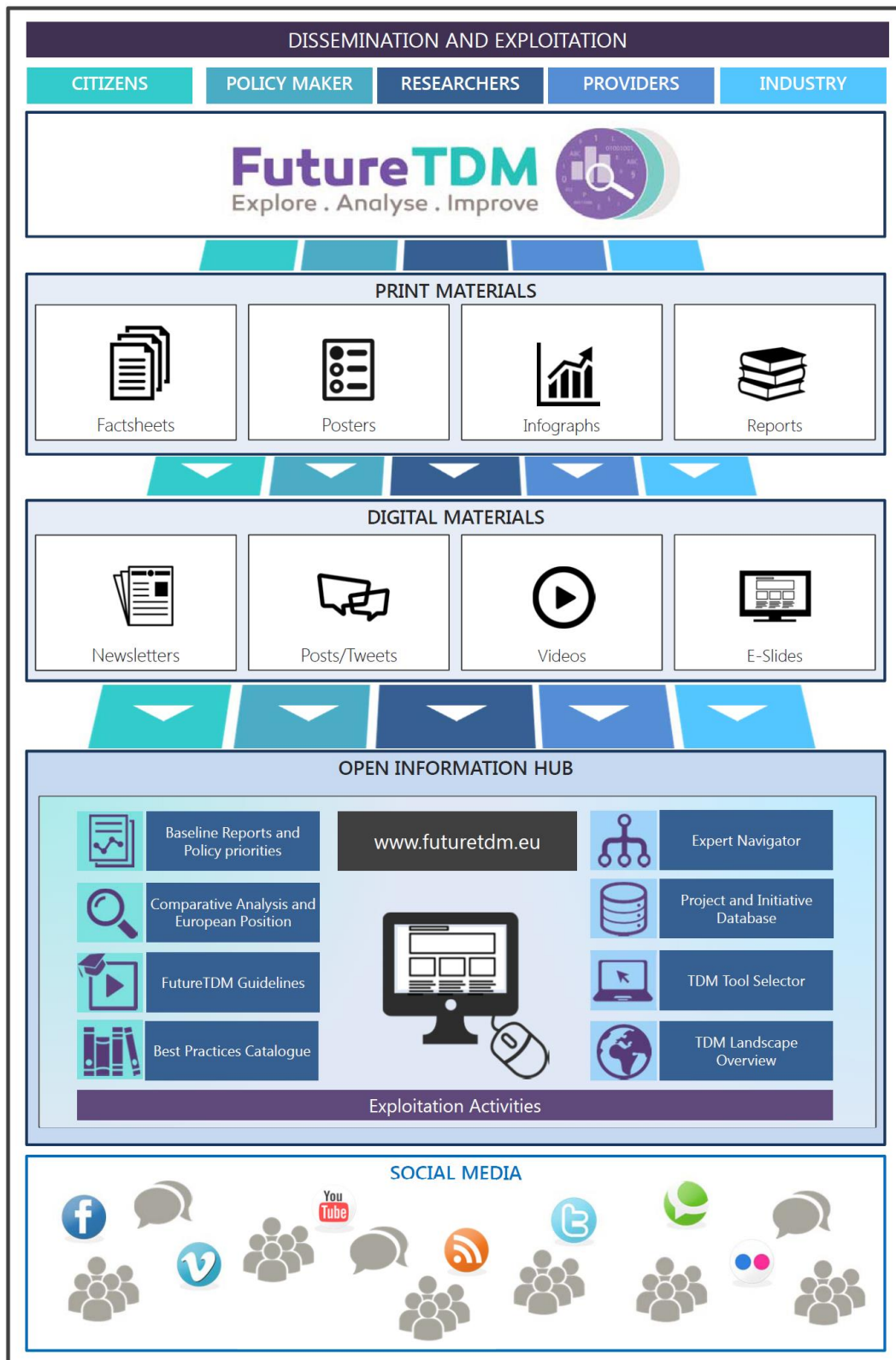


Figure 2: Communication and exploitation plan

2. OBJECTIVES OF THE DISSEMINATION STRATEGY

The overarching goal of the dissemination work was first of all to ensure the broadest possible awareness of the project and its results through clear and informative communications. By providing the necessary tools in the form of physical and virtual materials, the dissemination work package empowered project partners and stakeholders in the field with means to exploit the “multiplier effect”, sharing news and project developments with their own community of professionals, decision makers and end users. Secondly, the aim of the communication and dissemination work was to build an invested and growing TDM community to serve as an ever-larger soundboard for the project.

The FutureTDM project was divided into two phases with a specific focus. In the first instance, the aim was to generate a “buzz” around the project and encourage participation in the engagement activities, to build a sense of belonging and to get feedback from the TDM stakeholder community. With the main focus on gathering stakeholder feedback, the different user groups (researchers, TDM content providers, consumers of TDM, funders, policy makers, service providers, information aggregators and analysts and citizens, as previously identified in D7.2 Communication and Dissemination plan) were offered multiple opportunities to participate in the project’s work at EU events, FutureTDM workshops and knowledge cafés. Communications activities informed stakeholders of the TDM landscape and latest developments as well as the benefits of becoming involved in the project and collaborating with others.

In the second phase of FutureTDM, which started in the second project year, the focus shifted to spreading knowledge gained throughout the project (best practice achievements, recommendations and outcomes) through the FutureTDM platform (www.futuretdm.eu), to help stakeholders improve their own work processes. Stakeholders were also given the opportunity to share written feedback and expertise with the project via social media, blogs and surveys. Finally, they were invited to join two dedicated FutureTDM workshops at the European Parliament in Brussels, as well as the final Symposium during the International Data Science conference.

Throughout the project, our dissemination efforts sought to relate to the current political situation around text and data mining in Europe, especially around the proposed EC directive on copyright in the Digital Single Market (September 2016, see http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=17200). Apart from organising FutureTDM workshops directly at the EU parliament at strategic times (which lead to two press releases by MEPs¹), the work package planned specific blogs dedicated to current topics (see for example the blog ‘[CULT’s opinion on the TDM exception](#)’ by Marco Caspers, 17 February 2017) or featured such news prominently in the newsletter (for example featuring an interview with Julia Reda MEP (Digital Agenda Intergroup) in the September 2016 edition).

¹ Catherine Stihler, ‘Press release: FutureTDM Event at Scotland House’, 13 October 2016, http://www.cstihlermep.com/Press_Releases/id912.php and Marietje Schaake, ‘Future of Text and Data Mining’, 13 October 2016, <https://www.marietjeschaake.eu/en/future-of-text-and-data-mining>
© 2017 FutureTDM | Horizon 2020 | GARRI-3-2014 | 665940

3. A STRONG WEB PRESENCE

FutureTDM set up two domains for the online dissemination of project results: a project website (<http://project.futuretdm.eu>) with general project information, which was most active during the initial project phase, and the platform www.futuretdm.eu, which was gradually built up throughout the first months, launching as the main open information hub as of June 2016, and becoming the main point of dissemination in the second phase.

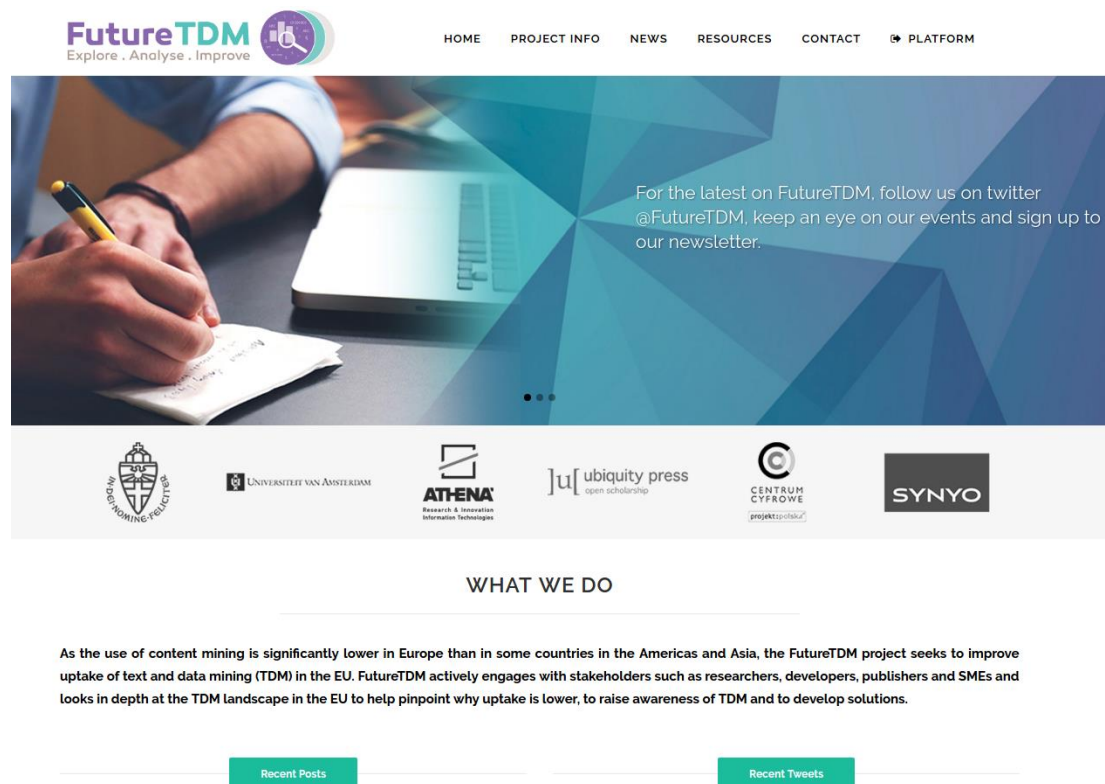


Figure 3: FutureTDM project website

3.1 Project Website

Until June 2016 (during the first phase of the project), the FutureTDM project website <http://project.futuretdm.eu> functioned as the main online communication space. Stakeholders could find out more about the project itself, as well as FutureTDM activities and resources (such as publications, flyers and awareness sheets). The website offered functionality such as commenting on blogs, sharing items across social media platforms and downloading specific items such as the project logo and flyers.

In the “Blog” section, consortium partners and invited stakeholders wrote about TDM as well as topics related to the scope of the project in general. These blogs were crossposted onto the FutureTDM platform (see next section) to ensure maximum visibility. As of June 2016, blogging was solely done through the platform: an [announcement](#) was posted on the project website.

Website statistics in the figure below show that the project website was well-visited from its’ start in early 2016. Clear spikes are visible for the Knowledge Cafes, which took place in the first half of 2016.

The largest peak after that is in end of May / early June 2016, when the first batch of awareness sheets was added and promoted. It is also clearly visible that traffic decreased after June 2016, when the FutureTDM platform became the main focus.

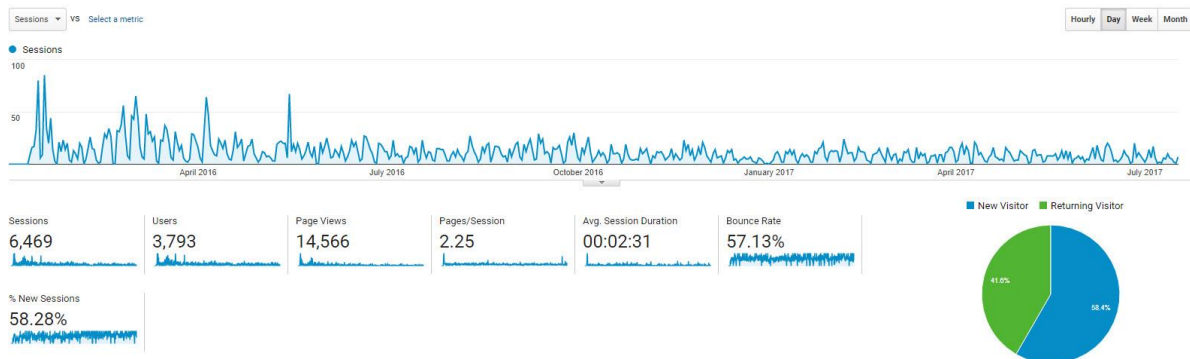


Figure 4: Google analytics of the project website throughout the duration of the project

Further information on the project website is available from D7.1: Project website, social media channels and communication activities (M2, October 2015).

3.2 FutureTDM Platform

The FutureTDM platform (<http://www.futuretdm.eu/>) was originally scheduled to be launched in M14 (October 2016), but since development progressed well it was already active since May 2016, with the project blog officially moving to the platform as of June 2016. The FutureTDM Platform consists of an Open Information Hub and a collaborative Knowledge Base. It aims to be a one-stop shop for evidence-based information on Text and Data Mining (TDM) developments for all stakeholders in Europe.

The Open Information Hub provides different sections such as Blog, Events and Knowledge Library. The blogposts are categorised, and authors (both project partners and invited stakeholders) write about FutureTDM outcomes and express opinions on important TDM developments which are relevant for the scope of FutureTDM. The Events page lists all FutureTDM Knowledge Cafés and other external events related to TDM. The Knowledge Library is a collection of awareness sheets, expert reports and other dissemination materials that have been created throughout the project.

The Knowledge Base displays the collection of resources on Text and Data Mining in a user-friendly structure, providing a coherent and up-to-date view on the European TDM landscape. The beta version was released in April 2017 and promoted through the newsletter, with the final version completed in June 2017 with the final project results incorporated.

