

CLARIN @ universities

Utrecht, 4-5 November 2019

Day 1

Opening and introduction

- **Participants introduce themselves briefly:** Name, affiliation and role.
- **Our mission:** CLARIN makes digital language resources available to scholars, researchers, students and citizen-scientists from all disciplines, especially in the humanities and social sciences.
- **Just making available is not sufficient for people to feel stimulated to use them effectively in their research:** In order for CLARIN to have a lasting impact, it is necessary for both methodological and infrastructural issues to become integrated in university curricula.
- **Part of CLARIN's strategic planning for 2018-2020:** Develop best practices for enhancing MA and PhD programmes to teach about the digital research methods needed to effectively use the CLARIN tools and resources.

Approach

- **Not wise to go for a one-fits-all approach:** legal and practical obstacles to influence curricula vary widely, across countries, universities or even departments.
- **Start with a (pilot) survey** of what is there in the CLARIN countries, collected by key persons from national consortia.
- **Present results at internal workshop**, comment, share obstacles and possible solutions.
- **Formulate best practice recommendations**, identify methods and materials that can be shared, make plans for targeted actions to improve the situation.

Workshop programme

4 November

12.00 - 13.00	Lunch
13.00 - 13.15	Opening and introduction (Steven)
13.15 - 13.30	Survey results (Darja)
13.30 - 15.00	Key problems and opportunities (all)
15.00 - 15.30	Coffee break
15.30 - 16.15	Breakout session 1 (all)
16.15 - 17.00	Feedback from working groups (all)
17.00 - 17.30	Next steps (all)
19.00	Dinner

5 November

09.00 - 10.00	Invited talk (Stefania Scagliola)
10.00 - 11.00	Breakout session 2 (all)
11.00 - 11.30	Coffee break
11.30 - 12.30	Feedback from working groups (all)
12.30 - 13.30	Next steps (all)
13.30	Lunch and closing (Steven)

Survey results

- **Feedback:**
 - 22 CLARIN member and observer countries
 - 4c separate feedback from 2 lecturers (France, Greece, Hungary, Slovenia)
 - 3c single feedback with experience from several colleagues (Austria, Poland, Netherlands)
 - 2c general summary of their activities without answering the survey (South Africa, Czech Rep)
- **Basic info:**
 - 55 individual courses, taught at 31 different universities in 20 countries
- **Course level:**
 - Masters (31), Bachelor (16), Doctoral (3), several levels combined (5)
- **Course size:**
 - up to 10 students (10), 10-30 students (34), over 30 students (11), a total of 6000 students
- **Taught since:**
 - 2008-2020, 78% courses 2015-2019 (43)

Course info

- **Discipline:**
 - Linguistics&language (34), DH (6), Language engineering (4), Cultural&Media studies (4), Information&Computer science (4), Cognitive science (1), Mixed (2)
- **Coverage:**
 - Full course (22), Single session (13), Few sessions (13)
- **Teaching type:**
 - Lecture+exercise (30), Lecture+exercise+assignment (12), Lecture only (6), Online course (3), Exercise only (2), Assignment only (1), Reading seminar (1)

Course names

Computational Linguistics
Corpus Linguistics
Digital Humanities
Literary Translation
Computer-Assisted Literary Translation
Recent advances in Linguistics
Computers and language
Lexical Resources
Contrastive English-Lithuanian Stylistics
Natural Language Analysis
Corpus analysis tools
Introduction to research methods
Cognitive Science 2 (Multimodality in
Communication and Cognition)
Language Technologies
Corpus Linguistics and Statistical Methods
Linguistic technologies for information extraction
Data Clinic
Methods
Digital Hermeneutics
Phonetics in language learning and teaching

Digital history in the 19th and 20th century
Television History Online
Introduction to the (computational)
lexicography
Introduction to Speech Analysis
Language and computers
Language resources
Language modeling
Digital Humanities: Web Resources, Tools
and Infrastructures
Language resources for machine translation
Digital Libraries
Language technologies for machine
translation
Digital Philology
Lexicology and Lexicography of Lithuanian
Language
Digital Texts and Multicultural Studies
Linguistics in the Digital Age
Digital Tools for the Humanities
Management of Big Information Sources

