D7.2: Second report on dissemination

UPVLC, XEROX, JSI-K4A, RWTH, EML and DDS

Distribution: Public

November 8, 2013
The partners in **transLectures** are:

Universitat Politècnica de València (UPVLC)
XEROX Research Center Europe (XEROX)
Josef Stefan Institute (JSI) and its third party Knowledge for All Foundation (K4A)
RWTH Aachen University (RWTH)
European Media Laboratory GmbH (EML)
Deluxe Digital Studios Limited (DDS)

For copies of reports, updates on project activities and other **transLectures** related information, contact:

**The transLectures Project Co-ordinator**
Alfons Juan, Universitat Politècnica de València
Camí de Vera s/n, 46018 València, Spain
ajuan@dsic.upv.es
Phone +34 699-307-095 - Fax +34 963-877-359

Copies of reports and other material can also be accessed via the project’s homepage:
http://www.translectures.eu

© 2013, The Individual Authors

No part of this document may be reproduced or transmitted in any form, or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission from the copyright owner.
1 Executive summary

1.1 Scope of this document

The purpose of this document is to outline dissemination strategy and activities carried out for the transLectures project. This is the second such report and will provide an overview of activities from year two of the project. We also evaluate activities and set goals for the next and final reporting period (Y3).

1.2 Recap of goals set for Y2

At the start of this period, a first set of subtitles had just been made available on the poliMedia repository, and user evaluations at both case studies were in their planning stages. The release of a first transLectures toolkit was also close on the horizon. All constitute key milestones both in their own right and in terms of future uptake. Consequently, the specific goals set for dissemination in year two were:

- To zoom in on opportunities for reaching out to potential new user/“prosumer” communities;
- To support and complement work being carried out in WP6 on user evaluations;
- To redesign and write new content for the project website to make it more accessible to non-scientific browsers.

A description of the steps we have taken towards achieving these goals follows in the main document.

1.3 Highlights

Some of the highlights in dissemination in this reporting period are as follows, in chronological order:

- Co-Creation of Emerging Trends in Academia workshop, a very successful event organised by JSI-K4A where number of prominent speakers were invited by transLectures to discuss the uptake and future direction of OERs, lecture capture, and language technologies [3.1];
- Improved website and web presence generally [4.1.3];
- EUCOIII/PASCAL2 IASD challenge, organised by the UPVLC and co-sponsored by transLectures [3.1];
- Release of the first transLectures toolkit, the UPVLC’s open source TLK toolkit for automatic speech recognition [3.3.1];
- JSI-K4A’s bid to host OCWC2014 was accepted, which, as co-sponsor, will bring transLectures into direct contact with key stakeholders at a timely point in the project [3.3.2];
- UPVLC’s external poliMedia pilots underway, which will not only increase the volume of feedback but also serve to test the feasibility of our business model, outlined in D7.4 [48] (3.3.3).

More details on all of the above will follow in the main document.
2 Strategy

2.1 Review of strategy as set in Y1

In D7.1 we identified transLectures’ main stakeholders as researchers, practitioners and other players in the following fields:

- Machine translation (MT), automatic speech recognition (ASR), natural language processing (NLP), machine learning;
- Lecture capture, audiovisual repository management;
- Education: online, open, technology-enhanced, digital educational resources.

The first group comprises mainly academic researchers, so dissemination in this area involves the publication of scientific papers, and poster sessions and presentations at academic conferences, with the aim of sharing findings and advancing the state-of-the-art. The stakeholder community corresponding to the latter two fields, meanwhile, is rather more broad and the goal of dissemination activities here is to expand the immediate circle of potential users and “prosumers” beyond the two case study sites and therefore pave the way for the uptake of transLectures tools. Not forgetting, of course, the general public, who are targeted through general press releases at local and community level.

It was established that dissemination activities should be aimed at all of the above stakeholders throughout the duration of the project.

See D7.1 [47] for a full outline of strategy and activities carried out in Y1.

2.2 Strategy going into Y2

Dissemination strategy going into Y2 was to continue much in the same vein, though this time with the emphasis on presenting first results rather than simply outlining the nature and goals of the project. Also, with the release of the first transLectures toolkit on the horizon, and talks at the Y1 progress meeting of running beta tools at poliMedia and OCWC, there was to be particular focus on our industrial stakeholders and, through them, the potential user community.

Dissemination activities were to attract interest in our tools, gain feedback from users and consolidate relationships with organisations that could help promote uptake and/or future development of project technologies.

2.2.1 Anticipated release of TLK tools

Inspired by talks at the third project meeting in Ljubljana (M12), discussions began about releasing a first transLectures toolkit. Transcription technologies in Spanish at least were at a point where this was viable and releasing a toolkit would put the project in direct contact with “prosumers”. It was decided that the UPVLC would work to release a first version of their toolkit for automatic speech recognition during this reporting period, which would of course have a direct impact on project dissemination.