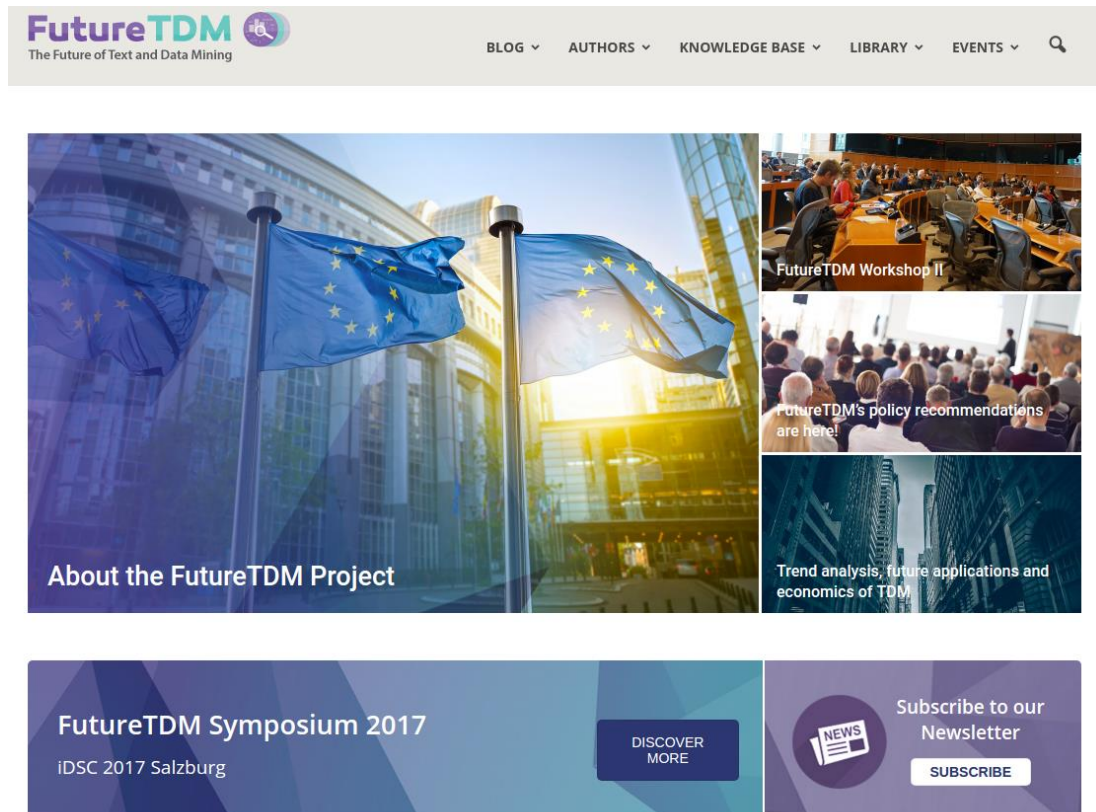


Figure 5: FutureTDM platform homepage

The FutureTDM platform has been visited by over 5,000 users since the beginning of the project (see figure below). Clear peaks are visible around the time of the first Knowledge Cafes (February-April 2016), the first workshop (27 September 2016) as well as for the increased dissemination throughout the second phase of the project (January - July 2017), when more final project results became available to be shared widely.

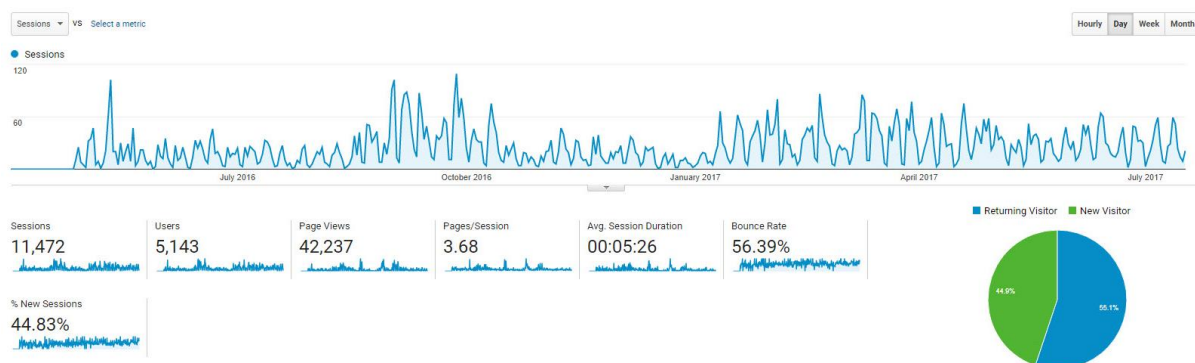


Figure 6: Google analytics of the FutureTDM platform throughout the duration of the project

Looking at the user statistics, the gender is divided fairly equally. Followers mainly come from the EU and the United States, but there is a good spread of users from other parts of the world.

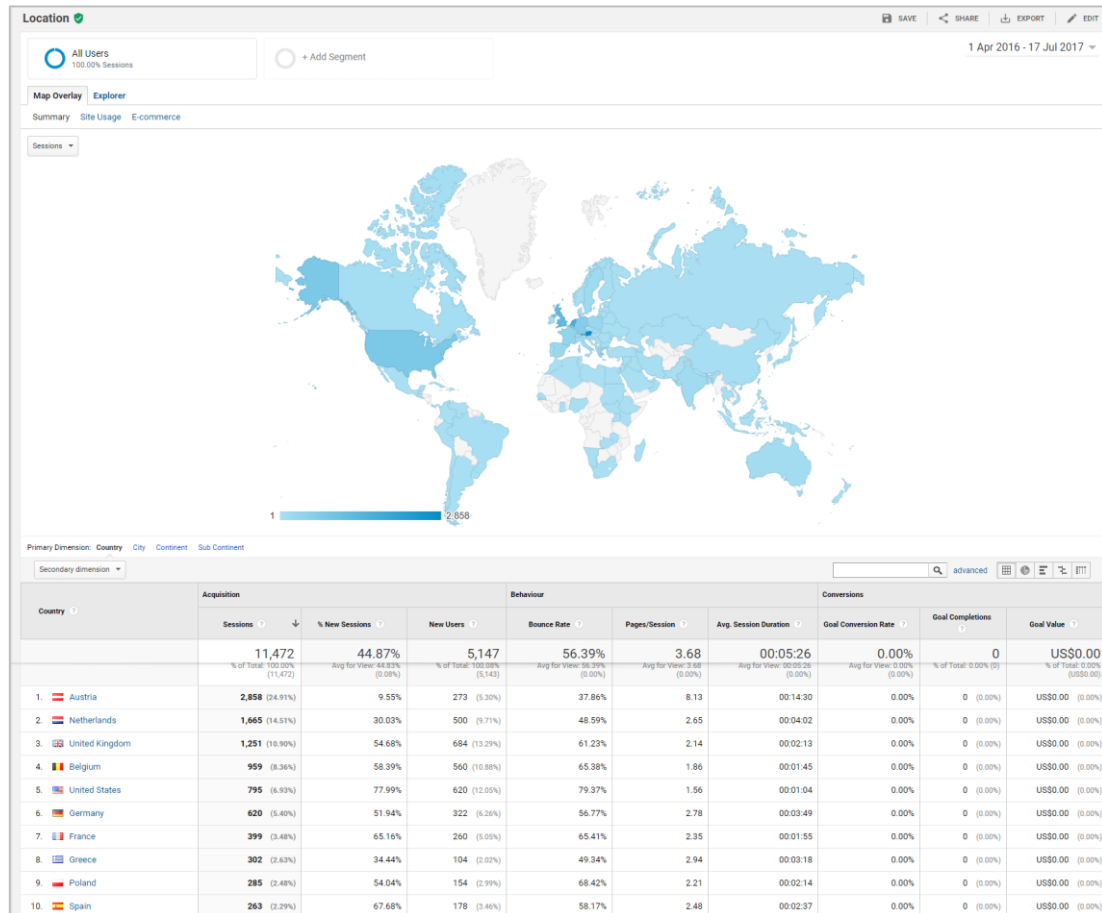


Figure 7: Country statistics for the FutureTDM platform users

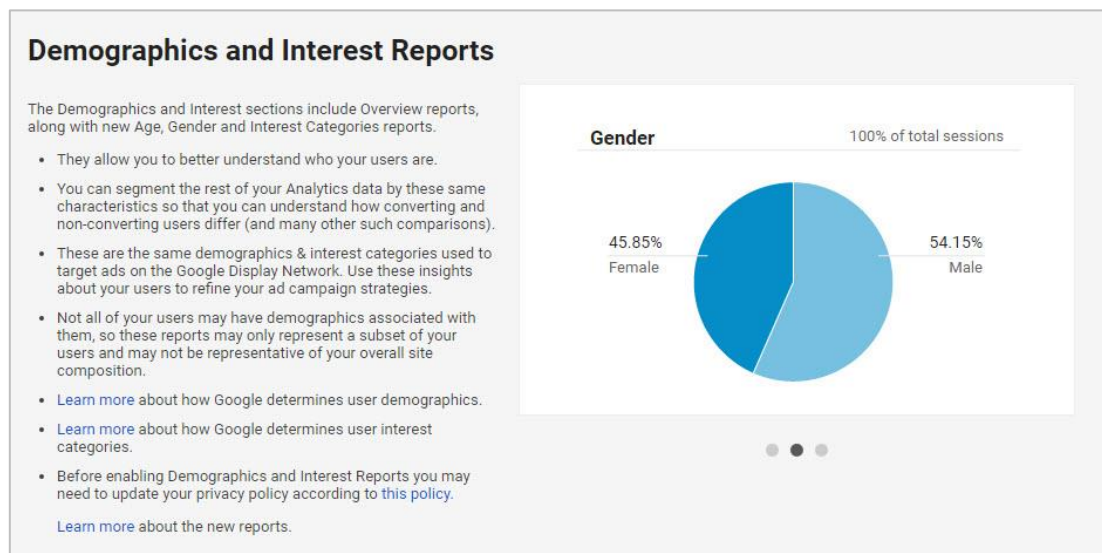


Figure 8: Gender statistics for the FutureTDM platform users

More information on the platform is available from D6.1 Web-based FutureTDM Collaborative Knowledge Hub and on the Knowledge Base from D6.2 Stakeholder mapping and TDM expert navigator.

4. EVENTS AND PUBLICATIONS

Since one of the most important aspects of the FutureTDM project was gaining feedback from stakeholders, the consortium partners made sure to attend many external conferences, events, workshops and meetings, also those hosted by other European Commission funded projects in the field, to talk about the project work. In addition, FutureTDM organised a programme of knowledge cafés, workshops and the FutureTDM Symposium. In this section an overview is given of participation at external events, publications, (poster) presentations and workshops.

4.1 External Events

Participation at external events happened throughout the project: first with a focus on gathering stakeholder feedback (from researchers, content providers, funders, policy makers, service providers and citizens) to incorporate into the project reports, and secondly as a means to spread awareness of the challenges and possible solutions for improving the uptake of text and data mining in the EU in the future. A Slideshare account for sharing presentations given at events was created at www.slideshare.net/FutureTDM. Relevant publications and presentations given by project partners were added to the project website - Publications page (<http://project.futuretdm.eu/publications>). The project poster was uploaded to Zenodo (<https://zenodo.org/record/56653>): the aim is to upload further relevant content there in August 2017 to ensure long-term preservation of the projects' outputs and its' persistent access for future reuse and citation (all content in the Zenodo platform is assigned persistent DOIs).

4.2 Publications

1. M. Eskevich, S. Piperidis, K. Pouli, M. Gavriilidou, D. Galanis, A. v.d. Bosch. Text and Data Mining: Past, present, and future. Data Science Journal, Ubiquity Press (resubmitted with corrections required by reviewers. June 2017).
2. M. Eskevich. Stakeholders in academic publishing: text and data mining perspective and potential. International Conference on Electronic Publishing (ELPUB), Goettingen, Germany, 7-9 June 2016 DOI: 10.3233/978-1-61499-649-1-51, <http://ebooks.iospress.nl/publication/42896>
3. M. Caspers & L. Guibault, L. (2016). A right to 'read' for machines: Assessing a black-box analysis exception for data mining. Proceedings of the Association for Information Science and Technology, 53(1), 1-5. <http://onlinelibrary.wiley.com/doi/10.1002/pra2.2016.14505301017/full>
4. FutureTDM poster: M. Eskevich, A. van den Bosch, F. van den Boom, H. Frew, B. Akinci, A. Bertone, I. Luca, M. Caspers. FutureTDM: Improving Uptake of Text and Data Mining in the EU. European Data Forum, Eindhoven, the Netherlands, 29-30 June, 2016, <https://zenodo.org/record/56653> and <http://project.futuretdm.eu/wp-content/uploads/2016/06/Poster-FutureTDM.pdf>
5. H. Frew, FutureTDM is all About Stakeholder Feedback, Signum, Journal of the Finnish Research Library Association, June 2016
6. C. Handke, L. Guibault & J.-J. Vallbé, Is Europe Falling Behind? Copyright's Impact on Data Mining in Academic Research, in Birgit Schmidt, Milena Dobрева (eds.), New Avenues for

Electronic Publishing in the Age of Infinite Collections and Citizen Science: Scale, Openness and Trust – Proceedings of the 19th International Conference on Electronic Publishing 2015, pp. 120-130, [link](#).

