



Deliverable Task 6

ELRC Workshop Report for the Netherlands



Author(s): Jan Odijk (Universiteit Utrecht)
Carole Tiberius (Instituut voor de Nederlandse Taal)

Dissemination Level: Confidential

Version No.: <V1.1>

Date: 2016-06-13



Contents

1	<u>Executive Summary</u>	3
2	<u>Workshop Agenda</u>	4
3	<u>Summary of Content of Sessions</u>	5
3.1	Session 1: “Opening, Welcome and Objectives”	5
3.2	Europe and Multilingualism	6
3.3	Languages and Language Technologies/Translation in the Netherlands	6
3.4	Panel: Multilingual Public Services in the Netherlands	7
3.5	Automated Translation: How does it work?	7
3.6	How can Public Institutions benefit from the CEF.AT Platform?	8
3.7	What Data is needed?	8
3.8	Legal framework for Contributing Data and the European Data Portal	9
3.9	Data and Language Resources in the Netherlands	9
3.10	Data and Language resources: Best Practices to Capitalize on your Valuable Data	10
3.11	How can we engage?	11
3.12	Wrap-up, on site Conclusions and Commitments	11
4	<u>Synthesis of Workshop Discussions</u>	13
5	<u>Workshop Presentations Material</u>	14
6	<u>Appendix</u>	15
6.1	Agenda in Dutch	15

ELRC Workshop Report for the Netherlands

1 Executive Summary

This document reports on the ELRC Workshop in the Netherlands, which took place in The Hague on 19 April 2016 at the European Commission in the Netherlands, Korte Vijverberg 5-6, 2513 AB The Hague. It includes the agenda of the event (section 2) and briefly informs about the content of each individual, interactive and panel workshop session (sections 3 & 4). The event was attended by 41 participants from the Dutch public sector, mainly from ministries and government offices, as well as independent data experts and freelance translators. The list of participants is included in the document. The dedicated event webpage can be found at <http://lr-coordination.eu/netherlands>. There one can also find the agenda and the presentations that were used during the workshop.

2 Workshop Agenda

Here we give the Agenda with English Titles. See the [website agenda](#) and annex 6.1 for the agenda with Dutch Titles.

9:00 – 09:30	Registration	
09:30 – 10:00	Opening, Welcome and Objectives	Steven Krauwer (CLARIN ERIC), Hugo Keizer (EC, Vert. in NL), Jan Odijk (UU)
10:00 – 10:30	Europe and Multilingualism	Hugo Keizer (EC, Vert. in NL)
10:30 – 11:00	Languages and Translation in the Netherlands	Kevin De Coninck (NTU); Michel Verhagen (Min BuZa)
11:00 – 11:30	Coffee break and networking	
11:30 – 12:00	Panel: Multilingual Public Services in the Netherlands	Paul Suijkerbuijk (Min BZK, ICTU/Open Data); Edwin de Koning (Min Justitie en Veiligheid)
12:00 – 12:30	Automated Translation: How does it work?	Véronique Hoste (Univ Gent; ELRC-organisator België)
12:30 – 13:00	How can Public institutions benefit from the CEF.AT Platform	Spyridon Pilos (DGT)
13:00 – 14:00	Lunch Break	
14:00 – 14:30	What Data is needed?	Carole Tiberius (INT)
14:30 – 15:00	Legal Framework and the European Data Portal	Khalid Choukri (ELDA, ELRC); Hayo Schreijer (KOOP)
15:00 – 15:30	Panel: Data and Language Resources in the Netherlands	Martin Everaert (INT), Kevin De Coninck (NTU), Hayo Schreijer (KOOP)
15:30 – 16:00	Coffee break and networking	
16:00 – 16:30	Data and Language Resources: Practical Aspects	Khalid Choukri (ELDA, ELRC)
16:30 – 17:00	How can we engage?	Henk-Jan Eras (Tweede Kamer); Michel Verhagen (Min BuZa)
17:00 – 17:15	Wrap-up, On site Conclusions and Commitments	Steven Krauwer (CLARIN ERIC), Jan Odijk (UU)
17:15 – 18:00	Drinks	

3 Summary of Content of Sessions

3.1 Session 1: “Opening, Welcome and Objectives”

Steven Krauwer (CLARIN ERIC) opened the event by welcoming the audience and introducing the key persons in conceiving and organizing the event, namely the ELRC consortium and the EC/DGT representatives.

