

Terminology: Don't only collect it, use it!

Vesna Lušicky & Tanja Wissik
University of Vienna

Translating and the Computer Conference

London, 29 & 30 November 2012

Overview

- ▶ LISE project
- ▶ Workflow research
- ▶ Highlights – needs analysis
- ▶ LISE Service Version Two
- ▶ Demo
- ▶ Outlook

The LISE project

- ▶ EU funded project within the ICT–PSP programme
- ▶ Consortium: academic, industrial and institutional partners
- ▶ Improve the quality of large (inter–institutional or inter–departmental) terminological resources by efficiently cleaning them in a collaborative online environment
- ▶ Efficiently spot inconsistencies and gaps
- ▶ Collaborate with peers from other institutions
- ▶ Benefit from proven best practices
- ▶ Achieve interoperability



REPUBLIK ÖSTERREICH
Parlament

Two directions of the project

- ▶ Workflow research and best practices in (inter-institutional) terminology work
- ▶ Development of a terminology service platform to support inter-institutional terminology work and to (semi)automate processes



User groups and evaluation

▶ User groups:

- IATE
- EU Institutions involved in IATE

European Commission, Parliament, Council, Court of Justice. Court of Auditors, Economic & Social Committee, Committee of the Regions, European Central Bank, European Investment Bank, Translation Centre for the Bodies of the EU

- Austrian Parliament

▶ Evaluation phase



REPUBLIK ÖSTERREICH
Parlament



Workflow research

▶ Interviews

- Interview guideline
- 17 conducted
- Recorded–transcribed–analysed

▶ Survey

- On–line
- Available until end of March 2012
- Disseminated via relevant networks
- 22 completed

Workflow research

Interviewed institutions

- ▶ Auswärtiges Amt (Ministry of Foreign Affairs, Germany)
- ▶ Translation Centre for the Bodies of the EU
- ▶ DGA3 Terminology and Documentation, General Secretariat, EU Council
- ▶ DG TRAD – Terminology Coordination at the European Parliament
- ▶ Court of Justice of the European Union, Terminology section
- ▶ DG Translation – D.03 – Terminology Coordination at the European Commission)
- ▶ Sektion Terminologie der Schweizerischen Bundeskanzlei
- ▶ Staatskanzlei des Kantons Bern – Zentraler Terminologiedienst
- ▶ Translation Workflow and Terminology Officer – Food and Agriculture Organization of the United Nations
- ▶ Ausschuss der Deutschsprachigen Gemeinschaft für die deutsche Rechtsterminologie – Ministerium der Deutschsprachigen Gemeinschaft Belgiens
- ▶ An Coiste Téarmaíochta – Foras na Gaeilge
- ▶ Terminologocentrum TNC
- ▶ Terminology Standardization Directorate/Translation Bureau, Canada
- ▶ TERMCAT
- ▶ Österreichische Nationalbank (Austrian National Bank)
- ▶ Ministerie van Buitenlandse Zaken (Ministry of Foreign Affairs, Netherlands)

Highlights-needs analysis

Maintaining TB	Harmonising and merging TB	Enriching TB	Review with peers – collaborative work
batch export - review – batch import	harmonising	monolingual entries in multilingual TB	communication via e-mails
discrepancy: working method of consolidation vs. workload	merging	crosscheck between TM and TB	communication via phone
redundant entries, redundant/missing information		language specific, subject-oriented projects	informal communication
basic collaborative work			formal communication
manual correction of misspellings			collaboration tool
			extranet

TB = Termbase
TM = Translation memory

Maintaining a termbase

“If we go about cleaning up in about 20 to 30 terms at the time, it will take about 100 years before we get anywhere. We would like to clean up, but frankly, I think, we need some assistance. Currently, with the cleaning up operation, I don't think we can go very far with that.” (INT6)

“...this is one side and cleaning is the other side. I mean, content-wise the database is well structured, but needs cleaning because we have lots of duplications, lots of things that are not necessary and others we need and are not there yet.” (INT7)

Harmonising and merging

“...of course there is a huge need for coordination and harmonisation of what is done. Especially with regard to feeding the database... and updating the database.” (INT9)

(as translated by V.L., T.W.; original: „natürlich großen Bedarf an Koordinierung und Harmonisierung dessen, was gemacht wird. Insbesondere im Hinblick auf die Einspeisung in die Datenbank und ... die Aktualisierung der Datenbank“ (INT9))

“And when it is larger and all that and if it involves mappings, scripts and things like that, this has to be done on a... Sometimes we would deal with that in batches, depending on the size of the project..” (INT17)