4.3 Presentations

1. Peter Murry-Rust (ContentMine): [High throughput mining of the scholarly literature: a new research Tool](#). MTO, University of Tilburg, NL, 2016-06-07 (Specialists in Scholarly Fraud detection)
2. Peter Murry-Rust (ContentMine): [Amanuens.is ContentMine + Hypothesis annotate the scientific literature](#). 100,000+ per day. IAnnotate!, W3C Berlin, DE, 2016-05-18 (W3C meeting on document annotation)
3. Peter Murry-Rust (ContentMine): [Open Content + Open Programs and Mining](#). MIOSS 2016, EBI, UK, 2016-05-17 (Industry/Academic meeting, European Bioinformatics Institute)
4. Peter Murry-Rust (ContentMine): [Automatic Extraction of Knowledge from the Literature](#). CILIP ISG, Cambridge, UK, 2016-05-11 (Chartered institute of Library and Information Professionals)
5. Peter Murry-Rust (ContentMine): [Automated Extraction of Knowledge from Biomedical Literature](#). Cochrane UK & Ireland Symposium 2016, Birmingham, UK, 2016-03-16 (Systematic reviews of clinical trials)
6. LIBER to the SPARC meeting on Openness in Education & Research, [The Open Science Agenda in Europe: Policy convergence & diversity of approaches](#), Texas, (March 7 2016)
7. M. Caspers, L. Guibault F van den Boom, [Thinking of copyright works as black boxes. A solution for TDM and machine learning activities](#), Conference TILting Perspectives 2017: 'Regulating a connected world', Tilburg University, Tilburg, the Netherlands (19 May 2017)
8. L. Guibault IPR, Technology Transfer & Open Science - Challenges and opportunities, Keynote speech, European Commission (DG Research and Innovation), Brussels (9 Mar 2017)
9. L. Guibault, 'Text & Data Mining: Barriers, Paths and Passable Roads', Symposium Designing and shaping open science, Amsterdam (5 April 2016)
10. L. Guibault, 'Mining Permits: Measuring copyright's impact on text and data Mining in academic research', workshop on Users' Rights in the Digital Economy, American University, Washington D.C. (18 March 2016)
11. L. Guibault, 'Is Europe Falling Behind in text and data mining?', Symposium of Science Europe (Brussels, 27 October 2015)
12. L. Guibault & J.-J. Vallbé, 'Is Europe Falling Behind? Copyright's Impact on Data Mining in Academic Research' ELPuB Conference – Valletta, Malta (1-2 September 2015)
13. M. Caspers, L. Guibault, F van den Boom (2016). Panel: A right to 'read' for machines: Assessing a black-box analysis exception for data mining. <https://www.asist.org/events/annual-meeting-2016/program/a-right-to-read-for-machines-assessing-a-black-box-analysis-exception-for-data-mining/>
14. F. van den Boom, FutureTDM: Improving Uptake of Text and Data Mining in the EU, Competitive Advantage in the Digital Economy Forum (CADE 2016) https://www2.warwick.ac.uk/fac/sci/wmg/research/business_transformation/ssg/ssgabout/sswmgactivities/cade2017/

15. F. van den Boom, FutureTDM: Improving Uptake of Text and Data Mining in the EU, unconference session at OpenGOv, Taiwan <http://g0v.asia/tw/>
16. F. van den Boom, FutureTDM: Improving Uptake of Text and Data Mining in the EU, unconference session at OpenCon2016, 12-14 November, Washington DC, <http://www.futuretdm.eu/blog/community-events/opencon-conference-2016/>
17. F. van den Boom, FutureTDM session 'TDM: Legal Aspects, Skills and Community Input' by Freyja van den Boom at Creative Commons 2017 Global Summit: Sharing and the Commons: What's Next. Toronto, <https://creativecommons-globalsummit2017t.sched.com/event/AF04/tdm-legal-aspects-skills-and-community-input?iframe=yes&w=100%&sidebar=yes&bg=no#>
18. F. van den Boom, Roundtable: 'Research and the right to mine (personal) data?' at CPDP 2017 conference, 25-27 January 2017, Brussels, <http://www.cdpconferences.org/27012017/MA2.html>
19. F. van den Boom, FutureTDM: Improving Uptake of Text and Data Mining in the EU, 2nd EDISON Data Science Champions Conference: Madrid <http://edison-project.eu/second-edison-champions-conference>
20. F. van den Boom: Joint session Teaching Text and Data Mining (TDM) on Education and Skill Development, RDA Barcelona 2017, <https://www.rd-alliance.org/group/teaching-tdm-education-and-skill-development-wg/case-statement/teaching-tdm-education-and>
21. F. van den Boom, FutureTDM: Best practices; Text and Data Mining in the EU, OR2017, Brisbane, <https://or2017.net/>
22. M. Eskevich, FutureTDM: Improving Uptake of Text and Data Mining in the EU, LT-Accelerate, 21-22 November 2016, Brussels. http://www.lt-innovate.org/sites/default/files/lt_accelerate_files/09.30%20Maria%20Eskevich.pdf
23. K. McNeice, invited plenary: Text and Data Mining: The Changing Nature of Research; panel discussion: Text & Data Mining Policy; Cambridge, UK, Text and Data Mining Symposium, 12 July 2017, <http://osc.cam.ac.uk/text-and-data-mining-symposium>
Planned in July - August 2017
24. F. van den Boom, August: presentation and poster on FutureTDM at the summer school course 'Text Mining the Great Unread: Data-intensive Methods and Digital Tools for Analysis of Texts in the Humanities and Social Sciences', Aarhus University, 24 July - 4 August 2017, <http://www.au.dk/en/summeruniversity/courses/text-mining-the-great-unread-data-intensive-methods-and-digital-tools-for-analysis-of-texts-in-the-humanities-and-social-sciences/>
25. LIBER, Session 199 Text and Data Mining (TDM) Workshop for Data Discovery and Analytics - Big Data Special Interest Group (SIG) at the IFLA World Library and Information Congress, 19–25 August 2017, Wrocław, Poland, <https://2017.ifla.org/>

4.4 Project Poster Presentations

1. FutureTDM poster: M. Eskevich, A. van den Bosch, F. van den Boom, H. Frew, B. Akinci, A. Bertone, I. Luca, M. Caspers. FutureTDM: Improve the Uptake of Text and Data Mining in the EU. Global TechMining Conference, co-allocated with International conference on science and technology indicators, Valencia, Spain, 13 September, 2016.

2. FutureTDM stand at Digital4EU Summit, Brussels, 25 Feb 2016, <https://ec.europa.eu/digital-single-market/en/digital4eu-2016-programme>
3. FutureTDM poster presented at LREC Conference, Portoroz, Slovenia (24 May 2016) <http://www.lrec-conf.org/>
4. FutureTDM poster presented at BDVA Summit 2016 (Valencia, December 2016) <http://www.bdva.eu/?q=node/412>
5. FutureTDM poster presented at DI4R Digital Infrastructures for Research event (Krakow, Sept 2016) <https://www.digitalinfrastructures.eu/>
6. FutureTDM poster presented at RDA Plenary 8 (Denver, Sept 2016) <https://www.rd-alliance.org/plenaries/rda-eighth-plenary-meeting-denver-co>
7. FutureTDM poster presented at Oreilly security conference (New York, November 2016) <https://conferences.oreilly.com/security/sec-ny-2016>
8. FutureTDM poster presented at APE2017, 17-18 January 2017, Berlin. <http://www.futuretdm.eu/blog/community-events/futuretdm-reports-on-ape-2017/>
9. FutureTDM poster presented at Sage Bionetworks, The 2017 Assembly, Mapping Open Research Ecosystems, <http://sagebase.org/events/congress-2017/>
10. FutureTDM poster presented at International Conference on Electronic Publishing (ELPUB), 6 June 2017, Limassol, Cyprus <http://www.elpub.net>
11. FutureTDM poster presented at Open Repositories Conference, 26-30 June 2017 at the Hilton Brisbane, <https://or2017.net/>
12. FutureTDM poster presented at International Digital Curation Conference 2017, Edinburg, <http://www.dcc.ac.uk/events/international-digital-curation-conference-idcc>
13. K. McNeice, B. Akinci, poster on FutureTDM poster stakeholder guidelines presented at LIBER 2017 annual conference, 5-7 July, Patras, Greece, <http://liber2017.lis.upatras.gr>

In addition to these activities, project partners also organised a number of workshops at relevant events, including:

1. Peter Murry-Rust (ContentMine): Content Mining of Science in Europe. OpenForum, Brussels, BE 2016-04-14 (Workshop, includes participation by Julia Reda MEP) <http://www.openforumeurope.org/>
2. L. Guibault, 'Text & Data Mining: Barriers, Paths and Passable Roads', Legal Workshop LREC Conference, Portoroz, Slovenia (24 May 2016)
3. F. van den Boom, Prof M. van Eechoud, K. Rogers, D. Krebs,, Workshop 'Mine the Government' at OGP Global Summit 2016, 7-9 December 2016, Paris. <https://en.ogpsummit.org/osem/conference/ogp-summit/program/proposal/22>
4. F. van den Boom and S. Kasberger, FutureTDM: Workshop on best practices for text and data mining, International Conference on Electronic Publishing (ELPUB), 6 June 2017, Limassol, Cyprus. <http://aims.fao.org/activity/events/elpub-2017-international-conference-electronic-publishing> and <https://github.com/ContentMine/FutureTDM/tree/master/workshops/elpub-limassol>
5. FutureTDM Open Repositories Pre-Conference workshop Mine your repository. together with Petr Knoth from OpenMinTed project. June 2017 <https://or2017.net/pre-conference-workshops/>

6. LIBER, Joint FutureTDM - OpenMinted workshop '[So You Want to Do Text and Data Mining? We'll Tell You How](#)', LIBER 2017 Annual conference, 5-7 July 2017, Patras.

FutureTDM also focused on collaborating with stakeholders in other ways: the project provided datasets to the [EU Hackathon](#), 15-16 November 2016, Brussels, and was present at the event to interview participants. Together with the OpenMinted project, FutureTDM organised a roundtable meeting at the [OpenMinted plenary meeting](#) on 7 February 2017 in Lausanne. Keira McNeice and Ben White held a talk for British Library staff about the FutureTDM project on 17 January, 2017 and led a discussion on TDM for startups at the Shack15 coworking space in London, 6 June 2017, <http://shack15.com/>.

Finally, following the project end, participation from FutureTDM partners LIBER and ARC is planned in the upcoming Open Science Fair event in early September (<http://openup-h2020.eu/2017/05/04/open-science-fair-2017/>), organised by EU-projects OpenMinted, OpenAIRE, Foster and OpenUP. FutureTDM results will be discussed and presented during an Open Science Cafe, for which planning is currently in progress.

4.5 Knowledge Cafés

Throughout the first half of 2016, FutureTDM organised a series of Knowledge Cafés across Europe, where stakeholders from across the TDM community were invited to meet face to face in an informal setting and share their ideas on what factors impact TDM uptake in Europe. Knowledge Cafés often took place alongside significant conferences and events, to meet with relevant stakeholders attending those. This was an important strategy to gather information that was fed back into the project where, after expert analysis of the TDM landscape, policy recommendations were made.



Figure 9: Discussion scene from Leiden Knowledge Café

A range of channels and methods were used to both attract relevant stakeholders to these events and to ensure sufficient feedback. Knowledge Cafés were announced through the FutureTDM platform,

the newsletter and the project Twitter account. A special promotion video titled 'Join the FutureTDM community, tell your TDM story!' invited people to give their feedback either online or through attending a Knowledge Cafe. Blogs reporting on the events, with images and video interviews with participants were added to the platform.

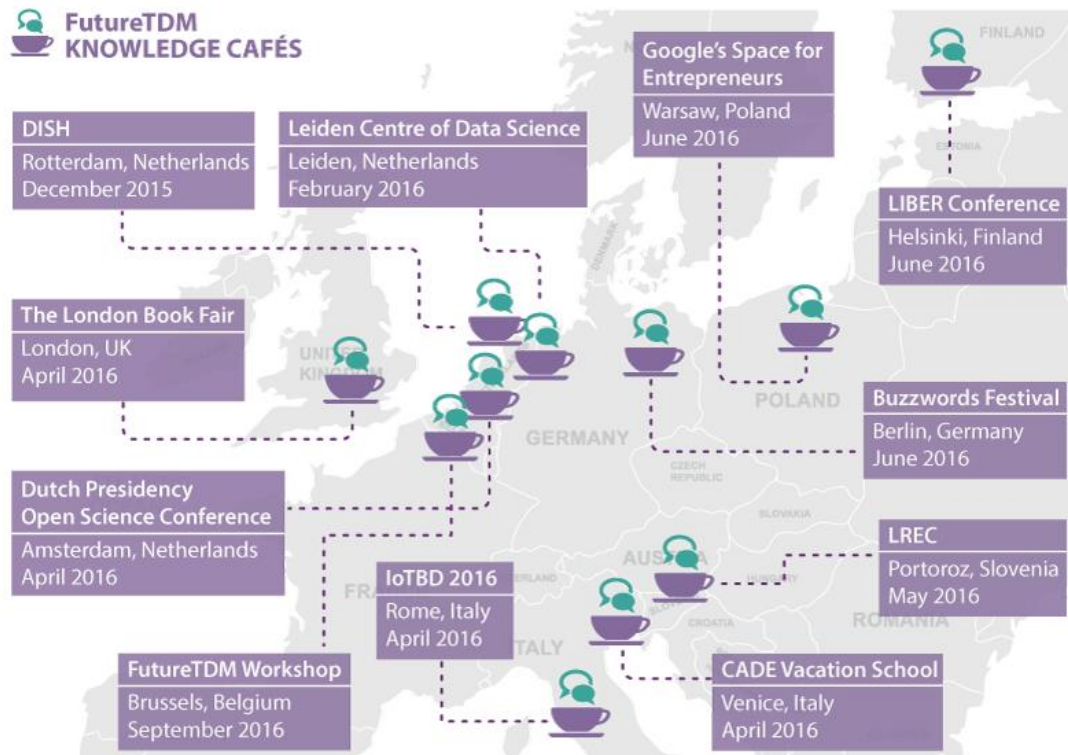


Figure 10: Knowledge Cafes mapped in Europe

More information is available from D2.3 Report on stakeholder mobilisation and perceptions.

4.6 Workshops

In the second project year, two FutureTDM workshops took place at the European Parliament in Brussels. The purpose of these thematic, multi stakeholder workshops was to provide the opportunity for stakeholders to discuss drivers and barriers to the take up of TDM and to feed into the core deliverables of the FutureTDM project. They were planned at strategic times during the ongoing work on copyright reform by the European Commission, and saw participation of several MEPs (as well as two press releases).

FutureTDM secured the European Parliament Digital Agenda Intergroup sponsorship of the first workshop to ensure that our message reached the EU policy audience and that it was done in a balanced and objective manner. MEPs from Intergroup chaired both of the workshop panels and gave the closing summary keynote. Both workshops covered all relevant stakeholder communities and were balanced with regard to gender diversity, as is shown in the figures below.