The transLectures-UPV toolkit “TLK” was ultimately released in M19. We would refer you to Section B.3.1 for the details.
2.2.2 Plans for poliMedia and OCWC pilots

Identified in D7.1 as a resource for dissemination, plans to tap into the poliMedia community were firmly on the cards at the start of Y2. As mentioned above, Spanish transcription technologies were at a point where they could be usefully exported to other universities, especially given this relatively controlled context. By doing so, we would broaden our pool of “prosumers”, get more user feedback and, as later cited in the Business Plan Draft (D7.4), test the feasibility of the proposed business model.

OCWC pilots, meanwhile, were and still are envisioned for later on in the project (Y3), dependent upon results in English transcription.

See Section [3.3.3] to see progress made in this area.

2.3 Goals set for Y2

In addition to general project goals and motivations, the specific objectives set for dissemination in year two were:

- To zoom in on opportunities for reaching out to potential new user/“prosumer” communities, to the extent possible creating a community around transLectures;
- To support and complement work being carried out in WP6 on user evaluations;
- To redesign and write new content for the project website to make it more accessible to non-scientific browsers.

2.4 Related milestones

2.4.1 End-of-year review

At the end-of-year review held in Luxembourg on 18th December, it was recommended that we leave the professional language sector out of the equation, at least for the immediate future. Given the specific context of the technologies being developed, it made more sense to zoom in on video authors, video providers and other users/“prosumers” in the context of the creation and management of educational audiovisual resources, lecture capture, video repositories, etc.

2.4.2 Business Plan Draft (D7.4 [48])

From around M14 until delivery in M18, we drew up the Business Plan Draft. This process led to more in-depth consideration of the most probable market(s) for transLectures technologies and exactly how we might achieve their use and, ideally, further development beyond project end. It therefore had a large hand in directing dissemination strategy from this point onwards and sparked greater internal communication on this point.

In this document, VideoLectures.NET, the extended poliMedia community, Opencast Matterhorn and OCWC were all identified as key targets for dissemination actions. VideoLectures.NET and poliMedia as case studies and subsequent showcases for project technologies, and the latter organisations as potential gateways to more widespread adoption among the international education community.

We also identified two main uses of transLectures technologies: one, to provide multilingual subtitles and two, to provide quality text data that can be used to power advanced repository
management and other functions related to the exploitation of Big Data. In the first instance, this means targeting the education sector, specifically the open and online education/technology-enhanced learning communities; and the audiovisual content management and research communities in the second.

This division effectively called for a reinterpretation of the first objective set for dissemination, where objective 1a would correspond to a community of users/“prosumers” who would use transLectures tools or services to enable transcription and translation of their videos and objective 1b to a community of researchers or developers who would exploit the underlying project technologies to create other innovative tools or services.

See D7.4 for full analysis of our potential market, closer examination of our likely “prosumers”, as well as proposed business models and plans at partner level.

3 Activities record (Y2)

Here we give a chronological list of all activities carried out in dissemination in this reporting period. We strongly recommend you consult our web entries, where available, for further commentary (links provided). Additionally, we have included more in-depth discussion of some of the highlights in Section 3.3.

3.1 Events timeline

M13: November 2012

1. 07.11.12: Organised Co-creation of Emerging Trends workshop - Ljubljana (Slovenia)
   Organised primarily by JSI-K4A, this event brought together students, researchers and other professionals from diverse technological, social and legal backgrounds to discuss the ways in which new technologies are shaping the future of education. From the point of view of project dissemination, the workshop was hugely successful and, above all, helped put what we are trying to achieve at transLectures into a broader context in the minds of many of our potential stakeholders and collaborators. It also served to consolidate relationships with key organisations within the field of (in particular, open) education, among them the OCWC.

   A description [43] of the workshop and a summary [40] of the talks given can be found on the transLectures website. The videos themselves can be viewed on VideoLectures.NET [49].

2. 13.11.12: Attended REC:all pre-conference workshop - Leuven (Belgium)
   As part of the event “Lecture Capture: current best practices and future developments in universities”, Carlos Turró (UPVLC) gave a talk entitled ‘transLectures: tackling the translation issue in a cost-effective manner’.

   See our blog entry [30].

3. 14/15.11.12: Attended Media and Learning 2012 conference - Brussels (Belgium)
   Carlos Turró (UPVLC) again gave his ‘transLectures: tackling the translation issue in a cost-effective manner’ presentation.

4. 21/23.11.12: Attended IberSpeech 2012 conference - Madrid (Spain)
   A key date for Iberian language speech technology research, UPVLC presented the project and early results for poliMedia in a poster session. They also entered the IberSpeech Audio Speech Segmentation competition, placing second overall.

   See our blog entry [27].
5. 21/23.11.12: Attended Languages and the Media 2012 conference - Berlin (Germany)

Davor Orlic (JSI-K4A) presented transLectures at this 9th international conference on language transfer in the audiovisual media.

6. 28/31.11.12: Attended Online Educa Berlin conference - Berlin (Germany)

transLectures was present at this international conference on technology-supported learning, presenting the project as part of a two-hour “talking head” session (UPVLC).