The welcome on behalf of the EC/DGT was expressed by Hugo Keizer (European Commission Representation in the Netherlands).

Jan Odijk (Utrecht University), the ELRC technical National Anchor Point in the Netherlands briefly described the purpose of this event. Multilingualism is a great asset of Europe, but it also creates barriers. These barriers can be overcome by means of translation, but there is too much to be translated to do it with human translators only. Support by technology is necessary. CEF.AT aims to offer this technological support. It will show good performance on text types that it is trained on, but will perform less well on other text types. Therefore, public services can benefit optimally from CEF.AT if they make the data (multilingual and monolingual texts, translation memories, dictionaries, etc.) that they use available for training of CEF.AT. The goal of this workshop is to explain this in more detail, to hear from the public service representatives whether they use / want to use EC translation services such as [MT@EC](#) currently and CEF.AT in the near future, and to make a start with getting data available.

The Netherlands already has a technical National Anchor Point (Jan Odijk) and hopes to identify its Public Services National Anchor Point in this workshop.

[\[presentation\]](#)

Questions from the audience:

Q: You spoke about MT@EC being situated in a secure environment, but that it is also accessible to all citizens. Isn't that contradictory?

A: The matter of security will come back in more detail later. However, it is not the case that citizens can use MT@EC, only governmental organisations can. Furthermore, all private information or information that is in other ways sensitive (e.g. because of ethical or legal reasons) must be excluded.

Q: Who can use CEF.AT?

A: Only governmental organisations can use CEF.AT. It is not accessible to individual citizens or commercial language service providers.

Q: What is the relation to Google Translate?

A: CEF.AT is an independent system, in a secure environment. It can produce context-dependent translations, tuned to the targeted users, which will make it yield better translations than is currently possible.

Q: For many languages the number of language pair combinations explodes. How are you going to make sure you will have enough material for the smaller languages?

A: CEF.AT is dependent on the data that will be delivered. Therefore workshops like this one are organized in all EU countries.

Q: Is CEF.AT only interested in data from governments? Texts from web shops might also be interesting.

A: That is true, but the focus is currently on governmental organisations and their data.

ELRC Workshop Report for the Netherlands

3.2 Europe and Multilingualism

Hugo Keizer (European Commission, Representation in the Netherlands) explained the multilingual situation in Europe and the EU policies towards multilingualism, especially in connection with the Digital Single Market.

[\[presentation\]](#)

3.3 Languages and Language Technologies/Translation in the Netherlands

Languages in the Netherlands

Kevin De Coninck (Dutch Language Union) explained the assignment to the NTU (Dutch Language Union, DLU) and the language situation of the Dutch language, inter alia in the Netherlands. He also sketched the upcoming changes in the INL (Institute for Dutch Lexicology), which is going to be turned into the INT (Institute for the Dutch Language). The HLT Agency, which currently resides under the DLU, will be split into two parts: the language resources will be managed by INT, the ‘makel-en-schakel’ (networking, liaising) role will be continued by DLU.

Kevin also gave a sketch of language resources available through the HLT Agency. In particular, he pointed out the *Woordenlijst van de Nederlandse Taal* / Word List of the Dutch Language (<http://woordenlijst.org>), which is the authority for proper spelling, the [ANS](#) grammar of Dutch, but also multilingual resources such as the [OMBI](#) (Omkeerbare Bilinguale Bestanden / Reversible Bilingual Resources) translation dictionaries for Dutch-Arabic, Dutch-Indonesian, and Dutch-Danish, and the [Dutch Parallel Corpus](#). All of these are available.

Translation in the Netherlands

Michel Verhagen (Ministry of Foreign Affairs, translation service) explained that most of the translation work is not done internally in the ministries but given out to external language service providers. Only approx. 5% of the material, the one considered to have a ‘hoog afbreukrisico’, i.e. a high-risk for image damage or to be confidential, is translated internally at the ministries. Most ministries deal with outsourcing translations themselves.

There are also translation departments at the “Talencentrum Defensie (TCD)” ([Ministry of Defense](#) Language Centre), at the [Sociale Verzekeringsbank](#) (Social Security Bank), at the [National Police](#), and at the ‘Mededingingssautoriteit’ nowadays called the [Autoriteit Consument en Markt](#) (ACM) (Authority Consumer and Market), which is part of the Ministry of Economic Affairs.