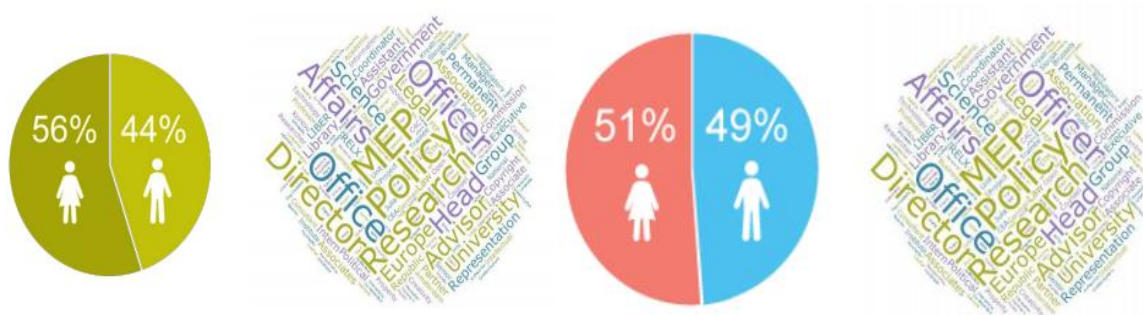


Figure 11: Word Cloud with Affiliations and Gender Diversity (%) of Attendees of FutureTDM workshop 1 (left) and workshop 2 (right)

More information on both workshops is available from D2.4 Workshop summary report 1 and D2.5 Workshop summary report 2.

4.7 Symposium

The 1st International Data Science Conference (iDSC 2017), which was organised on 12-13 June 2017 in Salzburg, was chosen as the most appropriate event to co-organise the FutureTDM Symposium. The audience targeted by the iDSC conference was the broad community of researchers and industry practitioners as well as other practitioners and stakeholders, and was therefore seen as the most suitable event to share the project outcomes at. In the FutureTDM conference track on 13 June 2017, consortium partners shared the policy recommendations and sector specific guidelines, the project's tools and services, practical tutorials, demos, online help and how-to guides to increase TDM uptake.



Figure 12: FutureTDM symposium announcement banner

In different sessions, the consortium presented the legal, technical and educational/expertise needs related to TDM, focusing on the main barriers hindering the uptake of TDM and the respective recommendations provided by FutureTDM.

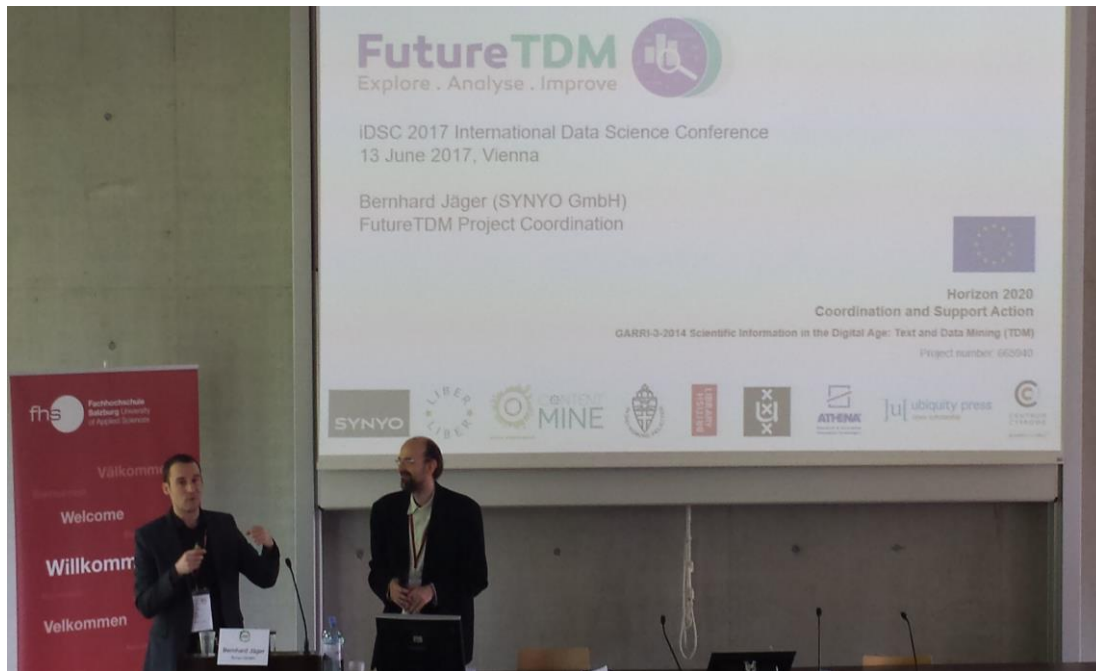


Figure 13: Bernhard Jäger opening the FutureTDM symposium with a keynote speech

Each session began with an introductory presentation by FutureTDM, followed by a panel discussion with external experts contributing their feedback, insights and their own activities on the issue, and finally an open discussion with the audience. During lunch, a demo session took place, during which participants could find out more about six projects and organisations working within the field of text and data mining.



Figure 14: FutureTDM Symposium Session

Dissemination material, such as flyers, stickers and awareness sheets was distributed among participants, and several participants were interviewed for the FutureTDM Youtube channel. The event succeeded both in making the outcomes of the project public and in bringing together experts from various backgrounds and parts of Europe. “Thank you” emails were sent to all the invitees participating in the Symposium and the received feedback was positive. A symposium summary was published on the [blog](#) and featured as the main subject in the [newsletter of June 2017](#).



Figure 15: FutureTDM awareness sheets from FutureTDM Symposium



Figure 16: Attendance to the keynote speech

More information on the symposium is available at <http://www.futuretdm.eu/blog/futuretdm-symposium-2017/>.

5. ENSURING ENGAGEMENT

The FutureTDM project focused heavily on stakeholder involvement: both in person during events, but also online through various tools and channels. Four aspects of the communications strategy were designed for this purpose:

- A clear and accessible video that explains the project and invites feedback (shown on the website, the platform, at events and promoted on Twitter)
- A blog section to which both stakeholders and project partners can contribute
- Online surveys for stakeholders to provide their opinion on the project
- A social media strategy focused on engagement rather than dissemination only

The following chapter outlines the achievements in each of these categories, as well as additional efforts that were undertaken to solicit stakeholder feedback. In addition to the online surveys that were foreseen at the start of the project, a number of different strategies were developed to obtain feedback on TDM barriers in the different sectors in other ways, for example through feedback cards handed out during events, stickers in the shape of leaves that could be posted onto the Knowledge Tree poster (see Annex 2) and video interviews. The experience gained during the project was that such in-person, direct feedback gave a larger response than solely online surveys.

During the engagement efforts, special attention was devoted to ensuring that participation (both from attendees and speakers) was balanced in terms of gender diversity, involved stakeholders from the different target groups in sufficient numbers, as well as engaged participants from across Europe. Additional research was started up to investigate gender and diversity issue more widely for the TDM field, but this turned out to be a too large field of study for the scope of the current project.



Figure 17: Tweet during the first FutureTDM workshop on 27 September 2016 by MEP Marietje Schaake

As will be further outlined below, the dissemination strategy was successful in reaching a wide variety of stakeholders, obtaining valuable feedback feeding into the deliverables, and reaching the targets set out at the beginning of the project (also see Annex 1). All dissemination outputs mentioned throughout this chapter have been uploaded during the project lifetime to the FutureTDM platform - Knowledge Library - Dissemination Outputs folder: <http://www.futuretdm.eu/knowledge-library>, as well as promoted through Twitter and our blog.

5.1 FutureTDM Blog

The FutureTDM blog was started on the project website in January 2016. The blog section was featured under News, but also promoted on the landing page of the website. Until June 2016, all blogs were cross-posted on the FutureTDM platform: in the final project website blog on 2 June 2016 the switch to the FutureTDM platform was announced, and the project website blog discontinued. On the platform, the blog had its own menu tab and was prominently featured on the landing page too. Each blog was categorised in one of the following categories:

- Community & Events
- Education & Skills
- Legal & Policies
- Research & Insights
- Stories & Cases
- Technologies & Tools

Project members received an Author account to be able to add blogs directly via Wordpress, and guest authors were invited to contribute as well, increasingly so during the second phase of the project. WP7 ensured the regular contribution of diverse material, including information on FutureTDM developments and events, opinion pieces that ask for feedback and create a discussion among the TDM community and blogs that increase awareness around the benefits of text and data mining and the current barriers that hinder further uptake in Europe.

As of July 2017, a total of 59 posts have been published on the platform and 13 posts on the project website. Since the end of 2016 there is an increased focus on soliciting guest blog posts, especially from the OpenMinted project (which were often crossposted on both project platforms), but also from other stakeholders in the field. On average, blogposts received between 300-500 views, with the most popular blogs being '[On the role of a university library in the TDM landscape](#)' (7920 views), '[Interview: Julia Reda](#)' (2329 views), '[ContentMine: A Practical Exploration of Mining the Scientific Literature](#)' (2309 views), '[The tension between TDM and data protection](#)' (2176 views) and '[FutureTDM's policy recommendations are here!](#)' (1781 views).

5.2 FutureTDM Videos

A Youtube channel for sharing videos of the project and events has been set up at <https://www.youtube.com/channel/UC4ebz-UMCvznjPIMvF1QV1A>. There are currently 28 videos available, with over 2000 views combined. They are a combination of specially produced animated videos and stakeholder interviews.

The first video was produced as a visualisation of the DoA, aimed at other projects wishing to gain an insight into FutureTDM operations, and second video encouraging all TDM interested stakeholders to get involved in the FutureTDM community and provide their feedback (key message 'Help us to help you, join the FutureTDM community to improve TDM uptake in the EU'). Both videos were added to either the project homepage (video 1) and the FutureTDM platform (video 2), as well as promoted via Twitter and played during FutureTDM event to ensure maximum impact and reach.



Figure 18: Screenshot of the video 'Join the FutureTDM community, tell your TDM story!' (June 2016), <https://www.youtube.com/watch?v=hwWEsmicew0>

In addition, video interviews were conducted at various events, such as the Knowledge Cafes and the final symposium. These interviews with stakeholders across different user groups focused on explaining the different barriers hindering wider uptake of text and data mining in Europe, as well as possible solutions from the stakeholder's unique perspective. Finally, a special video was produced for the Zika tutorial by ContentMine (also see 'Tutorials'), providing a walk-through of how to use ContentMine to find knowledge on the Zika outbreak from scientific articles.



Figure 19: Screenshot of the video interview 'Peter Knoth - FutureTDM Symposium' (June 2017), <https://www.youtube.com/watch?v=Zt61D8Gh1Qk>

5.3 Surveys

According to the original DoA, online surveys with a minimum of 50 responses were foreseen for soliciting stakeholder feedback. The project delivered two online polls, through the frontpage of the platform (as well as promoted via Twitter and the newsletter): the first quick poll asked 'What represents the biggest barrier to TDM uptake in the EU?'. Users were also asked to identify themselves as either Research, Education, Content Provider, Policy, Technical, Innovators, Economic, Start-ups, International Outlook or Other. The second survey asked 'How can we support the growth of TDM in your sector?'. Survey 1 received a total of 31 responses, and survey 2 20 responses.

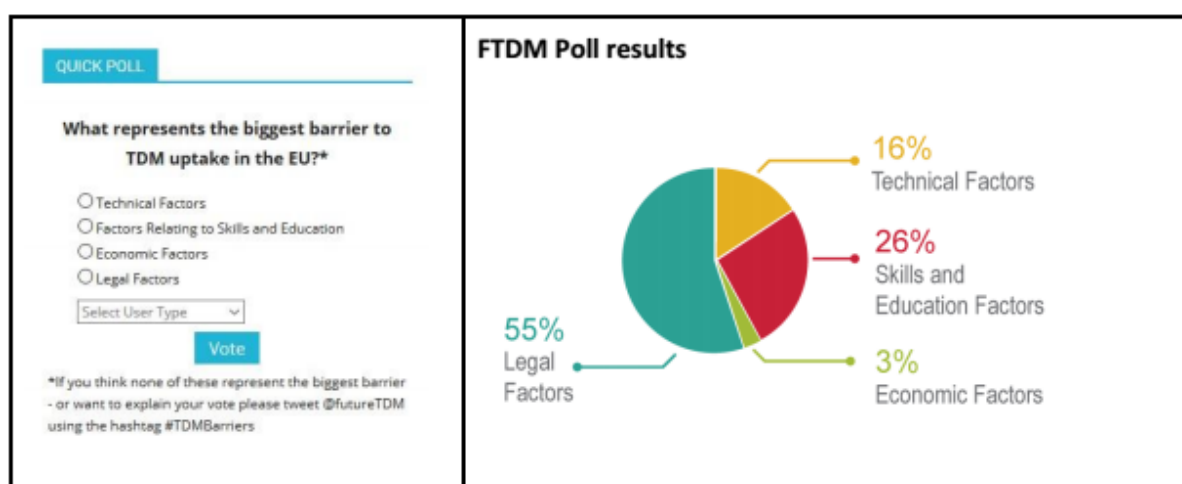


Figure 20: Results of online poll 1



Figure 21: Results of online poll 2

Since the number of responses received was not as high as envisioned at the onset of the project, a number of different strategies were developed to obtain feedback on TDM barriers in the different sectors in other ways, for example through feedback cards handed out during events, by asking people to put their challenge on a piece of paper in one of 4 glasses depicting the different areas (see image below), through handing out stickers in the shape of leaves that could be posted onto the Knowledge

© 2017 FutureTDM | Horizon 2020 | GARRI-3-2014 | 665940

Tree poster (see Annex 2), and by taping video interviews. The experience gained was that such in-person, direct feedback gave a larger response than solely online surveys.