M14: December 2012

M15: January 2013

7. 08.01.13: Delivered REC:all webinar

Gonçal García (UPVLC) presented the project in a one-hour webinar to an interested, non-specialist audience as part of the REC:all webinar series. Feedback was very positive and the presentation was followed by an article in the January edition of the Media and Learning News newsletter.

Watch the video [6], see our blog entry [35] or read the newsletter [5].

8. 09.01.13: Launched new-and-improved website [19]

UPVLC delivered on their promise to revamp the project website and use of social media. The new website went live at the beginning of January, with a more contemporary design and new content written with non-scientists in mind. It is very much an ongoing, collaborative process, and new ideas are continually being worked into the design. Resources are also being devoted to keeping the site up-to-date with project events and progress (UPVLC, JSI-K4A). The same is true of the transLectures Facebook [20] and Twitter [21] accounts.


Filed by: XEROX.

Description: This patent protects the adaptation method described in Section 4.1.1 of D3.1.1 [18].

M16: February 2013

10. 07.02.13: Opened a transLectures channel [45] on VideoLectures.NET

Listed alongside other European projects related to this video repository, this video channel includes a brief overview of the project and, naturally, a video channel which is a work-in-progress. So far we have five videos, including the two project video demos and the REC:all webinar.

11. 19/27.02.13: Press coverage of VideoLectures.NET and its upcoming WSIS+ award - Slovene

In light of the WSIS+10 awards (see below), though also as part of a broader dissemination campaign aimed at the OCWC and Opencast communities, VideoLectures.NET received a lot of press within Slovenia, on local and national TV and radio stations [11] [12] [13] [14] [15] [16] [3]. It is also reported the REC:all [17] blog.

12. 25/27.02.13: VL.NET received award at World Summit on Information Society 10 Year Review Meeting (WSIS+10) - Paris (France)
VideoLectures.NET was selected as one of eight best-of-the-best in eContent of the last decade at this gala held at UNESCO headquarters. JSI-K4A attended this gala, presenting transLectures results offline to a number of attendees from emerging third world markets where the main language of communication is English. Particular interest was expressed by UNESCO officials from Education and the Permanent Delegation of the Republic of Slovenia, and in relation to the use of transLectures tools in mobile learning, and in collaboration with UNESCO’s International Institute for Capacity Building in Africa.

See our blog entry [42]

M17: March 2013

13. 11.03.13: VideoLectures.NET in the Digital Agenda [2]

VideoLectures.NET’s presence at the WSIS+10 awards was also reported in the Digital Agenda for Europe.

14. 12.03.13: Attended 4th International Symposium on Live Subtitling conference - Barcelona (Spain)

Held at the Universitat Autònoma de Barcelona, this event is the fourth in a series of symposiums on live subtitling and respeaking technology. Yota Georgakopoulou (DDS) presented transLectures.

See our blog entry [34]

15. 20.03.13: Featured in the PASCAL2 brochure [2]

The project featured prominently in the PASCAL2 brochure, which highlights the success stories and resulting legacy of this Network of Excellence whose activity drew to a close during this reporting period. It is clear from this publication, and PASCAL’s support of the IASD challenge, that transLectures is considered one of PASCAL2’s main legacies, which puts the project in a strong position within the field of artificial intelligence, particularly as regards exploitation of results in future research projects.

M18: April 2013

16. 01.04.03: Press coverage of IASD challenge

The forthcoming competition was reported on by the UPVLC’s in-house science journalist, and has so far appeared in the web version of one Spanish newspaper [10].

17. 11.04.13: Organised EUCOGIII/PASCAL2 IASD challenge - Palma de Mallorca (Spain)

Launched on 12 February 2012, the aim of this challenge was to explore the different ways in which user interaction can be intelligently guided by the system to produce the best end results under varying scenarios of user effort. The motivations underlying this challenge were two-fold: first, to advance progress in intelligent interaction both in general and with particular relation to the project, and second to spread knowledge of this line of research and the project as a whole to a wider academic community.

In total, some 90 systems were submitted by eight participants from some of the most prestigious language recognition research centres in Europe. Two of the three winners were able to attend the awards presentation, held at the joint EUCOGIII/PASCAL meeting in Palma de Mallorca, on 11 April 2013. More importantly, the best system represented an improvement on existing transLectures techniques and will therefore help drive forward advances in WP4.

Key dates:
12.02.2013: First phase of challenge launched
25.03.2013: End of first phase: best three systems selected
26.03.2013: Second phase launched (for best three systems only)
05.04.2013: End of second phase: winners ranked and announced
06.04.2013: Deadline for reports

For full details please consult the challenge web page [22] and our two blog entries: call for entries [33] and winners announcement [32].