Enriching with new languages

“If you look at the figures for the monolingual, bilingual and trilingual entries, more than half of the database is made of monolingual and bilingual entries, which probably shouldn't be there, which are probably quite useless.” (INT6)

“It would be interesting, would be a comparison of the term candidates from the translation memories and a comparison of what is in there [the termbank].” (INT1)

(as translated by VL;TW; original „Was interessant wäre, einen Abgleich zwischen wirklich, was Termkandidaten aus der Translation Memory und Abgleich mit [der Termbank], was schon drinnen ist.“ (INT1))

Review with peers – collaborative work

“What we maybe lack, is this communication..., this platform, to communicate. ” (INT5)

(as translated by VL;TW; original „Was uns vielleicht fehlt momentan, ist so dieses Kommunikations..., diese Plattform, so ...zu kommunizieren.“(INT5))

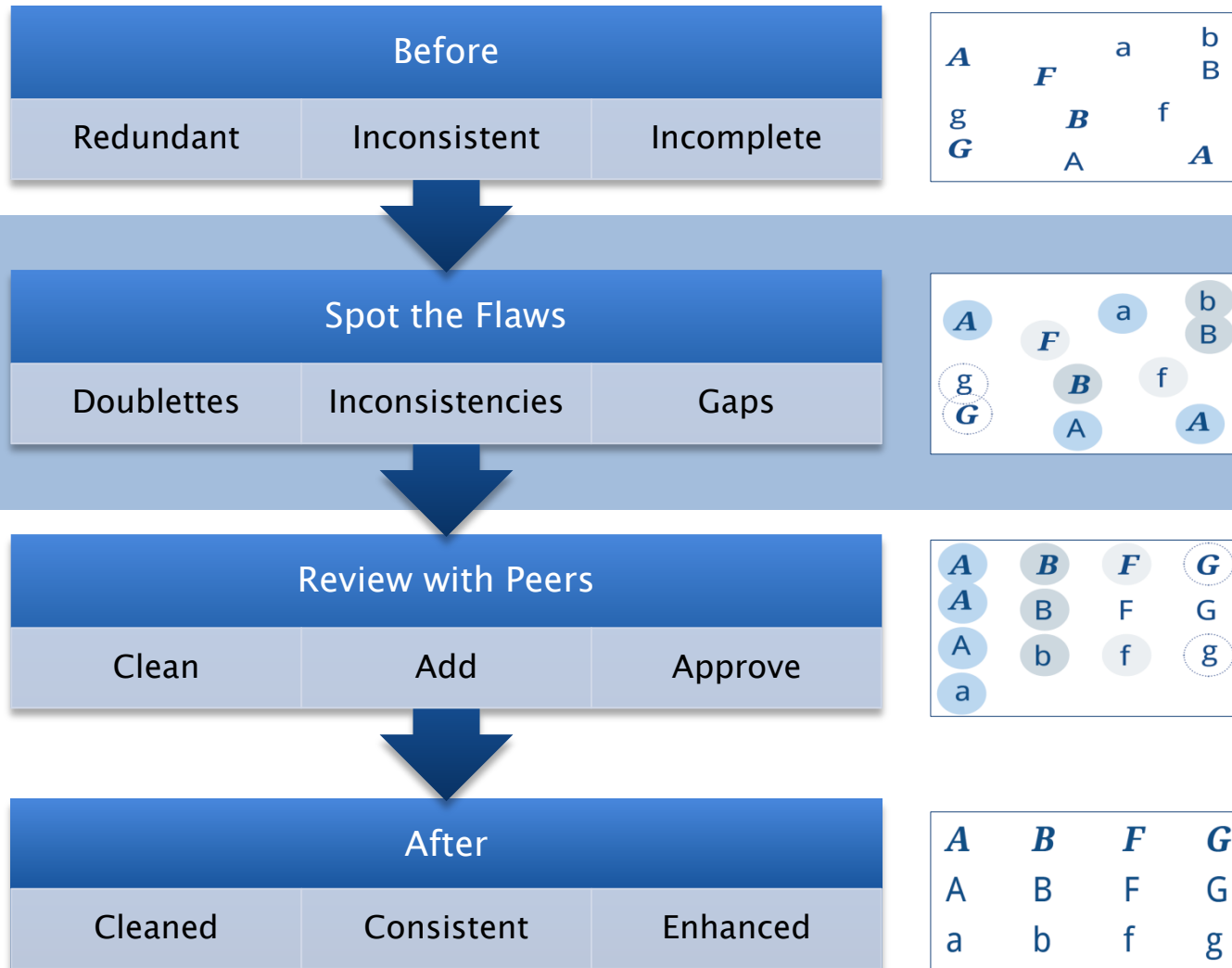
“They send us an e-mail or come and talk to us. We are in the same building.” (INT6)