Documents to be translated range widely in character, from dinner menus to treaties, speeches and diplomatic correspondence. SDL Trados Studio 2014 and Multiterm 2014 are used as well as a full text search engine (textstat). Within the Netherlands there are sometimes big differences in nomenclature and terminology, especially in parts of former colonies that have recently become part of the Netherlands (Antilles).

They use tools from other organisations such as [NTU termtreffer](#) and the [TBX standard](#) for terminology management. They use [MT@EC](#) for gisting, in combination with post-editing and for translation. MT@EC is not suited for generating high quality translations. It can be used in combination with post/editing but it is quite static.

Laws are published and searchable on <http://wetten.overheid.nl/zoeken>, speeches at <https://www.government.nl/>. Michel manages one big translation memory, but that contains confidential and/or copyrighted information, and cannot be provided to others.

The source texts are the property of the customers, as are the translations. Though the intellectual property on the translations resides with the translators, they cannot decide to make data available to external parties.

ELRC Workshop Report for the Netherlands

3.4 Panel: Multilingual Public Services in the Netherlands

Paul Suijkerbuijk (Ministry of Internal and Kingdom Affairs, dept. ICTU/Open data).

Paul Suijkerbuijk showed that Open Data are important because they create societal, economic, systemic and democratic value. He sketched more characteristics of Open Data, in particular for open public governmental data, which contain public governmental information and on which there are no restrictions on reusability: free, available “as-is”, free of rights, accessible without registration, computer-processable. The law of “Openbaarheid van Bestuur” already holds longer, generating open data, but the EU PSI directive on reusability has led to new legislation in the Netherlands (Wet Open Overheid, accepted in July 2015), which makes reusability of data much easier.

[\[presentation\]](#)

Edwin de Koning (Ministry of Safety and Justice)

Edwin de Koning stated that his department buys interpreting and translation services from third parties, with a clear focus on interpreting.

There is no clear common policy of translation departments within the ministries. The departments own their procedures and want to keep it that way, leading to fragmentation. There may be indications for changes in the right direction, but these still go slowly. Many translations are made by external parties. There is a growing demand for MT.

Many different types of documents are translated, e.g. web sites, regulations, ministerial decisions (in many different formats), as well as pictures or bad Xerox copies which are often very difficult to decipher for humans let alone machines.

European languages are dominant, but translation to or from languages spoken in war zones (currently e.g. languages from Syria) are also frequent.

The information provided at this workshop should be given at all ministries, so that each ministry can decide how it can benefit and how it can contribute by providing its data.

Discussion

Paul believes that the new regulations will make our lives easier. Reusability of data is made easier, data are open and are becoming available, and many new applications can be built around these open data. It was suggested that Open data can also help reduce the fragmentation that we observe today.

Jan Odijk sees a role for Paul Suijkerbuijk and his team to counter the fragmentation by educating the civil servants at the various ministries so that they will make available their data and to teach them how they should do this, and perhaps also to provide facilities, e.g. for the easy creation of appropriate metadata. Paul confirms that data.overheid.nl informs the ministries about such matters, but the quality of the reference to the data is crucial. Open data come from the organizations and they work in specific ways, which may not be optimal for being able to find data back on the basis of the metadata provided. It was suggested from the floor that there should be some standardization of metadata, which Paul confirmed. However, it was decided in an earlier stage to first set up a repository of links and not wait with that until standardization of the metadata was agreed upon. But with more needs from the society for finding the relevant data, standardization of metadata is becoming more urgent.

3.5 Automated Translation: How does it work?

Véronique Hoste (Ghent University) explained how statistical machine translation works. From this presentation it became fully clear why data are so important for statistical machine translation. A statistical machine translation system contains a translation model and a

ELRC Workshop Report for the Netherlands

language model for the target language. It derives its translation model from aligned multilingual data, and it derives the language model from monolingual data in the target language. These models work with probabilities that are estimated on the basis of frequencies: the more data, the higher the frequencies and the better the estimations of the probabilities.

The language model in a statistical machine translation system is a very simple model based on so-called n -grams, i.e. sequences of n consecutive words (where n usually is at most 4 and often lower). This works reasonably well in many cases but problems arise if there are dependencies between two words over longer distances (which happen to occur quite frequently in natural language) such as the dependency in Dutch between a verb and its separable particle (e.g. *Hij **had** het boek pas na enkele dagen **uit*** (Google Translate: “*He had the book until a few days off*”, correct: “*He finished reading the book only after a few days*”).