18. 09.4.2013: Meeting with Irina Bokova, Director-General of UNESCO - Brdo pri Kranju (Slovenia)

JSI-K4A were able to meet with Bokova during her visit to Slovenia for the Meeting of the Council of Ministers of Culture of South East Europe (CoMoCoSEE). In a private discussion, they presented the award-winning VideoLectures.NET and the K4A Foundation, also discussing transLectures in the context of deploying our tools on VideoLectures.NET. This is the second time UNESCO has expressed interest, albeit informally, in our transcription and translation tools, particularly for use in the African context.

M19: May 2013

19. 03.05.13: Released first version of the transLectures-UPV toolkit, TLK

Considerable effort was devoted during the M12-18 reporting period to the development of a first set of transLectures tools for transcription (UPVLC). A first version of these tools, TLK, was released on 3 May 2013 under the Apache License 2.0 [4].

A key milestone, further discussion can be found in Section [3.3.1].

20. 08/10.05.13: Attended OCWC Global 2013 conference - Bali (Indonesia)

JSI-K4A attended this global conference held by OCWC, one of our main targets for transLectures uptake. The organisers of one of two pre-conference workshops entitled “Advanced Methods and Tools for Web-based Education”, they presented transLectures alongside two other EU-subsidised projects related to their VideoLectures.NET: X-Like and Mediamixer.

21. 26/31.05.13: Attended ICASSP 2013 conference - Vancouver (Canada)

The world’s largest and most comprehensive technical conference in the area of signal processing and its applications, UPVLC and RTWH-Aachen attended this event, presenting the papers “Language model adaptation for video lecture transcription” and “Comparison of Feedforward and Recurrent Neural Network Language Models”, respectively, as part of the Spoken Language Processing: Language Modeling, for Speech and SLP poster session.

See our blog entry [28].

22. 25.05.13: Produced a new demo video of the transLectures interactive player

In advance of the ESWC2013 (see below), the UPVLC put together a new demo video to demonstrate the new intelligent interaction mode being used in the internal evaluations at both poliMedia and VideoLectures.NET.

See the video [39].

23. 28.05.13: Attended ESWC2013 conference - Montpellier (France)

A key event in Big Data and Web Semantics research, project lead, Alfonso Juan, and Jorge Civera (both UPVLC) attended the conference to present the project, an overview of results so far, plus the new demo video at the EU Project Networking session. At this event we made contact with other EU projects X-Like and ViSTA-TV.

See our blog entry [26].
M20: June 2013


Davor Orlic (JSI-K4A), alongside UCL lecturers, proposed a session entitled “Open Educational Resources: re-using, re-mixing and re-purposing content” at this one-day event to discuss the development of digital pedagogies in university teaching. The angle being worked at this event was transLectures as a provider of quality raw text data.

See our blog entry [20].

25. 24/26.06.13: Attended Opencast 2013 Unconference - Osnabrück (Germany)

Held in cooperation with the DIVERSE Conference at the University of Osnabrück, more than 100 people came to present and listen to talks on various issues relating to video in academia and the Matterhorn platform in particular.

A key community for transLectures dissemination, Carlos Turró (UPVLC) attended to present the latest upgrades to the Paella player, being used in WP5. Davor Orlic (JSI-K4A) was also there on behalf of transLectures.

26. 06.13: Ljubljana confirmed as the venue for OCWC2014

JSI-K4A’s bid to host OCWC2014 was accepted, which is great news for transLectures. As co-sponsor we will be well-positioned to approach potential adopters of project technologies at a timely stage in the project.

See our blog entry [37].

M21: July 2013

27. 02.07.13: Uploaded new videos to the transLectures channel on VideoLectures.NET.

Specifically, the two demo videos produced to date (UPVLC) and the REC:all webinar.


In collaboration with Olaf Schulte, we drafted a piece for the Matterhorn “success story” feature, describing the progress made towards the project’s third scientific goal of compatibility and integration with the Matterhorn platform.

See our blog entry [24] and the article itself [44].

29. 07.13: First feedback from external users

In July we began to get feedback regarding the quality of transLectures transcriptions from other Spanish poliMedia universities within Spain.

See Section 3.3.3 for further information.

See our blog entry [38] and follow-up post [36].

M22: August 2013

M23: September 2013

30. 2-6.09.13: Attended MT Summit 2013 - Nice (France)

Jesús Andrés-Ferrer (UPVLC) presented our latest results in a special European projects poster session. In a “poster boaster” round he also presented a 60-second video showing our intelligent interaction interface in action.

See our blog entry [31] and the sneak-peek video [23] shown at the event.

This theme for this year’s conference was “Connecting Europe for New Horizons”, aimed at powerful language technologies for the multilingual information society, the data value chain and the information market place.

Marion Mast (EML) attended this two-day event and presented transLectures in the session ‘Services for (Multilingual) Europe: The Industry Side’.

See our blog entry [29] for photos and a link to the presentation itself.


Organised in partnership with Intel and hosted by Danube University, the event focused on the concept of microlearning. Davor Orlic (K4A) attended with the talk entitled: “Micro Open Educational Resources (Micro OERs) - a case study in enabling creation, usage and remixing of video generated educational media fragments”. As part of this presentation, transLectures results and potential for exploitation in this field were presented.