“And then they send us the list, “Could you please delete or merge your data on to ours, so that we could eliminate a couple of duplicates in this area.” (INT7)

LISE Service Version Two

- ▶ LISE Collaboration Portal as a central hub:
 - Access software
 - Discuss issues
 - Exchange data
- ▶ Fully featured collaboration functionality
- ▶ Semi-automatic collaborative processing – the terminologist has the last word!

The LISE Workflow



Maintaining a termbase: Cleanup

- ▶ Misspellings
- ▶ Missing domain assignments
- ▶ Equivalents („doublettes“)
- ▶ Subsets (redundant concepts)
- ▶ Translation errors
- ▶ Wrong language assignments

ESTeam Cleanup@ User UI

File Settings Help

Pending

- Misspelling
- Canonisation
- Language
- Equivalent
- Domain
- Translation
- No English
- Subset
- Completed

EntryId	Domain
843027	FINANCE..Insurance
865785	FINANCE..Insurance
892966	FINANCE..Insurance
1129947	FINANCE..Insurance
1355673	FINANCE..Insurance
1550111	FINANCE..Insurance
1550135	FINANCE..Insurance
1550193	FINANCE..Insurance
765491	FINANCE..Insurance..Insurance
786034	FINANCE..Insurance..Insurance
790806	FINANCE..Insurance..Insurance
881163	FINANCE..Insurance..Insurance
1391293	FINANCE..Insurance+SOCIAL QUESTIONS..Social protection..Soc...

Total 56

Search Misspelling Advanced Search

Domain	Reference	Actions
Domain:FINA...	Reference - en	
en	investment income and charges	
Domain:FINA...	Reference - fr	
fr	produits et charges des placements	
EntryId:843027		
Domain:FINA...		
de	Ertäge aus bzw. Aufwendungen für Kapitalanlagen	✗ ✎ ✓
Misspelling	Ertäge - Use Right Click to select a potential Spell Suggestion	✗ ✎ ✓
en	investment income and charges	
es	ingresos y gastos de inversiones	
fr	produits et charges des placements	
it	proventi da ed oneri per investimenti	


Cleanup – spelling error

German spelling error „Ertäge“ – should be „Erträge“

Internal harmonising and merging: OMEEO

- ▶ Identical concepts should no longer be stored separately
- ▶ Identical terms in different entries are detected
- ▶ Term variants for one concept are discovered in order to be merged
- ▶ Results: a more complete collection of terminology, which can be harmonised semi-automatically
- ▶ Users can accept or reject the suggestions

IATE - Search Result



English (en) [Search Screen](#) [Help](#)

Haushaltszulage

de > en (domain: Any domain, type of search: All)

Result 1- 6 of 6 for **Haushaltszulage**

Pay policy [CdI] Full entry	
DE Haushaltszulage	★★★★ +@
EN household allowance	★★★★ +@
FINANCE, SOCIAL QUESTIONS, EMPLOYMENT AND WORKING CONDITIONS [EP] Full entry	
DE Haushaltszulage	★★★★ +@
EN head of household allowance	★★★★ +@
SOCIAL QUESTIONS, Administrative law [EP] Full entry	
DE Haushaltszulage	★★★★ +@
EN household allowance	★★★★ +@

Sample: IATE

Apparently doublettes, i.e. redundant entries

ESTeam Omeo User: Database : IATE - Language : German - Num. of Expressions : 1080

File Edit View Settings Help

Origin Id	Status	Synonym Id	# of Match Re	Last Visited
126722	Completed	220	2	
1391555	Completed	220	2	
1674641	Completed	220	2	
126628	Completed	218	2	
1391817	Completed	218	2	
1560003	Completed	218	2	
1265824	Completed	209	2	
1690954	Completed	209	2	12/11/2012 15
777520	Completed	209	2	12/11/2012 15
1265819	Completed	207	2	12/11/2012 15
146561	Completed	207	2	12/11/2012 15
1670731	Completed	207	2	12/11/2012 15
126433	Completed	206	2	12/11/2012 15
155540	Completed	206	2	12/11/2012 15
2113584	Completed	206	2	12/11/2012 15
125724	Completed	204	2	12/11/2012 15
208746	Completed	204	2	12/11/2012 15
880050	Completed	204	2	
123383	Completed	196	2	
151985	Completed	196	2	
185203	Completed	196	2	
1227937	Completed	193	2	
135220	Completed	193	2	
763921	Completed	193	2	