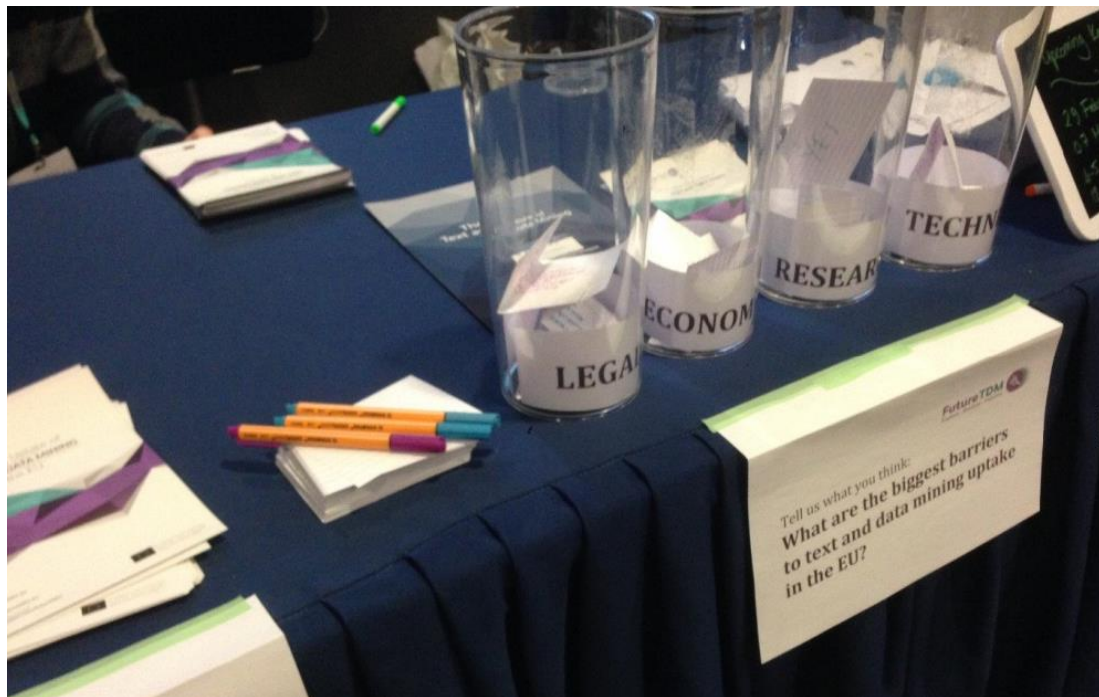


Figure 22: Inviting participants to provide feedback on TDM barriers, Digital4EU Summit, Brussels, 25 Feb 2016



Figure 23: Feedback through stickers on the Knowledge Tree poster, FTDM workshop, Brussels, 29.3.2017

5.4 Twitter

Since Twitter is the most widely used form of social media in the TDM community, this social media channel was established as the main social media channel. A Twitter account was set up at <https://twitter.com/futuretdm>, with tweets about the project including @futuretdm and #futuretdm where relevant. Partners of WP7 took part in Twitter duty: each week a different partner was responsible for ensuring sufficient content to the Twitter account. Specific guidelines for maximising Twitter benefits were sent round to all partners, and included recommendations on what to tweet, and which hashtags to include:

- Tweets should be mostly about the project/events/outcomes (reports, deliverables, newsletters, case studies, awareness sheets)
- Other tweets can be related to:
 - EU-related news (copyright reform)
 - News and developments around TDM
 - Other related projects (OpenMinted, OpenAIRE)
- Suggested hashtags to include / follow:
 - #futureTDM
 - #TDM
 - #textmining
 - #datamining
 - #openscience
 - #bigdata
 - #dataanalytics

Throughout the project lifetime there was an increase in usage and effectiveness of Twitter, with more tweets, more followers and a wider reach. As of 21 July 2017, the FutureTDM Twitter account has 902 followers and 1784 tweets. On average, FutureTDM tweets earn 30K impressions per month, with a record of 66K in September 2016 during the first workshop at the European Parliament. Peak times included the FutureTDM workshops and final symposium. Across the project duration, the potential reach is above 2 million people.

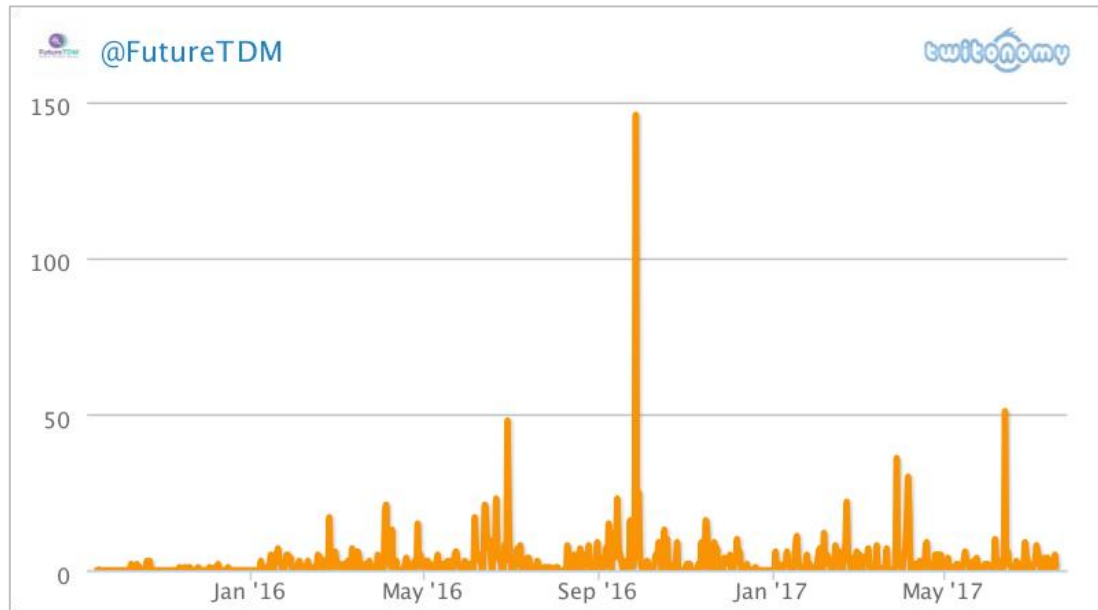


Figure 24: Number of tweets throughout project lifetime, with peaks for the Knowledge Cafes (April-June 2016), workshops (sept 2016 and March 2017) and symposium (June 2017).

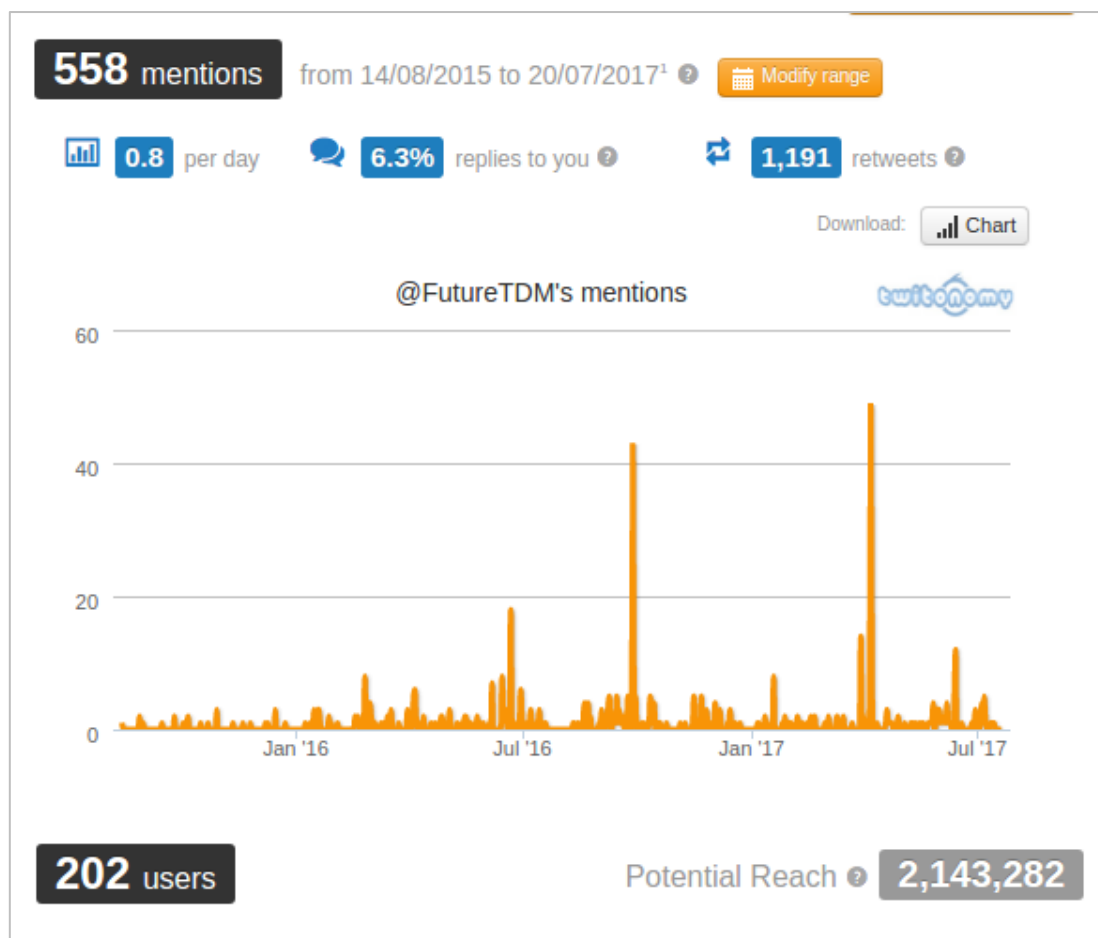


Figure 25: @FutureTDM mentions and potential reach

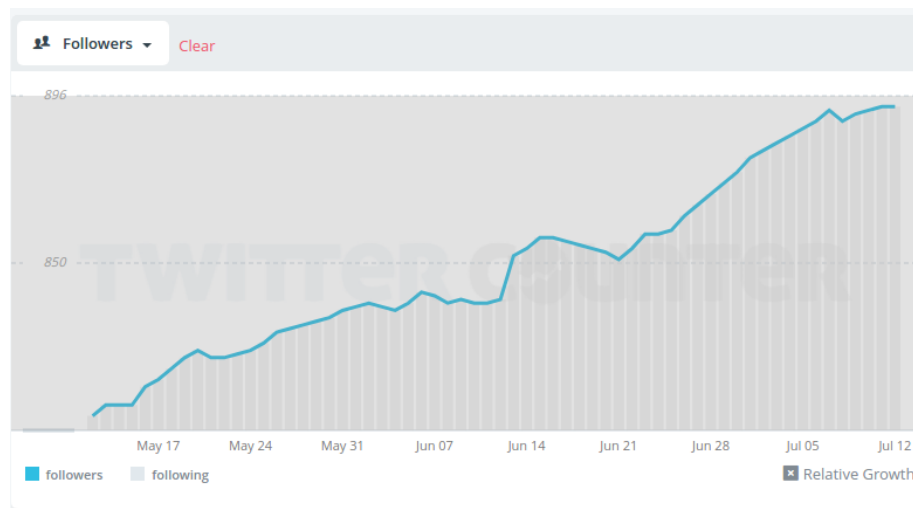


Figure 26: Twitter growth, increase in followers leading up to the end of the project (May-July 2017)

When analysing the audience, followers mostly come from interest fields such as technology, science and politics, with the gender being divided fairly equally. Followers mainly come from the EU and the United States, with English being the dominant language. This gives a similar picture to that of the website and platform statistics.

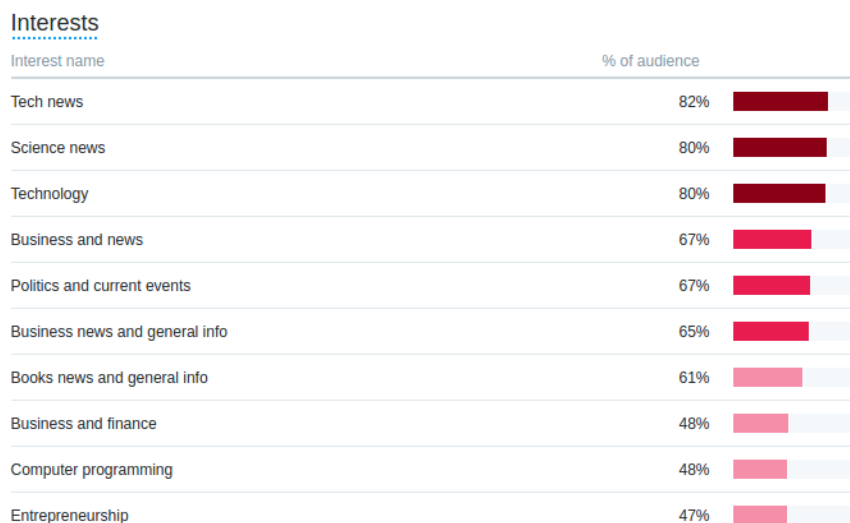


Figure 27: Twitter followers: interests

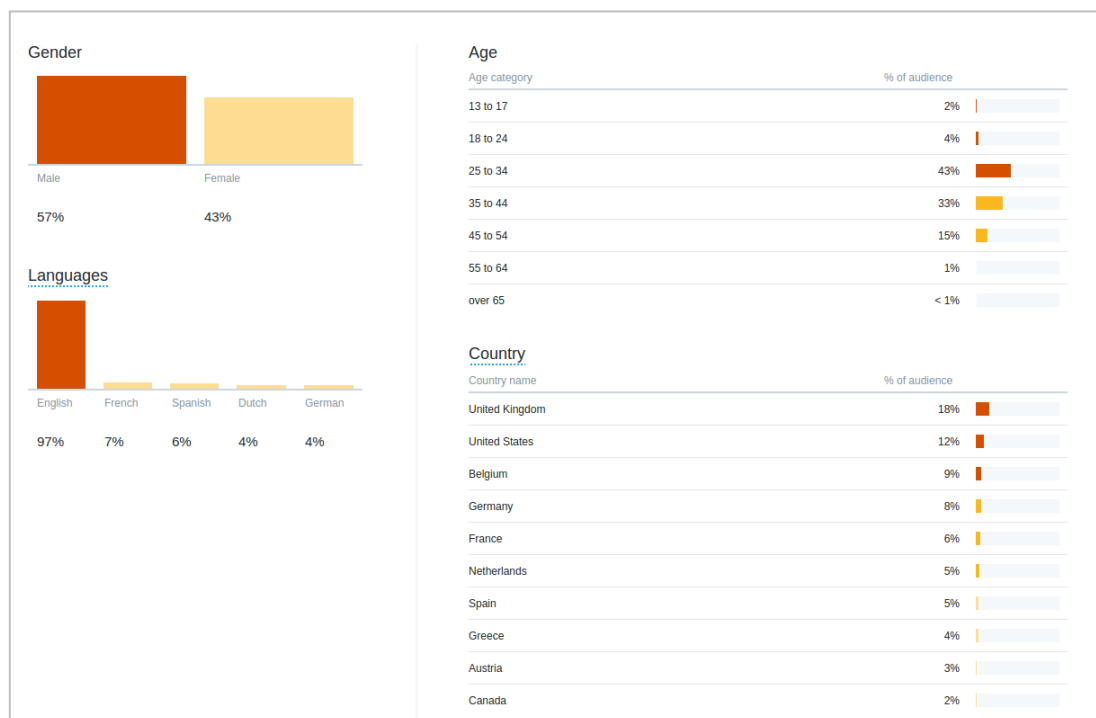


Figure 28: Twitter followers: gender, age and country



Figure 29: MEP Lidia Geringer tweeting on @FutureTDM during the second workshop (March 2017)

5.5 Branding and Print Material

Following the design and templates made in the first months of the project, a number of branded materials were produced for use at events. They included flyers, stickers, feedback cards, awareness

sheets, banners and posters. FutureTDM Templates have been made and available for all partners at the start of the project. A dissemination pack was distributed to all partners in June 2016: materials were also provided online for download. Examples of the various materials are included in Annex II.