Web entry coming soon.

33. **30.09.13: Science Dissemination and On-line Certification for All - Trieste (Italy)**

transLectures was invited to give a talk at this workshop organised by the Abdus Salam International Centre for Theoretical Physics’ (ICTP) Science Dissemination Unit and co-sponsored by TWAS, a global science academy based in Trieste. The aim of this event was to guide the scientific community in developing countries in the adoption of innovative, low cost, pedagogical tools for the creation and dissemination of online educational resources.

Carlos Turró (UPVLC) attended, to present project results and share our experience within this emerging field with an audience very interested in adopting transcription technologies in the near future. One attendee even downloaded our transcription tool, TLK, during the conference itself.

The ICTP itself was most interested in our English-language technologies, for the transcription of its entire video repository dating back some five academic years; other participants from around the globe were excited about the possibility of getting transcriptions in other languages which unfortunately do not fall within the scope of transLectures.

**M24: October 2013**

34. **1.10.13: Neth-ER seminar: Reflections on “Opening Up Education” - Brussels (Belgium)**

The goals of this seminar was to bring together education executives and management, policy-makers, educators and other practitioners to reflect on this new EC initiative. Davor Orlic (JSI-K4A) attended on behalf of transLectures.

35. **2.10.13: EADTU-EU Special Summit on Higher Education - Brussels (Belgium)**

Directly linked to the launch of the recent “Opening Up Education” initiative, the board of EADTU organised a special summit on open and online learning in higher education. Davor Orlic (JSI-K4A) attended on behalf of transLectures.


UPVLC will be presenting its paper: A System Architecture to Support Cost-Effective Transcription and Translation of Large Video Lecture Repositories.


Davor Orlic (JSI-K4A) will attend this two-day event, presenting transLectures to representatives from translation, linguistics, terminologists and other industry professionals. The event is aimed at participants from the countries of former Yugoslavia (including Slovenia (JSI)), and other countries in Eastern and Southern Europe.
And beyond

38. 11/12.11.13: Co-sponsor of Internet of Education 2013 conference - Ljubljana (Slovenia)
   A continuation of the hugely successful Workshop on Co-Creation of Emerging Trends in Academia (1, above), JSI-K4A is organising the first regional Internet of Education conference in Slovenia. The main goal of this three-day event is to identify how emerging technologies based on artificial intelligent and computer-human interaction tools can change and help create new trends in education. transLectures is co-sponsor, alongside the MediaMixer and Xlike projects.
   The event has received UNESCO patronage, is co-organised by the University College London (UCL) and so far has 19 speakers, including speakers from UCL, JSI, Stanford, SMEs, and MOOC start-ups; and 130 participants.
   Further information can be found [website].

39. 04.12.13: To attend Online Educa Berlin 2014 - Berlin (Germany)
   JSI-K4A will be holding a half-day workshop entitled “Artificial Intelligence Methods for Online-based Education” at this conference. They will also be also presenting transLectures.

40. 11.12.13: To attend REC:all lecture capture workshop - Leuven (Belgium)
   Carlos Turró (UPVLC) will be attending this pre-conference workshop.

41. 6/8.1.14: JSI-K4A to hold the 1st Winter School on Multimedia Processing and Applications - Dublin (Ireland)
   Although strictly a MediaMixer activity, it will also serve as an opportunity to disseminate transLectures technologies.
   See the official [website] for more details.

42. 04.14: Co-sponsor of OCWC Global 2014 conference - Ljubljana (Slovenia)
   JSI-K4A have been successful in their bid to host OCWC2014 in Ljubljana (Slovenia). Being held for the first time in South-East Europe, it is being held in synergy with the EC’s “Opening up Education” initiative.
   transLectures will have a dedicated satellite workshop immediately prior to this conference.
   See our online announcement [37].
   Please also refer to Section 3.3.2 for more information.

3.2 Publications

For the sake of convenience, here we give the complete list of all publications since project start.

3.2.1 2012

- Integrating a State-of-the-Art ASR System into the Opencast Matterhorn Platform.
• **Character-Based Handwritten Text Recognition of Multilingual Document.**

• **transLectures.**

• **Explicit length modelling for statistical machine translation.**

• **Does more data always yield better translations?**

3.2.2 2013

• **Comparison of Feedforward and Recurrent Neural Network Language Models.**

• **Language model adaptation for video lectures transcription.**

• **A System Architecture to Support Cost-Effective Transcription and Translation of Large Video Lecture Repositories.**
  Authors: J.A. Silvestre-Cerdà, A. Pérez, M. Jiménez, C. Turró, A. Juan, & J. Civera. Accepted at IEEE SMC 2013. (UPVLC)

3.3 Highlights

3.3.1 Release of transLectures-UPV TLK tools

In M19 the UPVLC released the first version of its toolkit for automatic speech recognition (ASR), the TLK toolkit. Available under an open source licence (specifically, Apache 2.0 [4]), it contains tools for building, training and applying acoustic models that can be used, by people or groups with a background in programming, to generate transcripts for video lectures.