Haushaltszulage

Match
Spell

Origin Id: 126433, Automatic, Type of Processing: Normal, Type of Matching: Equivalent

Haushaltszulage

Rej Acc

Haushaltszulage

Rej Acc

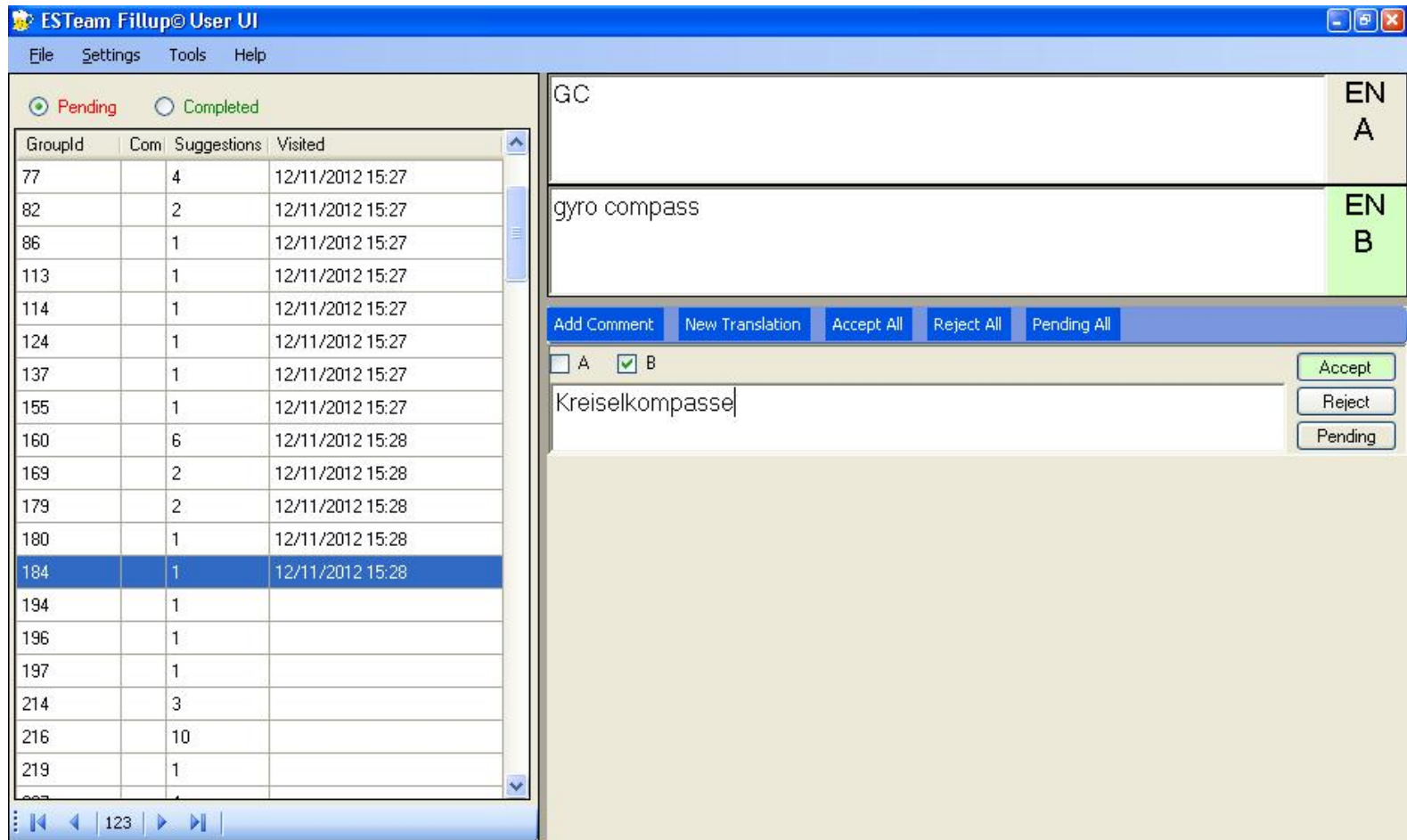
<<SynGroup<< <<Pending<< << >> >>Pending>> >>SynGroup>> ClearSelection

Omeo – Identifying Doublettes

Three IATE records suggested to become one.