[\[presentation\]](#)

3.6 How can Public Institutions benefit from the CEF.AT Platform?

Spyridon Pilos (DGT) explained via Skype how CEF.AT works and how public institutions can benefit from it. CEF.AT will be able to handle a wide variety of source document formats but not PDF.

CEF.AT is available for government organisations but not for translation companies and translation schools. Translation companies might still benefit if the ministry applies CEF.AT and provides the CEF.AT-generated translation, but this model has many disadvantages as well. The only institutions that it are currently available to are the ones that offer a translation master.

EC@MT which is currently in use will slowly grow into CEF.AT over time. CEF.AT will provide a faster system of better quality, and will run in a secure environment (login through ECAS) with secure document transfer. CEF.AT will have additional services as well.

[\[presentation\]](#)

3.7 What Data is needed?

Carole Tiberius explained what types of data are needed and can contribute to improvements of CEF.AT. This includes any data that contains “words”, even better is data that contains “sentences”, and still better is multilingual data with translations aligned at the sentence or paragraph level (parallel corpora) or about the same topics (comparable corpora). In addition, dictionaries, terminology banks, ontologies, and translation memories are highly useful.

Ideally the data provided are properly documented with formally structured metadata.

The data that are relevant ideally reflect the typical language use (jargon, terminology, constructions) in the public sector. This can be reflected in actual texts, dictionaries, term banks, translation memories, etc. Both data from a single language and data from multiple languages are sought for, and ideally texts that are translations of each other. The quality of these translations is less important, provided that they are human translations (not machine translation). The statistical techniques used by the machine translation systems are quite robust and will not be disturbed by translations that are considered of somewhat lower quality by humans’ standards.

Spoken data themselves are not interesting in this context, since CEF.AT focuses on written data. But written out versions of speeches, debates, etc. surely are interesting.

[\[presentation\]](#)

ELRC Workshop Report for the Netherlands

3.8 Legal framework for Contributing Data and the European Data Portal

Legal Framework

Khalid Choukri (ELRC, ELDA) sketched some aspects of the legal framework. In particular, on the reuse of public sector data (Dutch: Wet Hergebruik Overheidsdata). This makes public sector data and information available to third parties for reuse with a minimum of restrictions.

In the context of machine translation, provided texts will be used to develop translation and language models but not republished as such. The original texts cannot be reconstructed from the models derived from them.

Khalid further sketched 5 steps to be taken to release data that one has available. These steps ensure that any data that cannot be released for legal or ethical reasons is excluded or that acceptable (e.g. anonymized, reduced) versions of them are created before release.

According to Khalid, the Netherlands appears to have a policy that states that public sector data are open by default, and that specific justification has to be provided if such data are not to be treated as open data.

[\[presentation\]](#)

European Data Portal

Hayo Schreijer is from [KOOP](#) (Kennis- en Exploitatiecentrum Officiële Overheidspublicaties / Knowledge and Expertise Centre for Official Governmental Publications), which makes available (pointers to) public sector data via dedicated portals such as <http://www.overheid.nl> for governmental information, <http://www.dataoverheid.nl> as the national data portal, etc. These portals usually do not contain the data themselves, but only pointers (links, URLs) to the data, which reside elsewhere.

Many of the data are structured tables, though there are also datasets with text.

The portal <http://www.overheid.nl> contains a lot of links to laws, regulations, decisions, to the daily published official announcements from the 'Staatscourant', the 'Staatsblad' and the 'Tractatenblad', as well as all laws, regulations and decisions that were published earlier.

KOOP also makes recommendation on licenses. The Netherlands prefers CC-0 licenses.

The European Data Portal is an aggregation portal at the European level, which reads from all national data portals such as <http://www.overheid.nl>, <http://wetten.overheid.nl> and <http://www.dataoverheid.nl>.

KOOP takes requests for data seriously and takes them into account in prioritizing its activities. In the recent past, KOOP could point out directly where the data are, or they started up actions to make the relevant data available. Only 10% of the recent requests were turned down.

KOOP also tries to ensure the quality of the data and the metadata, though the metadata are made according to a limited scheme and therefore searching for data is also limited.

3.9 Data and Language Resources in the Netherlands

In this panel **Martin Everaert** (INL/INT) explained what data INL, which is transforming into INT, has available, and that INT serves a much wider target group than the group that the ELRC workshop now focuses on. He also emphasized that storing the data in a repository, and making them visible and accessible to researchers does not come for free and needs financing.