To disseminate this toolkit, we made use of specialist mailing lists, the project website and social media accounts, and LinkedIn groups. We also promote it at events. In fact, two downloads were the direct consequence of our presentations at the UCL’s Digital Pedagogies
It is available for download on our website, along with full documentation, in English. See our [Tools page](#). A second version is currently under development.

In order to monitor downloads, we have required downloaders to first contact us via a dedicated email address and state briefly how they intend to use transLectures tools. This step has proven its worth on a number of occasions: for example, we have been able to manage downloader expectations where non-IT specialists have got in touch expecting a “plug and play” transcription software.

So far we have been contacted by 27 individuals, of which 19 have proceeded with the download. The majority of downloaders are researchers within the field of NLP, our intended user group for these tools.

That said, there has also been a fair amount of interest from video repository owners, and other video users and “prosumers”, whose reasons for wanting subtitles confirm the very premise of the project as set out in the Description of Work: to allow access to the deaf and hard-of-hearing, and to enhance access to the material by non-native speakers.

**Note:** We have plans to extend our Tools page to include links to all of the open source tools used by transLectures partners to develop project technologies, as identified in Section 2 of the Business Plan Draft (D7.4 [48]).

### 3.3.2 JSI-K4A to host OCWC 2014

As mentioned above, JSI-K4A were successful in their bid to host OCWC2014 in Lubljana, Slovenia, in April 2014 (M30). They will also organise a themed satellite workshop to showcase project technologies at what will be a timely opportunity to promote the widespread uptake of transLectures tools.

The OpenCourseWare Consortium (OCWC) is an active community of higher education and related institutions committed to advancing open education and its impact on educational opportunities worldwide. Its OCWC global conference is a major event in open education and brings together policy makers, researchers, practitioners and educators from around the world to deeply explore open education and its impact on global education, and thanks to continued efforts on the part of JSI-K4A, transLectures will be directly involved in its organisation.

The theme for this year’s conference is “Open Education for a Multicultural World”. More information regarding this event can be found at the [conference website](#). Call for papers is now open, and will close on 1 December 2013.

This event ties in with the aims of the European Commission’s recently-launched “Opening up Education” initiative to more fully exploit the potential of ICT and OERs, events related to which we have attended during this reporting period (34, 35).

### 3.3.3 poliMedia pilots

Taking advantage of the fact that poliMedia has been adopted by other Spanish universities, in M18 the UPVLC contacted the Faculty of Medicine at the Universitat Autònoma de Barcelona (UAB) and the Universidade de Vigo (UVIGO) about integrating transLectures tools into their poliMedia video repositories. In order to demonstrate the kind of quality they could expect, we ran some videos from each university through the transLectures transcription system and
sent back the resulting subtitles. On the whole, feedback was extremely positive from both universities and both were interested in getting more of their videos transcribed.

Further information about this stage in the process can be found in our two web entries: Feedback from First External Users [38] and New Feedback from External Users [36].

Since then, we have maintained a closer dialogue with UVIGO and will soon (all being well, in M25) embark on a technology transfer with this university. The exact nature of this transfer is yet to be decided, and may likely depart from the business model proposed in Section 4.2.1 of D7.4 [48]: the model outlined in D7.4 requires UVIGO to purchase equipment, set up the required infrastructure and produce manual transcriptions (in-domain data). Similarly, it requires the UPVLC to travel to Vigo (Spain) to implement and train the acoustic and language models.

The alternative we are considering, at least as an interim step, is to exploit the in-domain data and the models already available at the UPVLC and simply provide UVIGO with the transcriptions via the transLectures web service. UVIGO, then, would simply have to submit their videos via this web service, and the transcriptions would be generated using the UPVLC’s own computational resources. No cost is incurred to either party, except computationally, a cost which at this point the UPVLC is willing and able to assume.

This way, the UPVLC keeps hold of its models; they are not released under any contractual basis to UVIGO. We would also have more immediate access to the text data generated, as well as to the user interaction statistics (and configuration thereof) (see D6.3.1 [48]) logged by the web player. At least until the tools, interfaces and workflows are finalised, it seems a more sustainable approach to testing transLectures technologies in a real-life setting and getting feedback from real-life users. Pending approval and commitment of resources by the Universitat Politècnica de València, this model might also make the uptake of transLectures tools more appealing (economically/computationally viable) to a greater number of universities.

3.3.4 Fourth project meeting, with Advisory Board attendance

The fourth project meeting took place on 24-25 October at the XEROX headquarters in Grenoble, France. Project partners met to discuss progress in the different work packages, including the newly-instigated WP8: Demonstration.

As at the end of Y1, the three Advisory Board members were invited to this meeting, with Prof. Harald Höge and Meena Hwang ultimately attending. Both congratulated those present on the results obtained so far, commenting on the huge progress made over the past twelve months.