Figure 30: FutureTDM dissemination pack (June 2016)



Figure 31: FutureTDM banner next to Julia Reda at the first FutureTDM workshop, Sep 2015

5.6 Newsletter and Mailing List

The FutureTDM mailinglist was set up in February 2016 through Mailchimp. A subscription button was added to both the project website and the platform, as well as promoted on a regular basis through the Twitter account. Newsletters were sent out quarterly since February 2016, with a total of 7 newsletter being sent out throughout the project lifetime:

1. 02/18/2016 - [FutureTDM Newsletter, ed. 1, Feb 2016](#)
2. 06/21/2016 - [FutureTDM Newsletter ed. 2, June 2016](#)
3. 09/12/2016 - [FutureTDM Newsletter ed. 3, September 2016](#)

4. 01/19/2017 - [FutureTDM Newsletter ed. 4, January 2017](#)
5. 04/27/2017 - [FutureTDM Newsletter ed.5, April 2017](#)
6. 06/29/2017 - [FutureTDM Newsletter ed.6, June 2017](#)
7. 08/2017 - *FutureTDM final newsletter, August 2017 - planned*

The content of the newsletter consisted of a mix of various elements: relevant news from the wider TDM field (such as around the EC copyright reform process), an interview with a stakeholder, updates from recent FutureTDM events, a summary of recent blogposts and tweets and pointers to project results, such as deliverables or tools. The interview section was 'newsletter-exclusive' content, ie. published first through the newsletter before being more widely spread through the blog. This ensured fresh content for subscribers that follow the project closely. The following stakeholders were interviewed:

- 'TDM is Simply the Research Method of the Future!', An Interview with Dr Lucie Guibault (University of Amsterdam)
- Interview: Julia Reda MEP (Digital Agenda Intergroup)
- An open-access publisher perspective on TDM - Interview with Frederick Fenter, Executive Editor (Frontiers Media SA)
- Behind the scenes of the 1st International Data Science conference - Interview with Peter Haber and Manfred Mayr (Salzburg University of Applied Sciences)



Figure 32: Interview with Julia Reda at the European Parliament

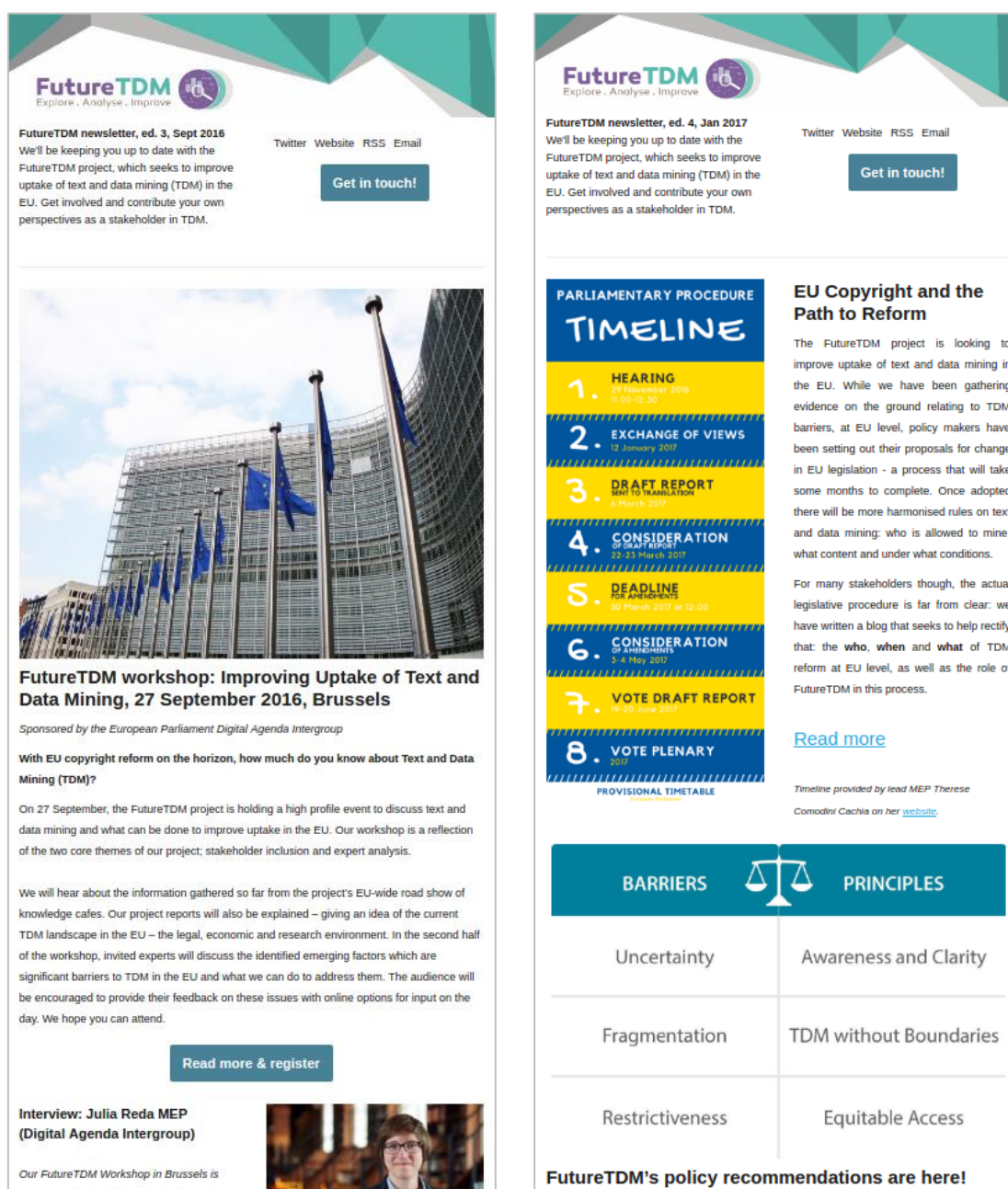


Figure 33: FutureTDM newsletters, edition 3 and 4

Due to the promotion of the newsletter on social media and during events, the amount of subscribers increased from 38 with the first campaign to 181 subscribers for campaign #6. The average open rate is around 40% and the average click rate around 10%, which is relatively high compared to the average email campaign stats of Mailchimp customers per industry² (for the sections government and non-profit, the average open rate is around 25%, with an average click rate of 3%).

List performance

Average open rate	40.4%	Average click rate	10.6%
-------------------	-------	--------------------	-------

² <https://mailchimp.com/resources/research/email-marketing-benchmarks/>

6. PRESS

The project was mentioned **21 times** in external press, most notably **twice by EC MEPs**, following the FutureTDM workshop in Brussels in September 2016:

- Marietje Schaake, 'Future of Text and Data Mining', 13 October 2016, <https://www.marietjeschaake.eu/en/future-of-text-and-data-mining>
- Catherine Stihler, 'Press release: FutureTDM Event at Scotland House', 13 October 2016, http://www.cstihlermep.com/Press_Releases/id912.php

A press release detailing the achievements of the project and all available resources will be produced in August 2017, and sent round to the FutureTDM mailinglist of stakeholders as well as press contacts.

An overview of the press mentions of FutureTDM (not included in this overview are general announcement of FTDM events, or pages on FTDM at partner websites):

1. 02/10/2015, OpenMinTed project site - FutureTDM has kicked off, <http://openminted.eu/futuretdm-kicked-off/>
2. June 2016, eData - FutureTDM goed op weg, http://www.edata.nl/1003_010616/pdf/FutureTDM_goed_op_weg.pdf
3. 05/07/2016, AIMS - FutureTDM : Reducing Barriers and Increasing Uptake of Text and Data Mining for Research Environments, <http://aims.fao.org/activity/blog/futuretdm-reducing-barriers-and-increasing-uptake-text-and-data-mining-research>
4. 14/09/2016, EU Horizon magazine - Copyright shift would put Europe ahead in 'future of research' data mining, http://horizon-magazine.eu/article/copyright-shift-would-put-europe-ahead-future-research-data-mining_en.html
5. 28/09/2016, Copyright4Creativity - Text and Data Mining: how the Future TDM workshop highlighted the draft exception must be improved for TDM to have a future in Europe, <http://copyright4creativity.eu/2016/09/28/text-and-data-mining-how-the-future-tdm-workshop-highlighted-the-draft-exception-must-be-improved-for-tdm-to-have-a-future-in-europe/>
6. 29/09/2016, Science | Business - New EU text and data mining proposal would not benefit everyone, <http://www.sciencebusiness.net/news/79942/New-EU-text-and-data-mining-proposal-would-not-benefit-everyone>
7. 13/10/2016, Marietje Schaake - Future of Text and Data Mining, <https://www.marietjeschaake.eu/en/future-of-text-and-data-mining>
8. 13/10/2016, Catherine Stihler - Press release: FutureTDM Event at Scotland House http://www.cstihlermep.com/Press_Releases/id912.php
9. 18/10/2016, Implisense blog - Interview: Implisense at the FutureTDM Berlin <http://blog.implisense.com/en/interview-implisense-at-the-futuretdm-berlin/>
10. 21/12/2016, UvA news "'Provide a licence that clarifies what others can do with your article and data' interview Lucie Guibault", <http://uba.uva.nl/en/news/news/content/2017/01/guibault.html>
11. 31/01/2017, Data IT law - DATA IT LAW JANUARY 2017: THE RIGHT TO EXPLANATION & EPRIVACY REGULATION DRAFT, <http://www.dataitlaw.com/data-law-january-2017-right-explanation-eprivacy-regulation-draft/>

12. 10/02/2017, Michael Upshall - What can machines discover from scholarly content? <http://consultmu.co.uk/blog/what-can-machines-discover-scholarly-content>
13. 23/02/2017, Éanna Kelly, Science|Business - Post-Brexit, UK could seize competitive edge in text and data mining, <http://sciencebusiness.net/news/80154/Post-Brexit-UK-could-seize-competitive-edge-in-text-and-data-mining>
14. 31/03/2017, Allied for Startups - Newsletter mentioning FTDM workshop Brussels 29 March (email)
15. 03/04/2017, EC infocentre - Breaking down barriers to EU text and data mining uptake, http://ec.europa.eu/research/infocentre/article_en.cfm?artid=43716
16. 06/04/2017, AIMS team - Legal issues on Open Research Data : How Open and FAIR is your dataset? <http://aims.fao.org/activity/blog/legal-issues-open-research-data-how-open-and-fair-your-dataset>
17. 19/04/2017, Zabala Innovation Consulting - Breaking down barriers to EU text and data mining uptake, <http://www.zabala.eu/en/news/breaking-down-barriers-eu-text-and-data-mining-uptake>
18. 05/05/2017, Danny Kingsley - Strategies for engaging senior leadership with RDM – IDCC discussion, <https://unlockingresearch.blog.lib.cam.ac.uk/?tag=policy>
19. 08/06/2017, Copybuzz - Caroline de Cock, 'Ensuring Text and Data Mining has a Future In Europe' <http://copybuzz.com/copyright/ensuring-text-data-mining-future-europe/>
20. 22/06/2017, Claire Sewell - Mining for Data: Skills for TDM <http://www.librarianintraining.com/2017/06/mining-for-data-skills-for-tdm.html>
21. 23/06/2017, CILIP - Copyright exception for text and data mining (referencing FutureTDM expert reports) <https://www.cilip.org.uk/news/copyright-exception-text-data-mining>

In addition, FutureTDM was mentioned on the social media channels of projects in the field, such as OpenMinted, OpenAire and Foster Open Science.

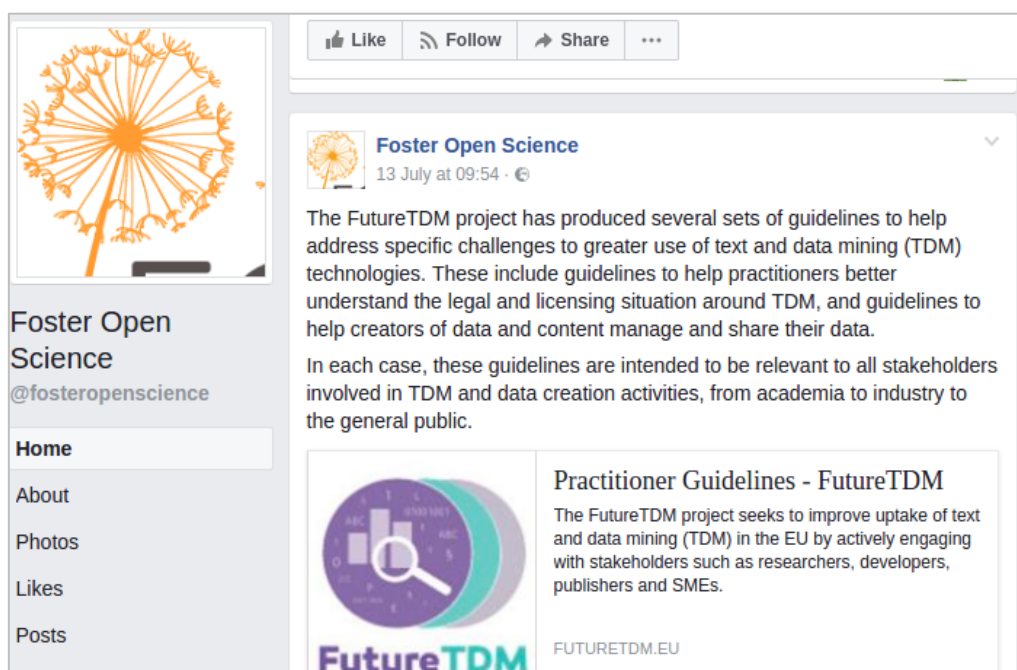


Figure 34: FOSTER Open Science Facebook post, 13 July 2017,
<https://www.facebook.com/fosteropenscience/posts/1427939340576462>

7. PROMOTING BEST PRACTICE AND EXPERTISE

Apart from engaging stakeholders and obtaining their feedback, an important part of the dissemination strategy (especially for the second phase of the project) was focused on promoting the best practice and expertise coming from the project. For this reason, WP7 worked on developing a collection of awareness sheets, a knowledge base and tutorials.