ELRC Workshop Report for the Netherlands

Kevin De Coninck (NTU) repeated his offer of data that he made earlier.

Hayo Schreijer (KOOP) emphasized that the Public Sector Information (PSI) directive was formulated as it is on purpose: making data open really means making data open: no costs should be involved in making data that must be open accessible in a later stage. All new public sector data produced must be visible and accessible from the start.

He also emphasized that KOOP may also benefit from an initiative such as ELRC, since it uses quite some language technology as well, in particular to get to search by concepts rather than by keywords, and for extracting links to other data. Kevin De Coninck also pointed out the Spelling API that NTU offers.

Digitalisation of the full [Staten Generaal Digitaal](#) will be finished by the end of 2016.

Q: Aren't you providing more advantages to big companies, which have means to invest, and isn't there a danger that small companies will suffer?

A: KOOP makes sure that the basic information is available to all parties, whether they are small or large. The market can then add services operating on these data and other data on top of that.

Q: Many ministries have data such as texts and their translations. Can they publish them via KOOP?

A: Ministries publish under their own mandate, but they can do it via KOOP.

Q: Are there facilities for web services in KOOP (i.e. for access to the data via software programs)?

A: These exist, but not for all datasets. APIs (Application Programmers Interfaces) are not required by law, but in some cases they exist. The focus has always been on priding the raw data.

Q: Will any data from the commission on terminology be made available?

A: The EU wants data for improving their MT. But CLARIN, INT, ELRA, etc. can make available all these data for research. Hugo Keizer adds that DGT NL strongly focuses on terminology. The output generated by this workshop will find its way to IATE etc.

Khalid Choukri urged everybody to help ELRC get the data in a usable format, not only the things that are accessible via the web, but also data 'hidden in the deep web', i.e. data which are not openly accessible through the visible websites.

[\[presentation\]](#)

3.10 Data and Language resources: Best Practices to Capitalize on your Valuable Data

Khalid Choukri (ELRC, ELDA) points out how important it is to make sure that the data that you have are reused. It will improve the machine translation systems. And with your data it will improve your translations. There are many practical aspects that must be taken into account when you make data available, and ELRC can provide assistance, for example with anonymizing sensitive data. That is very important for the Sociale Verzekeringsbank (Social Security Bank), which has many data in many languages that they are willing to make available provided that the data can be anonymized reliably.

[\[presentation\]](#)

ELRC Workshop Report for the Netherlands

3.11 How can we engage?

Henk-Jan Eras from the Reporting and Editing service of the Dutch Parliament makes data (transactions of the parliament, commissions, etc.) available and all these data are open. The data are on the sites of the “Eerste Kamer” and the “Tweede Kamer”, and are also accessible via KOOP. However, different locations have different versions of the data. The Parliament has uncorrected data, KOOP publishes the corrected data. Furthermore, KOOP publishes the data per activity, while the Parliament publishes the data per day. In addition, KOOP has an archiving function, the Reporting and Editing service has no such function.

The Parliament is developing a data store (‘datamagazijn’), which is accessible to external parties, and INL/INT already uses it. If other parties are interested in participating, a similar cooperation as with INT can be set up.

Most of the parliamentary texts are in Dutch, but occasionally reports are made in different languages, e.g. recently an English report was made on the debates in connection with the MH17 disaster.

Michel Verhagen explained that only a small part of the translations they produce can be made available. However, the versions at <http://wetten.overheid.nl> may differ from their own versions, because after the translation has been made, the results are checked by a legal expert.

The two contributions in this panel illustrate two extremes: on the one hand, an organization that makes public data in large quantities, which they can easily make available; on the other hand, a service that has no control over the data that it deals with, many of which are dealt with by this service because they are or can be sensitive/confidential; Other data are distributed over many different ministries.

It is evident that there is a clear willingness to make data available, but there are many obstacles, both of a practical and of a legal nature, to actually do this.

Is there a way to streamline the process of making produced public data and their translations available, e.g. via KOOP?

The importance of good metadata is also stressed during this panel. Rich metadata of high quality are a prerequisite for making data findable.

The National Archive makes open data available, they are accessible via an API, in particular material from approx. 1950 that has been digitized. There is only little born-digital material. Data must move from the ministries to the National Archive after 10 years. The National Archive also supports ministries in deciding on making data public, for example by automatic classification so that parts of certain documents can be made public. The National Archive has data in multiple languages, but no translations.