Höge participated actively throughout the day, making valuable suggestions across the different work packages and generally showing much enthusiasm at the results presented.

Hwang was particularly excited about the open nature of the project, and the huge commercial and exploitation potential that this will allow. She gave us some ideas as to the directions we might take, mostly for English-language transcription, even reaching out personally to her contacts and inviting them to take a look at the project. She stressed the importance of transcriptions in particular as the basic building block for any kind of derivative product of educational videos. The only limit, in her view, to the potential of transLectures was how far we are prepared to take it! This is something we will be discussing more seriously over the next six months (see below).

Finally, of course, we invite you to check out our web entry (photo included).
4 Evaluation

4.1 Overview

The simplest way to evaluate activities carried out in work package 7 would be to do so in relation to the objectives set for this reporting period.

4.1.1 Y2 Objective 1: to zoom in on user and “prosumer” communities

The main objective set for Y2 was to get closer to our potential user and “prosumer” communities, and if possible build a community around transLectures through event attendance and other dissemination activities.

Both the recommendations following the first year review and the Business Plan Draft (D7.4) were fundamental in this process. The first refocused our attentions on the immediate use scenario, the video lecture, and away from other potential uses in professional translation for example. While in D7.4 we ultimately defined two main applications of project technologies within this use scenario: one, to provide multilingual subtitles and two, to provide quality text data that can be used to power advanced repository management and other functions related to the exploitation of Big Data. As already discussed in Section 2.4.2, this led to a dual focus for this objective, zooming in on:

- Users of transLectures subtitles, the transcription and translation tools; namely, the education sector (events timeline 2, 3, 5, 6, 7, 20, (24), 25, 27, 28, (32), 33, 34, and 35);
- Users within the researcher/developer communities, and audiovisual management sector, interested in the underlying technologies and data (24, 25, 28, 32).

A second knock-on effect of drawing up D7.4 was a search for more events and other dissemination opportunities within the education sector. This is reflected in the nature of events attended throughout the year, particularly the ratio of academic versus industry-focused events, with the balance falling in favour of the latter. That said, we have not overlooked our academic stakeholders, attending a number of prestigious events in machine translation and speech technologies (4, 21, 30, 36 and 37), even organising a scientific challenge in partnership with EUCOGII and PASCAL2 (17).

Meanwhile, and as already discussed in Section 3.3.2 our relationship with the OCWC (identified in the DoW and, again, in D7.4 as a gateway to a large potential market for transLectures tools) has been further strengthened this year. This is thanks mainly to the outreach activities of JSI-K4A, who have also established a dialogue with UNESCO (12, 18) regarding the application of transLectures technologies in domains not contemplated in the DoW, to the benefit of the least developed and developing member states; as well as with universities in China (Tsingua University) and India (IIT Bombay) regarding the possible incorporation of new language pairs.

The actions carried out in pursuit of Objectives 2 and 3 also contribute to this objective.
4.1.2 Y2 Objective 2: to support and complement work carried out in WP6 on user evaluations

As already discussed in more detail in other sections of this report, during this reporting period the UPVLC released their TLK toolkit for automatic speech recognition. The response has been remarkable and has led to dialogue being established with some private universities in the United States regarding interest in implementation.

On a related note, the first release of the transLectures player is scheduled for the first few months in Y3.

Secondly, and also as already reported (see Section 3.3.3), the UPVLC has begun work on the first external pilots: the poliMedia pilots, which is a direct extension of the internal evaluations carried out as part of WP6. Feedback so far has been positive and both parties are keen to get started with the transcription.

4.1.3 Y2 Objective 3: to redesign and write new content for the project website

Our third objective was to redesign and rewrite the project website [19] to make it more accessible to non-scientific browsers. This was carried out at the end of 2012. We also committed to making more regular web entries in order to increase traffic, the time spent on the website and, by doing so, increase awareness and understanding of the project and related issues by a broader audience set.

Based on figures extracted from Google Analytics, total traffic during the Y2 reporting period has increased 362% on Y1: 7,976 visits compared to 1,727. Similarly, average visit duration has increased from 2 minutes 44 seconds, to 7 minutes 10 seconds, with browsers viewing an average of 5.5 pages per visit, up from 3.3. Further analysis reveals that the IASD challenge and the tools download section both account for significant traffic over the past twelve months. However, so too does the transLectures player used in the internal evaluation and quality control stages, which may be skewing figures.

Even so, these figures speak to the improved navigability, design and content of the new website.

We would also point out referrals from Facebook have increased from 1 to 556, which reflects our improved presence on this social network. LinkedIn and Twitter referrals have seen similar increases, from 0 in both cases to 126 (LinkedIn) and to 92 (Twitter). The former can be explained by our announcing the release of the TLK toolkit in various related LinkedIn groups.

Our Facebook page [20] itself has also seen improvements: we have kept a more active profile, with more photos and links to the website. At present count, we have 92 likes, of which 64 are since the revamped website and Facebook activity.