Fillup

- ▶ Fillup: Enhance the database into other languages by using translation memory resources
- ▶ Extraction of translation candidates



Fillup – Suggesting translations

Out of a translation memory, Fillup suggests for the existing English terms „GC“ / „gyro compass“ the German term „Kreiselkompass“

The LISE collaboration platform

- ▶ Achieve interoperability & collaboration
- ▶ Improve quality of multilingual legal term banks
- ▶ Facilitate (inter-institutional) meta-communication
- ▶ Workflow-oriented

The LISE collaboration platform

The screenshot displays the LISE (Legal Language Interoperability Services) collaboration platform interface. At the top left is the LISE logo and the text "Legal Language Interoperability Services". On the top right, there is a "Sign out" button with a downward arrow. A navigation menu below the logo includes "HOME", "PROJECT", "ABOUT", "CONTACT", and "PROFILE".

The main content area is titled "MY STREAM" and contains several sections: "INBOX" (with a dropdown arrow), "BOOKMARKS" (with a right-pointing arrow), "AREAS" (with a dropdown arrow), and "ALL TOPICS". Under "AREAS", there are five categories: "General" (selected with a green circle), "Harmonization", "Standards", "Technology", and "Terminology Advice".

The stream shows a post by Michael Wetzel from 31 Jan 2012 at 14:34. The post title is "*termbase*" or *term base*?" with a reply count of 2. The post content includes a "Recipients" section and an "Areas" section with "General" selected. Below the post is a text input field with a left-pointing arrow and the text "reply to this topic...".

A second post by Michael Wetzel from 02 Apr 2012 at 14:31 is visible below. The post content reads: "Brilliant feedback! So, I suggest we go for 'termbase' as the short form and for 'terminological database' as the long form. ... well, and in acronym form we anyhow also see 'TB', don't we?"

The project has received funding from the European Community (ICT-PSP 4th call) under Grant Agreement n° 270917.

LISE in action: demonstration

Outlook: LISE Service Version Three

- ▶ Available in July 2013
- ▶ Functionalities to be implemented:
 - Workgroups
 - File attachments to posts
 - Full-text search & filtering
- Feedback on adaptation to functioning with IATE metadata

Further information

Project website: <http://www.lise-termservices.eu>

LISE

Legal Language
Interoperability
Services

[HOME](#) [ABOUT](#) [CONSORTIUM](#) [NEWS & ACTIVITIES](#) [DOWNLOADS](#) [CONTACT](#) [MEMBERS LOGIN](#)

HOME

*Welcome to the
LISE project*

Consortium

- *University of Vienna,
Centre for Translation
Studies*
- *EURAC, Institute for
Specialised
Communication and
Multilingualism*
- *Austrian
Parliamentary
Administration
(Parlamentsdirektion)*
- *CrossLang*
- *Esteam AB*

Supported by:

WELCOME TO THE LISE PROJECT!

LISE (Legal Language Interoperability Services) is a European project which addresses the urgent need for consolidated administrative nomenclatures and legal terminologies as tools to enhance interoperability and cross border collaboration. Without high quality and standards-based terminologies, it is impossible to reach precision, efficiency, and transparency within and across any services, processes or systems in the areas of legal and administrative work.

The key points of the project are

- identifying current problems and future options for optimization in legal and administrative terminology work and terminology management workflows
- developing a web-based terminology service platform for collaborative inter-institutional work
- analysing real-life processes and workflows in close cooperation with the relevant user groups
- active participation of and cooperation with the relevant user groups

Visit LISE Service Version Two at app.lise-termservices.eu or register for it [here](#).

LISE Twitter Feed

LISE **LISEtermervice** Very interesting LISE meeting with IATE on 30th October 2012 in Luxembourg City. 1 day 21 hours ago.

LISE **LISEtermervice** First day of LISE meeting with guests from the University of Iceland (with public section)! 4 months 2 weeks ago.

LISE **LISEtermervice** LISE 3rd EMB/TMB Meeting next week in Iceland at the University of Iceland (Háskóli Íslands)! Planning the next steps! 4 months 2 weeks ago.

LISE **LISEtermervice** LISE poster presentation on the 20th June @ TKE 2012! Come and visit us! t.co/DpcZjmPk 4 months 4 weeks ago.

The project has received funding from the European Community (ICT-PSP 4th call) under Grant Agreement n° 270917.

Thank you for your attention!

Tanja Wissik

tanja.wissik@univie.ac.at

Deputy Project Coordinator & Researcher, University of Vienna

Vesna Lušicky

vesna.lusicky@univie.ac.at

Researcher, University of Vienna