

**European Language
Resource Coordination**
Connecting Europe Facility



Context for language data collection in Norway

Sarah Jane Hails Gjelsten & Koenraad De Smedt

6 July 2016 – LRB Meeting, Lisbon



Presentation outline

1. A bit about Norway
2. Our involvement prior to ELRC workshop
3. ELRC workshop in Norway – participation and useful findings
4. Plans for the road ahead



Norway - more than fjords and mountains



Our involvement prior to the ELRC workshop

University of Bergen



- Partner in META-NORD, 2011 – 2013
- Shared 49 datasets in Norway with metadata and licenses
- Produced 2 white papers (Norwegian Bokmål, Nynorsk)
- Organized awareness events
- Participated in European meetings
- LRB scientific contact: Koenraad De Smedt



Agency for Public Management and eGovernment (Difi)

- Shall develop and modernise the Norwegian public sector
- Subordinate to Ministry of Modernisation and Local Government (KMD)
- Responsible for coordinating CEF Digital in Norway
- LRB public sector contact: (previously KMD v/Gard Jenssen) now Difi v/Sarah Hails Gjelsten

ELRC workshop in Oslo, June 8th



- 50 attendees
- Good range of speakers
- Informative presentations
- Coincided with new report
- Positive feedback

- Useful findings for us...

Useful finding (1)



Responsibility for multilinguality in Norway rests with public agencies.

Requirement to provide information and services **for all**.

Useful finding (2)



Kartlegging av behovet for automatisk oversettelse i statlig sektor

Utarbeidet for Kommunal- og moderniseringsdepartementet

OE-rapport 2016-15

Interest at
Ministry level
for Connecting
Europe Facility.

**Report
published in
June 2016**

Machine
translation
recommended
to cut costs.

Norwegian
needs
strengthening
within system.

Simple data
provision and
collection
recommended

Useful finding (3)



Interest for developing multilingual cross border services is budding.

Online Dispute Resolution

HOME ABOUT THIS SITE GET HELP [SIGN IN >](#) [Register](#)

About this site

If you've had a problem with something you've bought online, you can use this site to try to reach an out-of-court settlement. You can only use it if you live in the EU and the trader is based in the EU. In some countries, you can also use this site if you are a trader and you want to complain about a consumer over a good or service you sold online.

I am a consumer
I want to submit a complaint against a trader

I am a trader
I want to submit a complaint against a consumer

A silhouette illustration of several business professionals in a meeting, standing and talking, set against a light blue background.

“Translation needs in the public sector are real and growing.”

Ministries and public agencies who were active at the workshop

- Ministry of Foreign Affairs
- Ministry of Children and Equality
- Directorate for Integration and Diversity
- Directorate of Immigration
- Directorate for Labour and Welfare
- Brønnøysund Register Centre
- Norwegian Public Roads Administration
- Lovdata (online legal database)



UDI



Most public agency in-house translators already use translation tools

Public Agency	Digital translation tool (CAT-tool)
Ministry of Foreign affairs – EEA unit	SDL Trados Studio
Ministry of Foreign affairs – translation unit	MultiTrans Prism
Public Roads Administration	Memsources
Norwegian Maritime Authority	SDL Trados Studio
Financial Supervisory Authority	<i>No tool used</i>
Parliament	<i>No tool used</i>

Useful finding (6)



Translation activities vary from agency to agency, and within agencies.

Need to raise awareness, cooperate and work smarter – requires coordination.

Useful finding (7)



Several entities exist with knowledge and expertise to assist in Norway.

Established language data banks and research institutions to draw upon.

Norway promotes and enables the sharing of open public data.

Register over åpne datasett i Norge

Her finner du publiserte datasett og åpne API-er fra offentlig sektor i Norge. [Les mer om data.norge.no](#)



FINN VIA TEMA:

Offentlig forvaltning (179)

Transport og kommunikasjon (35)

Utdanning og forskning (26)

Kommuner og distrikter (70)

Mat, fiskeri og landbruk (31)

Miljø (24)

Økonomi (35)

Helse og omsorg (26)

Kultur, fritid og livssyn (22)

[SE ALLE TEMA >](#)

[Registrer datasett >](#)

[Registrer applikasjon >](#)

What data can be shared?



Red: **Closed data** – Sensitive information – not to be shared

Amber: **Shared data** – Personal data - only shared through secured systems for case processing

Green: **Open data** – free and may be shared and used by all (public and private)



Help yourself!
(at your own risk)

Potential data holders



- Agencies with translation activities
- Public web site administrators
- Public Open Data, data.norge.no
- Accessible annotated language resources: corpora, termbanks etc.
- Raw non-fiction and fiction texts (National Library?)

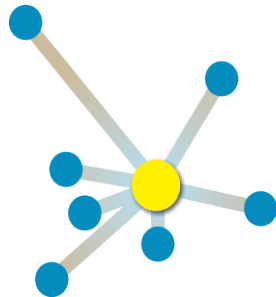
Plans for road ahead

- In Norway
 - Map existing needs and available language data
 - Promote cooperation and sharing in public sector
- In cooperation with ELRC
 - Data collection from potential data providers
 - Establish routines for providing segments and language data files
- Follow up on CEF.AT
 - Promote CEF.AT, funding call with potential users



Thank you for your attention

Tusen takk!



**European Language
Resource Coordination**
Connecting Europe Facility



difi Direktoratet for
forvaltning og IKT