Digital Tools in Literature and
Translation
Methods of digital enrichment
Discourse Analysis
Natural language processing
ELDP Training
Programming and Curation
Film, Text, Context
Specialized Translation
German Corpus Linguistics
course
Bulgarian Syntax in non-native
environment
Introduction to programming

CLARIN content covered

- **RI:**
 - Intro to RIs, finding, accessing and reusing CLARIN language resources and tools, legal and ethical issues, open science, research data management, metadata standards, standards for language resources, depositing research data
- **Tools:**
 - ELAN, Korp, e-magyar, Weblicht, Praat, Tundra, VLO, CLARIAH Media Suite
- **Resources:**
 - CLARIAH Media Suite, HuComTech corpus, Language Bank of Finland, plWordNet, UD Treebanks
- **Methodology:**
 - Usage of text preprocessing tools, usage of language resources, language data research methods, development of language resources and/or tools

Initiative

- lecturer's initiative
 - o Bulgaria
 - o Croatia
 - o Denmark
 - o Hungary
 - o Iceland
 - o Netherlands
 - o Norway
 - o Portugal
 - o UK
- lecturer's initiative but with consortium support when needed
 - o Austria
 - o Germany
 - o Greece
 - o Slovenia
- strategy to leverage individual efforts into a common action
 - o France
- driven and supported by consortium
 - o Finland
 - o Greece
 - o Italy
 - o Lithuania
 - o Poland

Satisfaction of integration

- Not satisfactory
 - Austria
 - Croatia
 - Denmark
 - France
 - Iceland
 - Italy
 - Portugal
 - Slovenia
 - UK
- Not bad but could be better
 - Bulgaria
 - Germany
 - Greece
 - Hungary
 - Latvia (LRs<s widely used but not recognized as CLARIN)
 - Lithuania
 - Netherlands
 - Norway
 - Sweden
- Satisfactory
 - Finland (languages and DH, could be better for other fields)
 - Poland

Key positive experience

- Advantages of RIs
 - the interfaces are usable, reliable, state of the art; students like the idea that they can access services that are hosted in other countries with their university identifier
- Findability of resources
 - search the VLO for concrete corpora and to show the power of well curated metadata
- Reuse of resources
 - students use CLARIN corpora and tools in master theses, projects and internships with companies
- Boost to research
- Benefits for the infrastructure
 - a hard lesson how incomprehensible and partial are our user guides, UIs and also technical documentation
 - the integration of CLARIN content in teaching motivated some of the lecturers to archive their materials and make them accessible to the research community and the general public
- Student feedback
 - a student who explored Fin-CLARIN said that it is paradise for researchers

Key obstacles

- Lack of resources and/or tools for a specific language
 - Lack of resources for French, in particular of corpora and NLP modules
- Lack of standardization
 - Lack of standardization that hinders reuse of resources and tools
- Lack of teacher expertise
 - There is a potential to use lots of other tools that I am insufficiently familiar with. There is insufficient documentation of available resources and no tutorial on how to effectively use them in the classroom.
- Lack of training materials
 - Lack of time and other resources to locate or create materials for all potentially relevant technologies.
- Too difficult for students
 - Many corpora are in TEI, and there are not many tools which allow graphical exploration or analysis of TEI corpora