7.1 Awareness Sheet Collection

The awareness sheets (previously called ‘factsheets’) are one of the major dissemination materials of FutureTDM. Throughout the project, they were published to raise awareness of an outcome which may be of high interest to the relevant field, specifically, to the people who work in text and data mining, big data and data analytics. The awareness sheets were created from FutureTDM expert reports, expert interviews and discussions through the Knowledge Café events and cover a range of factors that have an impact on TDM uptake.

Content was revised for the awareness sheet: the sheets were concise (1-2 pages maximum length), the language was accessible and additional images and infographics were used to visualise the issues at hand. Awareness sheets were produced in six different categories: Challenges, Organisations, Stories, Projects and Tools. They were printed out and distributed at relevant events, introduced on the blog and offered through the Knowledge Library on the platform, as well as through the project website, for download (see <http://www.futuretdm.eu/awareness-sheets/>). The collection consists of the following 20 awareness sheets:

Challenges

- Barriers Academic Researchers Experience
- Text and Data Mining as an Economic Asset
- Legal Barriers and Recommendations
- Impact of Text and Data Mining on European Economic Growth
- Legal Issues of TDM: the Relevant Questions
- *TDM as Market Value (August 2017)*

Organisations

- CORE: Aggregating the world’s open access research papers
- Kconnect: Search Technologies for Medical Information
- Paperhive: a Start-up Perspective
- TDM and Contentmine - Barriers and Enablers of TDM
- The Plazi Approach

Stories

- TDM Spotlight: the Start-up Perspective
- TDM Spotlight: the University Perspective
- Microsoft: the International Perspective
- TDM Spotlight: the Space Industry Perspective

Projects

- Openminted: Open Mining Infrastructure for Text and Data
- Research Projects & Infrastructures for TDM in Europe
- *Summary factsheet on the FutureTDM project (August 2017)*

Tools

- Techniques, Tools and Technologies for TDM in Europe
- ContentMine: Zika tutorial



Figure 35: FutureTDM awareness sheets available at the FutureTDM workshop on 27 Sep 2016

7.2 Knowledge Base

The Collaborative Knowledge Base showcases structured collections of resources on Text and Data Mining (TDM) that has been gathered throughout the FutureTDM project phase, especially derived from WP4, in an attractive online format as a reference point for researchers, practitioners and any stakeholders interested in the field. It is accessible through the platform at <http://www.futuretdm.eu/knowledge-base>. The collections encompass experts as projects or organisations focusing on TDM, as well as technologies and resources that are useful for TDM practitioners (i.e. TDM methods and TDM tools).

The beta version of the Knowledge Base was launched in April 2017 and promoted through the newsletter and Twitter, asking the community for their feedback to help improve the collection. At the final symposium, a short walk-through guided participants through using the knowledge base.

More information is available from D6.3 Research projects Directory and Best Practice Library.

7.3 Tutorials

In the second project phase, following the availability of additional research and case studies, project partner ContentMine delivered three hands-on tutorials about three use-cases where text data mining can help practitioners and to show the power of it.

The first tutorial, called 'Zika Virus' shows some basic and advanced text and data mining methods with the ContentMine toolchain, to explore the scientific publications and get a better understanding of the research around the Zika virus. It contained a dataset of more than 1100 publications from EUPMC about the Zika virus. At the start, all the publications were downloaded, then the needed facts extracted from the corpora. The data then was used for all kinds of exploratory analysis - especially with a look on the relations between entities. The tutorial was extended with a video detailing 'How content mining can discover scientific papers about Zika?' (<https://www.youtube.com/watch?v=XtRqohbrlKg>).

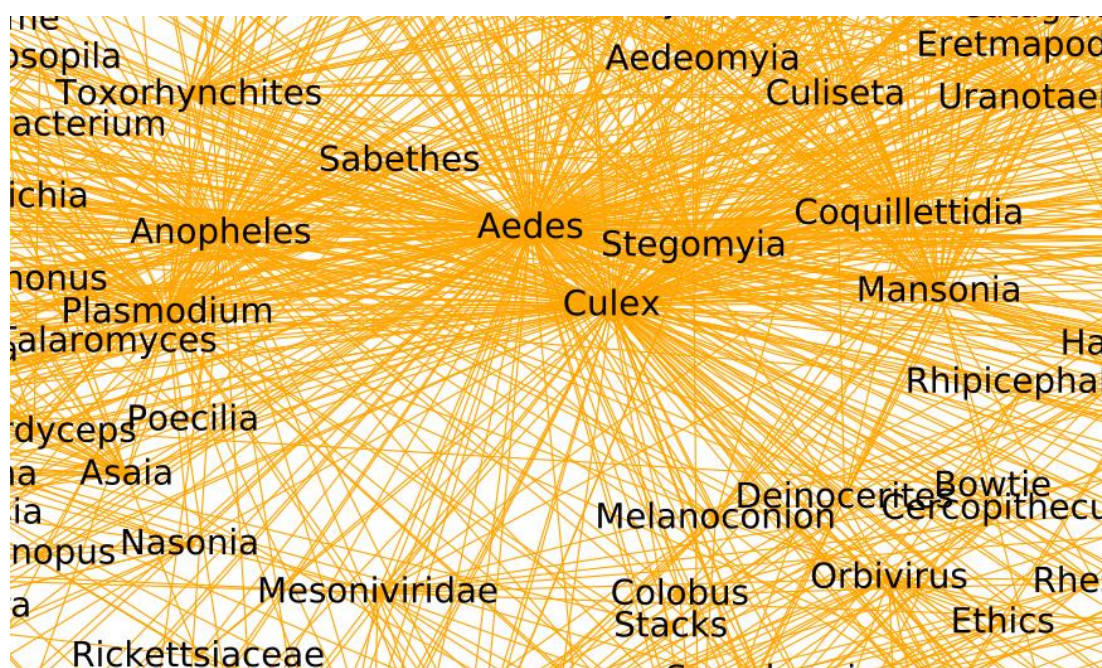


Figure 36: ContentMine

The second use-case, 'Systematic Literature Review (Train the Trainees for Librarians)', was about Systematic Literature Review, which helps Librarians to filter and analyze big number of publications from a research field. We extended the functionality to get the needed data and extract the relevant facts - in a fully open and reproducible way.

The last tutorial, 'P-Hacking', was about the extraction of statistical measures, like p-value, sample size and mean, to get to know more about the usage of statistical methods and to find indicators of p-hacking.

The tutorials were also given during a hands-on workshop at ELPUB in Limassol and presented at the final FutureTDM Symposium in Salzburg. All content and code is openly licensed and available under <https://github.com/ContentMine/FutureTDM>.

8. ANNEX I: METRICS

	<i>Goal for M10 (June 2016)</i>	<i>Result M10 (June 2016)</i>	<i>Goal for M24 (August 2017)³</i>	<i>M24 (August 2017)</i>
Twitter followers	400	374	800	937
Tweets	500	575	1500	1817
Blogs	20	28	60	73
Newsletter subscribers	100	75	200	182
Newsletters	Quarterly from M5 (2 total)	2	Quarterly (7 total)	7 in M24
Website visits	500 unique, 1000 total/month	Website: 1500 Platform: 500	500 unique, 1000 total/month	Website: 3700 Platform: 5100
Publications	At least 2	5	At least 4	5
Conference posters	At least 3	4	At least 6	12
Video (amount)	3	5	20	28
Video (views)	500	418	1500	2200
Banner stands	4 (2 with LIBER, 2 with OK)	4	No new banners planned	4
Awareness sheets (Factsheets)	10	5	20	20

Table 1: FutureTDM KPIs Metrics

³ This goal was set by WP7 in July 2016, following the first EC review of the project.

9. ANNEX II: SELECTION OF PRINT MATERIAL PRODUCED DURING THE PROJECT

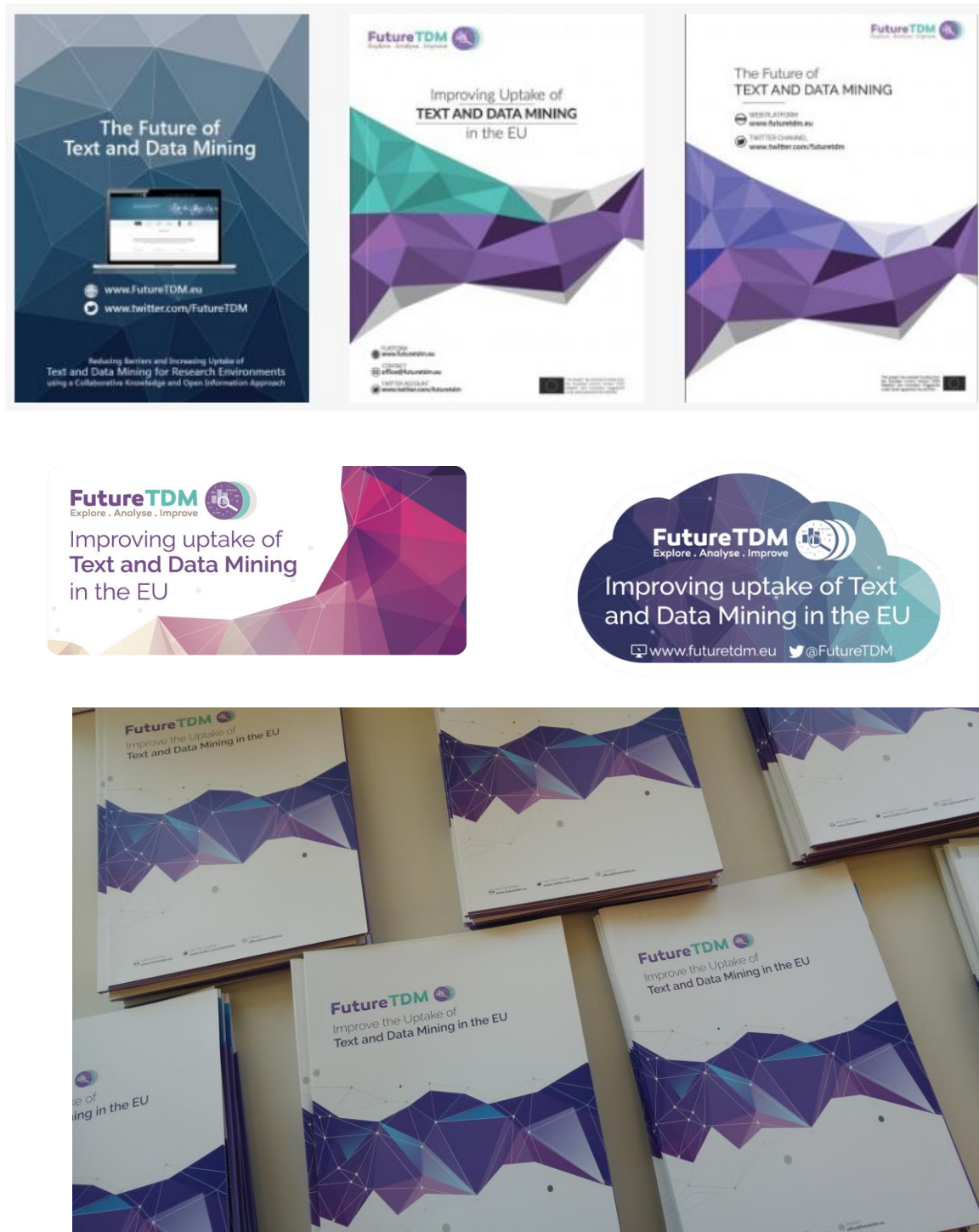


Figure 37: FutureTDM print materials collection

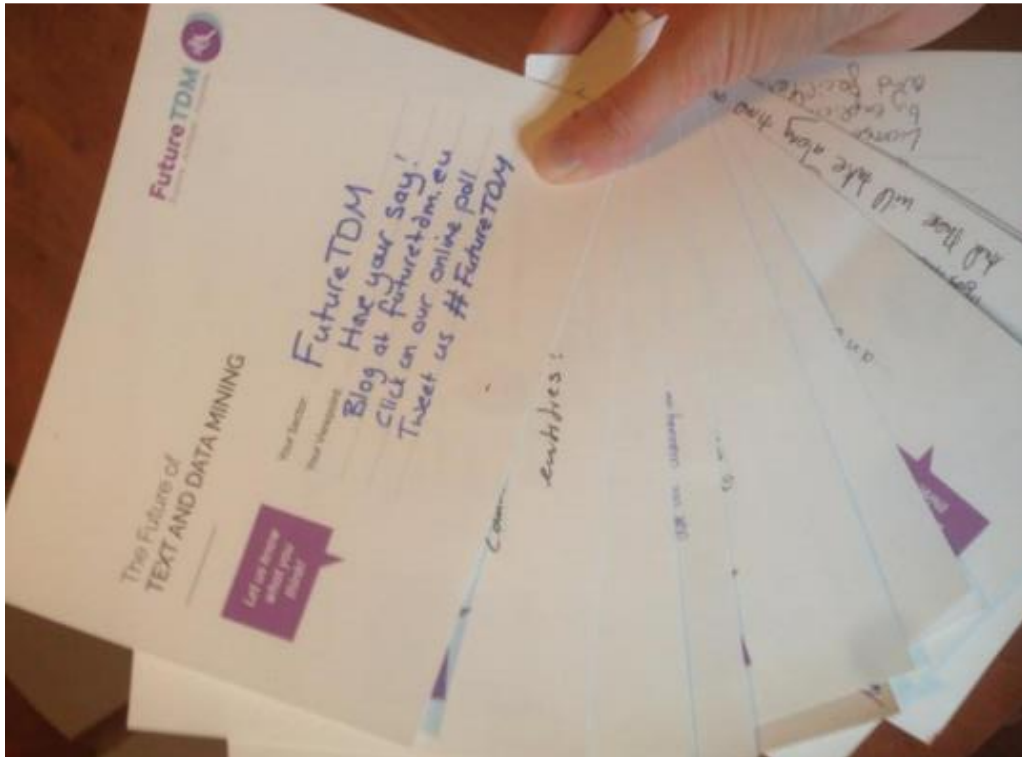


Figure 38: FutureTDM Feedback Cards

Awareness Sheet Collection

**CORE: AGGREGATING THE WORLD'S OPEN ACCESS RESEARCH PAPERS**

CORE is a global large-scale Open Access aggregation platform that offers access to a large volume of free and open access content. Its aim is to aggregate all open access research outputs from repositories and journals worldwide and make them available to the public. In this way, CORE facilitates free unrestricted access to research for all. [Read More](#)

The CORE system harvests metadata records and the associated full-text content from Open Access repositories and journals listed in CORE.