The National Archive offers services to governmental organisations but is not the owner of the data. The Regional historical centres reside under the National Archive, but each municipality has its own archive.

3.12 Wrap-up, on site Conclusions and Commitments

Jan Odijk summarized his major findings of the workshop. The public sector that we tried to reach was outside of the organizers’ networks, so it was not easy to find the relevant persons and to understand what role they play. Organising this workshop has therefore been difficult but by organizing it a lot more clarity on who the relevant players are has been created.

Jan Odijk sees a lot of interest in the workshop from many different parties. There is a clear willingness to share data, especially since people see that they may benefit from sharing their data, but there are still many practical and legal obstacles. The owners of the data are

ELRC Workshop Report for the Netherlands

distributed over all ministries, which makes getting these data difficult. The situation is even worse at the municipal level.

It would be ideal to set up a system through which there is a continuous influx of public data from the public sector into CEF.AT, but many issues have to be resolved before that can be created, inter alia:

- Distributed production of texts
- Translation services are distributed
- Confidentiality of certain data has to be dealt with
- The format and the quality of the data en associated metadata

Jan Odijk expresses the hope that KOOP can assist to resolve at least some of these issues, e.g. by providing a single access point from which data can be 'harvested' by CEF.AT.

Another goal of the workshop was to identify a public National Anchor Point in addition to the technical National Anchor Point (Jan Odijk), and the organisers have identified several candidates that could fulfil this role.

The workshop is just the beginning. The next steps in the process are:

- end of April: analysis of your feedback and a report on the workshop
- mid May: identification of potential data providers
- end of May: contacting potential data providers

Workshops of this type were organized in each EU-country, plus Norway and Iceland. The experiences gained will be shared by all workshop organisers in a meeting in the beginning of July.

Steven Krauwer reminds everybody to fill out their engagement form, invites all for the drinks, and closes the workshop.

[\[presentation\]](#)

4 Synthesis of Workshop Discussions

The results of the discussions have been included in the preceding section, since there they are in their proper context. In some cases the questions and answers have been written down more or less verbatim, on other cases discussions have been briefly summarized in one sentence or paragraph.



5 Workshop Presentations Material

The presentations are on the [ELRC Nederland website agenda page](#) and the current report has links to these presentations in each relevant subsection in section 3.

6 Appendix

6.1 Agenda in Dutch

9:00 – 09:30	Registration	
09:30 – 10:00	Opening en Welkom, Doelstellingen	Steven Krauwer (CLARIN ERIC), Hugo Keizer (EC, Vert. in NL), Jan Odijk (UU)
10:00 – 10:30	Europa en Meertaligheid	Hugo Keizer (EC, Vert. in NL)
10:30 – 11:00	Talen en Vertalen in Nederland	Kevin De Coninck (NTU); Michel Verhagen (Min BuZa)
11:00 – 11:30	Coffee break and networking	
11:30 – 12:00	Panel: Meertalige Openbare Diensten in Nederland	Paul Suijkerbuijk (Min BZK, ICTU/Open Data); Edwin de Koning (Min Justitie en Veiligheid)
12:00 – 12:30	Geautomatiseerde vertaling: Hoe werkt het?	Véronique Hoste (Univ Gent; ELRC-organisator België)
12:30 – 13:00	Overheidsinstellingen en CEF.AT	Spyridon Pilos (DGT)
13:00 – 14:00	Lunch Break	
14:00 – 14:30	Wat voor Data zijn Nodig?	Carole Tiberius (INT)
14:30 – 15:00	Wettelijk kader en het Europese DataPortaal	Khalid Choukri (ELDA, ELRC); Hayo Schreijer (KOOP)
15:00 – 15:30	Panel: Data- en Taalbronnen in Nederland	Martin Everaert (INT), Kevin De Coninck (NTU), Hayo Schreijer (KOOP)
15:30 – 16:00	Coffee break and networking	
16:00 – 16:30	Data- en Taalbronnen: Praktische Aspecten	Khalid Choukri (ELDA, ELRC)
16:30 – 17:00	Hoe Kunt u Bijdragen?	Henk-Jan Eras (Tweede Kamer); Michel Verhagen (Min BuZa)
17:00 – 17:15	Afronding, Conclusies en Afspraken	Steven Krauwer (CLARIN ERIC), Jan Odijk (UU)
17:15 – 18:00	Drinks	