Other infrastructures taught

- DARIAH
 - o Digital Archives und Data Management (Lecture and Exercise): The infrastructure is introduced, services are tested together with the students.
 - o Textmining for Digital Humanities (Lecture and Exercise): Usage of the DARIAH-DE TopicExplorer
 - o [DARIAH content is being used at the Faculty of Classical and New Philologies](#)
 - o DARIAH is taught at Université de Lorraine
 - o in France some lecturers use Voyant, for instance those by Aurélien Berra, in Paris
 - o Elena Pierazzo in Tours may use the DARIAH teach infrastructure for which she contributed an introduction to TEI
- CESSDA and ESS
 - o My institution is the coordinator of CLARIN national consortium, but it is also a coordinator of two more RIs from Social Sciences: ESS and CESSDA. The colleagues sociologists integrate the information about their relevant RIs in their curricula
- KB Delpher (newspapers)
 - o There are many courses making use of KB Delpher (newspapers), that will be part of the CLARIAH infrastructure (WP6) and which have also been embedded and made searchable in the Media Suite
- Media Ecology Project and Cinemetrics
 - o Eric Hoyt's Media Lantern Project, Mark Williams' Media Ecology Project and Yuri Tsivian's Cinemetrics are used in other courses. With CLARIAH WP5 we are looking into collaborations with the two first-mentioned projects and have already to some extent exchanged software and data.
- Korp and Sparv
 - o Swe-Clarín/Språkbanken Text also uses a range of tools, such as Korp and Sparv, in teaching situations in several other courses in Swedish

Key next steps

- Teaching materials
- Training
- Dissemination
- Technical development

The role of CLARIN ERIC

- Provide materials for the teachers
- Support the exchange of teaching materials
- Provide teacher training
- Support teacher exchange and teacher collaboration

The role of national consortia

- Provide/improve documentation
- Improve integration of resources and tools
- Intensify dissemination
- Offer target-specific training
- Offer small funding for teaching pilots
- Support collaboration among universities

What can we learn from other infrastructures

- General
 - data sharing and data discoverability as key for interdisciplinary research and innovation
- DARIAH
 - Workshops and summer schools
 - DARIAH-Teach
 - Interdisciplinary topics of working groups
- SSK
 - adapt it to build teaching scenarios
- Europeana
 - funding scheme for integrating content and tools into teaching
- Ranke 2.0
 - inspiring examples of tutorials and/teaching materials

Key problems and opportunities

13.30-13.35	Austria	Gerlinde Schneider
13.35-13.40	Bulgaria	Petya Osenova
13.40-13.45	UK	Vera Ferreira
13.45-13.50	Czech Republic	Pavel Stranak
13.50-13.55	Finland	Mietta Lennes
13.55-14.00	Germany	Andreas Witt
14.00-14.05	Hungary	Tamás Váradi
14.05-14.10	Italy	Federico Boschetti
14.10-14.15	Latvia	Inguna Skadiņa
14.15-14.20	Lithuania	Jolanta Kovalevskaitė
14.20-14.25	Netherlands	Christian Olesen
14.25-14.30	Norway	Koenraad De Smedt
14.30-14.35	Poland	Maciej Piasecki
14.35-14.40	Portugal	Amália Mendes
14.40-14.45	Croatia	Marko Tadić
14.45-15.00	all	discussion

Key problems and opportunities in Austria

University of Graz

BA-Level Optional course modules: 'Information modelling' and 'Digital science'
(Open for all Humanities undergrad programs)

MA Digital Humanities (since 2017)

PhD Digital Humanities (since 2017)

Traditional Humanities programs Romance studies, Slavistics, Translation studies, Linguistics...

Austria

Individual courses on DH

Traditional Humanities programs

Key problems and opportunities in Austria

Challenges

- **Dedicated DH courses and programs**
- **Traditional philological and linguistic study programs**
- CLARIN, its services and benefits are relatively unknown
- Research infrastructures are not integrated in the curriculum
- It depends on the lecturer, if they integrate CLARIN in their courses
- Changing syllabi and developing new content is often too much effort for lecturers

Key problems and opportunities in Austria

Opportunities

Digital Humanities Austria Consortium (CLARIAH-AT):

- Strategy: Development of study and continuing educational programs
- Anchoring of modules on DH methods in already established degree programmes in the humanities
- Agreement between the universities and the ministry: Integration of the DH in the universities