**TEXT AND DATA MINING AS AN ECONOMIC ASSET**

What is TDM? TDM facilitates the extraction of useful and instrumental pieces of information from typically large corpora of essentially unstructured text and other types of data; it also allows for the translation of this information into actionable intelligence for advancing a specific process – be it public policy intervention, market actions or actions performed by other entities for various reasons... [Read More](#)

BIG AND OPEN DATA IN EUROPE

It has been estimated that Big and Open Data will give an incremental boost of 1.9% to European economic growth by 2020.

Buchholtz, S., Bukowski, M., & Sniegocki, A. (2014)

**KCONNECT: SEARCH TECHNOLOGIES FOR MEDICAL INFORMATION**

Radiologists are drowning in images. At larger hospitals more than 100.000 images (over 100 Gb) are produced per day. **What is Kconnect?** Radiologists and other clinicians are facing an information overload caused by an increasing number of images and an increasing complexity of radiological protocols... [Read More](#)

KCONNECT

As a researcher you can even build your own pipeline to combine and process medical texts!

Figure 39: FutureTDM Awareness Sheets Collection

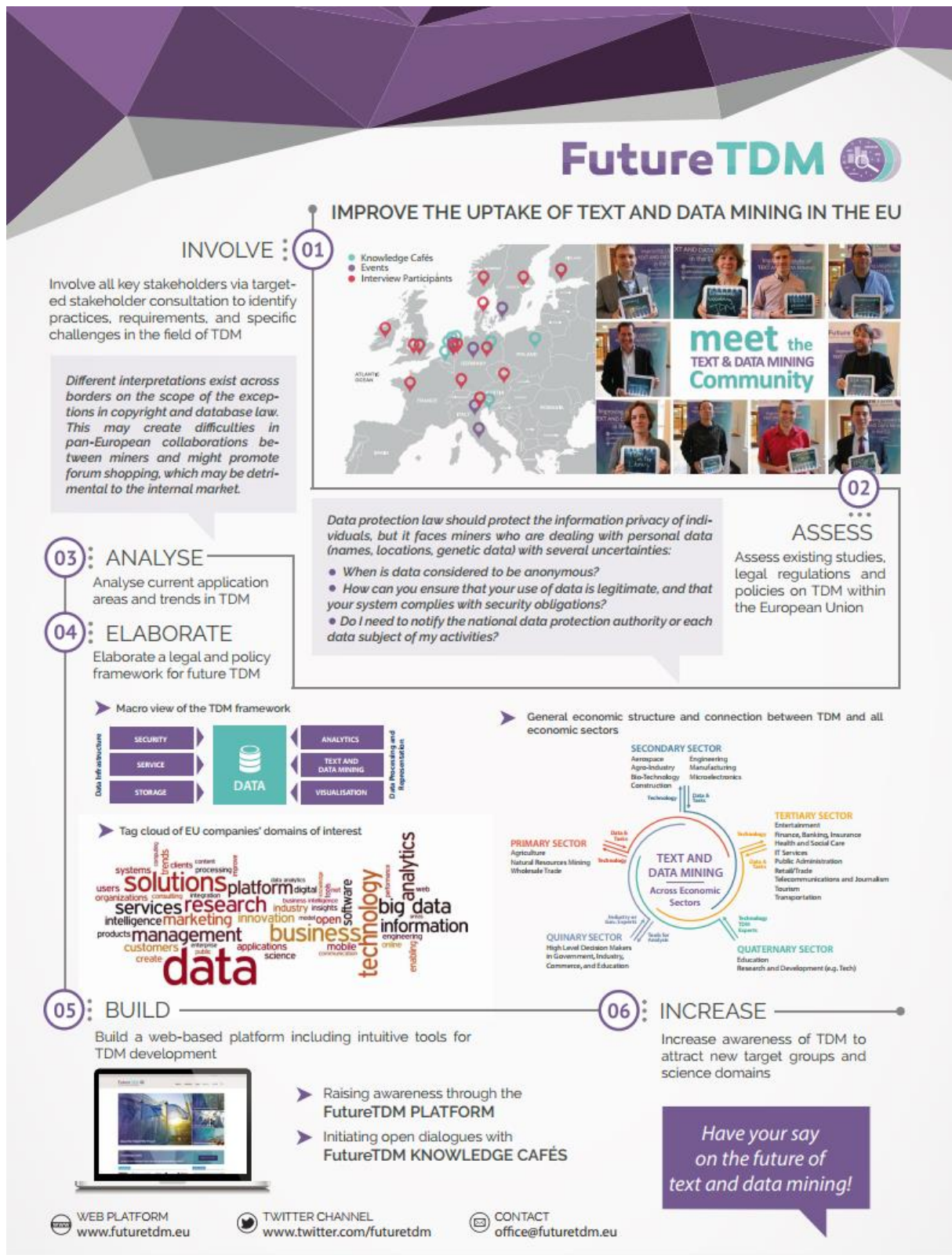



Figure 40: FutureTDM Poster






SUPPORTING GREATER UPTAKE OF TEXT AND DATA MINING IN EUROPE

FutureTDM OBJECTIVES

- INVOLVE all key stakeholders via targeted stakeholder consultation
- ASSESS existing studies, legal regulations and policies on TDM within the EU
- ANALYSE trends in TDM
- ELABORATE a legal and policy framework for future TDM
- BUILD a web-based platform including intuitive tools for TDM development
- INCREASE awareness of TDM to attract new target groups and science domains



Raising awareness through the
FutureTDM PLATFORM

Initiating open dialogues with
FutureTDM EVENTS

FutureTDM Guidelines for STAKEHOLDERS

FutureTDM has just published a series of practical, plain English guidelines to support the uptake of TDM across Europe. You can read them in full on the FutureTDM platform.
www.futuretdm.eu

LEGAL GUIDELINES FOR TDM PRACTITIONERS
GUIDELINES FOR CONTENT LICENCES
DATA MANAGEMENT GUIDELINES FOR RESEARCHERS
GUIDELINES FOR SUPPORTING TDM AT UNIVERSITIES

TDM and the Law

Questions every TDM practitioner should ask

- Is my TDM project lawful?
- Is the content I want to use protected by laws or regulations?
- Are the things I plan to do with the content subject to specific regulations?
- How can I minimise the risk of my TDM activity being unlawful?
- Should I seek expert legal advice?

Data Management for TDM

Machine reading is different to human reading!

Computers need metadata!
Working with hundreds or thousands of documents – or more – means using automated processes to identify which ones are relevant and legally available.

Content repositories should make available key metadata on:

- What licences apply to content
- Data type, format, size
- Any specific tools needed to work with the data
- Data provenance and rights holder information
- Data changes, with versioning

For best practice, permit the widest re-use possible – and use standard, open licences like Creative Commons.

Licences for TDM

Libraries can help!

TDM may involve:

- Copying content that you want to analyse
- Publishing results of your analysis
- ...and practitioners can't always rely on copyright exceptions. Libraries can therefore play a crucial role by ensuring licences are reasonable and appropriate to support researchers carrying out TDM.

When are licence terms reasonable and appropriate in the context of TDM?

Supporting TDM at Universities

Supporting TDM at your institution

Universities are key stakeholders in the TDM value chain
Supporting big data analytics should be an important strategic goal!

THINGS TO CONSIDER:

- Does it make practical sense to distinguish between "commercial" and "non-commercial" TDM research?
- Is usage and activity monitoring intrusive enough to affect researchers' academic freedoms?
- Can researchers reproduce reasonable, illustrative excerpts of content with the results of their TDM analysis?
- Is it practical for researchers to attribute credit to every piece of content used in TDM analysis?
- Do technical protection measures or limitations prevent researchers from carrying out TDM at reasonable scales?

How can you help develop policies and strategies supporting data science and analytics?

- DEMONSTRATE NEED** Gather evidence from research and education faculties on how data management and data science are key skills for researchers.
- INVOLVE STAKEHOLDERS** Get everyone involved in discussions, across all departments and roles, to consolidate support.
- UNDERSTAND YOUR INSTITUTION** How does its organisational structure allow for new policies to be introduced, or new skills added to education streams?
- CONSOLIDATE INFORMATION** Set up a hub to bring together people across the institution with an interest in TDM, and links to any resources you may have.
- IDENTIFY PROMOTERS** There are likely to be people across many fields in your institution who are already interested in supporting TDM – work with them!
- INTRODUCE INCENTIVES** If funding is limited, find out what else is valuable to stakeholders: internal recognition, public promotion online, volunteering for public good?
- SHARE YOUR PROGRESS!** The more success stories we can share about supporting TDM, the easier it will be to build momentum and support new initiatives.

DISCOVER MORE www.futuretdm.eu | office@futuretdm.eu | [@futuretdm](https://twitter.com/futuretdm)
















Figure 41: FutureTDM poster for Practitioner Guidelines

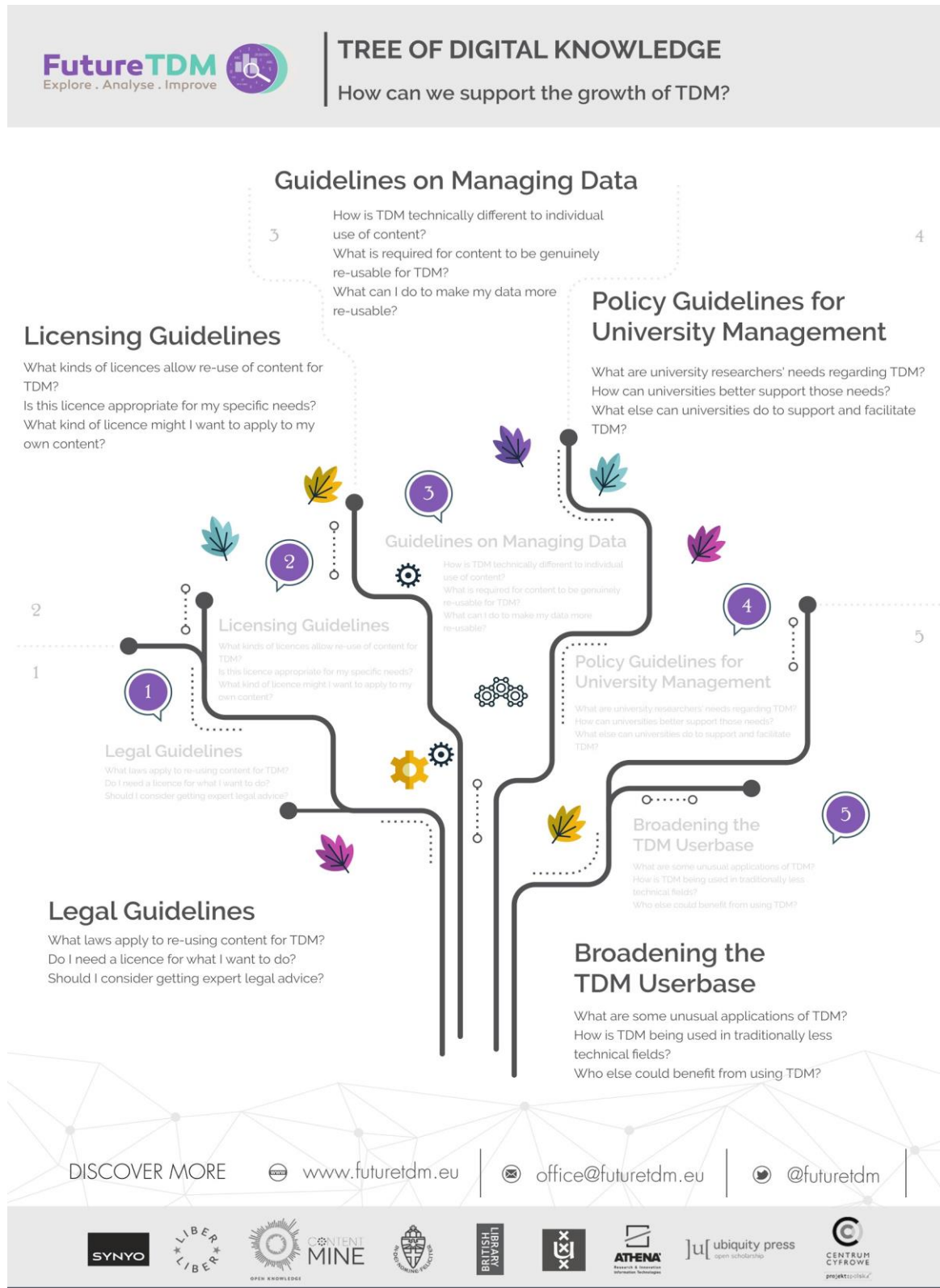


Figure 42: Tree of Digital Knowledge