For the most part, our web entries consist of event write-ups, but we have made efforts over this reporting period to appeal to an audience interested though relatively uninformed regarding the broader context of video lectures, open education, translation technologies, etc. Titles such as “Happy International Translation Day!” and “The Video Lecture: an Overview”. Analytics PDFs illustrating the above findings can be found in Appendix A.

4.2 Goals set for Y3

Given the value of the reviewers’ comments following the first end-of-year review in guiding strategy, we feel it prudent to wait until we have received their recommendations based on Y2 results before setting specific goals. That said, the general direction work in this area should take is clear. We offer a preliminary discussion of our immediate aims and intentions below:
4.2.1 D7.5. Final Business Plan (M32)

A key priority in Y3 is to more precisely define each partner’s intentions beyond project end, whether it be exploitation in the form of future research or a more commercial approach. Given that the Final Business Plan is due in M32, this is going to be one of the main areas of work over the next six months.

A timeline was established at the meeting in Grenoble, beginning with a Skype conference in mid-November between the main authors of the deliverable. By the end of January we hope to have collected all the necessary information and final “declarations of intent” from the transLectures consortium members and start drafting what will most likely be individual business plans.

4.2.2 OCWC 2014

Similarly, given the importance of the OCWC global conference for transLectures, focus will also lie on planning, coordinating and optimising our presence, message and goals at this event, including at the dedicated transLectures workshop.

4.2.3 Publications

Being the final year of the project, this would be the year to see the joint publications we talked about in M18 come to fruition. Discussions in Grenoble in this regard were productive and two key areas were identified as possible subjects of joint papers between the UPVLC and RWTH: data selection techniques and the results obtained from collaborative experiments.

A tentative list of conferences for submission of these papers would include:

- *CoLing*, organised under the auspices of the International Committee on Computational Linguistics (ICCL)
- *ACL*, the Association for Computational Linguistic’s (ACL) annual meeting;
- *EACL*, the European chapter of the ACL’s meeting;
- *EAMT*, the European Association of Machine Translation’s meeting;
- *IWSLT*, the International Workshop on Spoken Language Translation;
- *EMNLP*, the Empirical Methods in Natural Language Processing conference;
- *InterSpeech*, international event in the field of speech communication sciences.

4.2.4 Events

Beyond the academic sphere, and as a continuation of the first objective set for Y2 to zoom in on the user community, key among the events being targeted for 2014 are:

- *EADTU 2014*, the European Association of Distance Teaching Universities’ annual conference (Cyprus, September 2014);
- *European Data Forum 2014*, (Athens, March 2014);
4.2.5 Other

Not forgetting of course continued work on the second release of transLectures-UPV TLK, the release of the transLectures player and the poliMedia pilots.

5 Conclusion

On the whole, the work carried out in WP7 during this reporting period has been strong. Dissemination activities have been proactive, varied, and impactful. We feel that people are really starting to see the value of what we are trying to achieve, and real interest has already been expressed from various sources in using transLectures tools. This leaves us in a very strong position at the start of the final year of the project.

Work in the first half of Y3 will be focused on drawing up a solid business plan, in which we hope to map the way forwards through to the end of the project and beyond in terms of exploitation and, hopefully, commercialisation of project technologies.
References


A Web analytics

Attached.
## Audience Overview

### All Visits

**Overview**

Nov 1, 2012 - Oct 31, 2013: 

- Visits

Nov 2, 2011 - Oct 31, 2012: 

- Visits

![Graph showing visits over time](image)

**2,928 people visited this site**

- **Visits**: 362.13% increase from 7,381 to 12,013
- **Unique Visitors**: 166.91% increase from 2,928 to 5,097
- **Pageviews**: 676.85% increase from 43,752 to 306,532
- **Pages / Visit**: 68.10% increase from 5.48 to 8.36
- **Avg. Visit Duration**: 61.36% increase from 00:07:10 to 00:11:54
- **Bounce Rate**: -30.81% decrease from 38.54% to 26.63%
- **% New Visits**: -42.53% decrease from 36.37% to 20.84%

**Returning Visitor** vs **New Visitor**

### Language

<table>
<thead>
<tr>
<th>Language</th>
<th>Visits</th>
<th>% Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. en-US</td>
<td>Nov 1, 2012 - Oct 31, 2013</td>
<td>2,571</td>
</tr>
<tr>
<td></td>
<td>Nov 2, 2011 - Oct 31, 2012</td>
<td>796</td>
</tr>
<tr>
<td></td>
<td>% Change</td>
<td>222.99%</td>
</tr>
<tr>
<td>2. es</td>
<td>Nov 1, 2012 - Oct 31, 2013</td>
<td>2,064</td>
</tr>
<tr>
<td></td>
<td>Nov 2, 2011 - Oct 31, 2012</td>
<td>143</td>
</tr>
<tr>
<td></td>
<td>% Change</td>
<td>1,343.36%</td>
</tr>
</tbody>
</table>