Ideas

- Integrate Research Infrastructure in general in the Curriculum
- Workshops and trainings for traditional researchers
- Info desks and tutorials at relevant conferences
- Teaching material that could easily be integrated into ongoing courses

Key problems and opportunities in Bulgaria

Key problems:

- Crossing and thus - duplicating - policies by the Ministry
- No awareness of the services that CLARIN provides
- The CLaDA-BG services still at a very initial stage
- No enough courses in English

-

Key problems and **opportunities** in Bulgaria

Opportunities:

- Localisation, adaptation and a targeted re-use of the many developed courses within the CLARIN community
- A more focused connection between the course content and the infrastructure knowledge base: viewing topics in a context

Key problems and opportunities in Croatia: context

- University of Zagreb, Faculty of Humanities and Social Sciences (FFZG)
 - Dept. of Information and Communication, Programme in Informatology
 - Study of “SSH Informatics” established in mid ‘80, but somehow it didn’t follow the development in the field in last 10 years, particularly regarding RI
 - Dept. of Linguistics, Programmes in Linguistics: CLARIN content found its way in
 - B.A.: Course in Corpus Linguistics
 - M.A.: Course in Language Technologies (orientation Computational Linguistics); course in Translator and Computer (orientation Translation in all philological studies)
 - PhD: Courses in Computational Linguistics and Corpus Linguistics
 - Dept. of Classical Philology
 - Substantial collection of Medieval Latin corpora: initiative by a professor there
 - Right now FFZG is going into a new round of reform of our study programmes
 - Initiative to have an interdepartmental M.A. study in Language Technologies
 - CLARIN content already exists in many of these programmes, but it should be presented in more details (not just lectures, but also hands-on sessions)

Key problems and opportunities in Croatia 2

- Many traditional (philological) departments and their staff are reluctant to use LT at all
 - Inadequately informed about CLARIN RI
- Almost 7000 students have only 3 larger and 2 smaller computer rooms available
 - Although many have their own and there are over 200 desktops in the library
- HR-CLARIN consortium (and repository) still *in statu nascendi*, waiting for the approval by Ministry
- CLARIN RI should be more visible in different research communities
 - CLARIN Ambassador concept could be reshaped, maybe set at the lower, national level:
 - Selected researcher from a research community (e.g. History, Anthropology, Ethnology, Psychology...)
 - Familiar with the eScience paradigm and DHSS
 - Educated about CLARIN RI
 - possible central courses for Ambassadors (online, F2F)
 - Spreads the information about CLARIN RI in his/her research community
 - HR-CLARIN would like to try with this concept and check how it will work
- Translate, adapt and re-use many courses developed by the CLARIN community (Petya)

General problems in Finland

- Many overlapping efforts in teaching and other activities related to Digital Humanities
- Confusion due to recent and forthcoming changes in
 - university entrance methods (exams/ no exams; MOOC courses may replace entrance exams)
 - the contents and structures of study programmes at universities
 - organization between universities
- Pressure towards universities to increase their output (number of degrees, "life-long learning")
- Teachers have less and less time and energy for planning and adopting new content.

Opportunities in Finland

- **FIN-CLARIN** and the **Language Bank of Finland** are nationally well established.
- **MA Programme in Digital Humanities and Social Sciences** (University of Helsinki; related to DARIAH) is planned to start in 2020 – FIN-CLARIN courses included.
- A promising new national **service for open educational resources (OERs)**: <https://aoe.fi>
 - to open for uploads in Nov-Dec; to be integrated with library systems during 2020
 - offered by the *Ministry of Education and Culture* and the *Finnish National Agency for Education*
- Aoe.fi could be used by FIN-CLARIN for sharing open content.

More general opportunities for CLARIN

- An open library of small pieces of learning content might be useful
 - e.g., *practical* issues in Data Management that are time-consuming to teach
 - Metadata descriptions and citeability are essential
 - CLARIN has the expertise here!



University of Oulu

FIN-CLARIN

KIELIPANKKI

The Language Bank of Finland

University of Vaasa

University of Jyväskylä

University of Eastern Finland

Tampere University

University of Turku

Institute for the Languages in Finland

CSC -IT Center for Science (Technical center)

University of Helsinki (National coordination)

Sweden

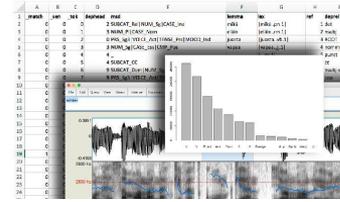
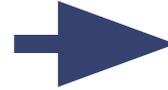
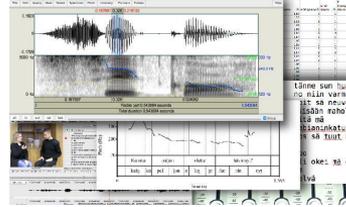
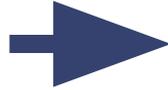
Finland

Russia

Estonia

Google My Maps

Online FIN-CLARIN courses in programmes at UHEL



Corpus Linguistics and Statistical Methods (5 cr)



BA programme in Languages:



optional in common studies

Introduction to Speech Analysis (5 cr)



MA programme in Linguistics:

mandatory in Phonetics, optional in other study tracks

Data Clinic (5 cr)



MA programmes:

- Linguistics (Finnish)
- LINGDA (international)

Also optional in doctoral studies



The courses are open to all students and researchers within and outside the University of Helsinki, even abroad.



DIGITAL HUMANITIES
COURSE REGISTRY

Key problems and opportunities in Germany

Problems:

- The teaching landscape is very heterogeneous
- In the process of (re-)accreditation the content of study programs is defined in a general way, to allow a certain amount of freedom
- The professors have the power to integrate content as they like

Opportunities (see problems):

- Teach the teachers
- Provide teaching material that can be adapted and reused by the teachers

Key problems and opportunities in Hungary

PROBLEMS

- Digital Humanities does not exist at universities
- CLARIN partners mostly research centres, less direct contact with universities
- University humanities research is fragmented and pursued mostly individually
- Infrastructure at universities is very poor
- HUN-CLARIN operates without any national funding

OPPORTUNITIES

- There is a palpable demand for and interest in digital technology
- HUN-CLARIN Roadshows very successful
- Research projects may recruit and educate young researchers in digital humanities

Key problems in Italy

- Teachers are not yet sufficiently informed about CLARIN opportunities
- E-learning platforms (mainly Moodle) are used in many universities: a better integration between CLARIN resources and them is desirable (e.g. creation of SCORM pluggable modules)
- Many “traditional” courses (e.g. in philological disciplines) are reluctant to use digital resources in class
- An integration between primary sources (e.g. literary texts) and secondary sources (e.g. critics, scientific articles) is necessary

Opportunities in Italy

- **Institute** for Computational Linguistics “A. Zampolli” (host of CLARIN-IT), which promotes many activities with Universities (European and national projects, seminars) and High Schools (internship, training)
- Increasing interest in **Digital Humanities**
 - New **Courses** in Digital Humanities in many Universities (e.g. Bologna, Venice, Genoa, Catania, etc.)
 - New **Centers** for Digital and Public Humanities (e.g. VeDPH at Ca’ Foscari University, DH Center - University of Salento)
- **Associations** of Computational Linguistics (AILC) and Digital Humanities (AIUCD)
- **Workshops** within the annual conferences of the associations
- **Networks** (e.g. DiCultHer)
- Permanent **Seminars** (e.g. LabCD in Pisa) and **Summer Schools** (e.g. in Pisa, Verona, Palermo, and Venice [next year])

Key problems and opportunities in Latvia

- Some initiatives at the initial stage in Digital Humanities (1 course in Faculty of Humanities at the University of Latvia, Master Course at the Riga Technical University)
- Computational linguistics course at the University of Latvia for master students of English Philology
- Some lectures on corpus linguistics and some other topics of computational linguistics for doctoral students of University of Liepāja and Ventspils University of Applied Sciences
- Lack of financial resources for expanding courses related to use of language resources and tools for SSH
- Lack of CLARIN repository (under construction)
- CLARIN and its services are relatively unknown in the traditional disciplines

Key problems and **opportunities** in Latvia

- Interest from linguists and social science researchers
- Interest from administrative bodies (e.g., Ministry of Education and Science)
- CLARIN ERIC and NC can together create relevant knowledge and share know-how

Key problems and opportunities in Lithuania (Vytautas Magnus university)

Possibilities of integration of CLARIN content into university curricula

CLARIN content (VLO) at the VMU library (next to licenced databases), could increase the visibility and findability of CLARIN resources for lecturers and students.

Resource materials for lecturers of different profiles and for different target groups of students (foreign languages, sociology, history, literature, etc.):

- discipline-based working groups at the national level;
- basic reusable material (in EN) as a MOOC (international level);
- knowledge and best practice sharing initiatives at the CLARIN events (e.g., pre-conference workshop at the Annual Conference) or separate platform "CLARIN@universities".

Key problems and opportunities in the Netherlands

(experiences based on CLARIAH Media Suite experiences)

Key problems

- English language content
 - Media Studies programs partly or often fully in English
- Integration into traditional course modules Film & Media Studies
 - Resistance DH
 - Digital tools are not perceived as accommodating for high level of abstraction, source criticism or reflexivity
 - Time consuming (teachers need training)
 - Not always acknowledged that digital tools are already integral to Hum research (databases, etc.)
 - Shortcoming tools:
 - Offline tools are often more advanced/better for analyzing media resources

Opportunities

- Content! (film and AV archives)

Key problems and opportunities in Norway

- Teachers and courses currently using CLARIN: very few
- Teachers and courses potentially using CLARIN: impossible — teachers lack willingness and courses lack methodological component
- My immediate colleagues and I use mostly our own language resources and tools since we are familiar with them
- Students need to be able to find resources for term papers, master thesis, etc.
- It is not easy to identify which resources and tools are optimally suited to exploring some linguistic aspects
- Many relevant tools and data are not in CLARIN (students don't care)
- Use of other resources may require too much effort (lack of documentation, lack of examples, unusable formats etc.)
- ***We must cooperate to prepare demos, walkthroughs and "recipes", replication packages with scripts and suggestions for exercises and student projects.***

Poland: key opportunities

- Academic users already interested in applying CLARIN tools in teaching
 - enthusiasm towards digital methods
 - however many of them are junior staff members if not PhD students
 - training workshops initiated or even organised by users
- Large diversification of users' areas
- Great interest in CLARIN tools, not only resources
 - several tools available
 - but this imposes quite heavy expectations on CLARIN-PL
 - concerning support and robustness under stress of intensive use
- CLARIN is perceived as the major site for LT for Polish

Poland: key problems

- Teaching is not perceived as a primary application of “big research infrastructure”
 - focus on research applications
- Lack of training materials for students
 - even user guides and documentation is partial and not fully comprehensible
- Lack of any linguistic background among our students
 - from Computer Science but also Social Sciences

Poland: key **problems**

- Formalised procedures for introducing and getting accepted new courses
 - one two years ahead, formal path: a department, faculty commission, faculty board
 - predefined lines of studies and specialisations
 - including fixed skeleton of obligatory courses,
 - only 30-40% of the schedule is not pre-determined

Key problems and opportunities in Portugal

- School of Arts and Humanities:
 - courses in Corpus Linguistics (MA)
 - Discourse Analysis (MA and BA)
- Students respond positively to content on corpus compilation, concordances, collocations, but less so to more advanced contents (processing, unix usage, annotation)
- Traditional courses in Humanities show little interest in CLARIN contents (at least until recently)

Key problems and **opportunities** in Portugal

- The infrastructure PORTULAN CLARIN will soon be online and it will increase the visibility of the resources for Portuguese
- The School of Arts and Humanities is one of the PORTULAN CLARIN proponent partners: the infrastructure was presented at an internal interdisciplinary research event
- It became evident that there is a strong interest in the Digital Humanities and the sharing of standards in the field
- This interest translated mostly at the research level and not yet in the teaching curricula

Key problems and opportunities in the UK

Key problems:

- Lack of awareness of the services that CLARIN provides
- Lack of specific training for the integration of CLARIN content and resources
- Lack of time to plan and adjust to new content

Key problems and opportunities in the UK

Key problems:

- Lack of background knowledge regarding the importance of creating and archiving good primary data for research and community purposes
- Making research data accessible in linguistics and philology is still very biased

Key problems and **opportunities** in the UK

Key opportunities:

- ELDP training (every year)
 - Increase in quality and amount of data collected
 - Increase in the quality of the collections archived
 - Strengthen the dialogue between researchers, also form different disciplines

Key problems and **opportunities** in the UK

Key opportunities:

- Courses on data management, data processing and archiving at SOAS
 - Increase of interest in archiving primary data
 - Increase of interest in data mining by using archived materials (MA and PhD thesis)

Key problems and opportunities

discussion

Breakout session 1

- (How) should a follow-up survey be conducted to obtain information at the level of national consortia?
- What is your dream CLARIN curriculum?

Feedback from working group 1

Question 1

- Yes
 - To make CLARIN content more visible
 - We want to know:
 - if people are aware of what is out there
 - what they are using
 - if they are not using any tools why
 - what do they need for their work

Feedback from working group 1

Question 1

- Maybe a broader survey
 - people who do not belong to the national consortia but we will probably be dependent on people that we know (our network contacts) or use conferences / workshops to share the survey among participants
 - Ask about resources available in general, if they use them and if not why
 - It should be limited to 5-6 questions, combination of multiple questions with open options
 - The questionnaire should entail a tick box for getting more information about CLARIN (we would like to have a non-anonymised survey to be able to provide feedback)

Feedback from working group 1

Question 2

The ideal curriculum should allow:

- Stable part
 - Reference to research infrastructure, metadata, legal issues, open/fair research data
- Flexible part
 - Toolbox like approach: discipline and research specific tools, data sets, exemplary resources
 - Workflow between tools (how each tool interact with each other, to achieve better results / different purposes and cover all the steps for instance from data collection to data processing, data archiving and data analysis)
 - Having a flexible organisation structure to be able to integrate different formats like summer schools or intensive trainings on particular topics

Feedback from group 2

Q1: Follow-up survey?

No.

- Too much work
- Information about available courses and curricula can already be shared by teachers via Digital Humanities Course Registry.
- Better to ask people directly at suitable situations?

Feedback from group 2

Q2: Dream curriculum?

- Introducing the stages of a research data analysis workflow (applicable to any field of science)
- Finding some elementary concepts and tools that “everyone” should know within Digital Humanities and the Social Sciences – same for ALL languages!
- What is “the best morphological analyzer for Finnish”?
- Text editors, plain text, regular expressions
- Processing text or video to make it more accessible and searchable
- Exporting data, data mining tools, more advanced NLP methods (required from all language document intensive studies?)
- Using digital tools during the entire studies
- Hackathon: talk among digital humanists, computer scientists, statisticians, ...
- Making data citeable, open/accessible
- Data and tools in Real Life: after university

Feedback from working group 3

Scale up the survey to all institutions in each country offering relevant curricula? Maybe.

- A survey with direct questions about CLARIN potential presupposes that respondents are sufficiently informed about CLARIN
- A survey not asking about CLARIN potential, but directed at getting an overview of current goals and methods in (digital) humanities education, would require a follow-up to assess CLARIN potential for change
- A survey would require significant efforts by the national coordinators, e.g. in making a list of selected respondents (lecturers and/or study program coordinators?) at selected institutions and in evaluating results
- A survey could result in new possibilities to use CLARIN, but their realization is far from guaranteed.

Feedback from working group 3

What is the dream curriculum? This is a relative question without a single answer. However, some useful elements of courses and programs may be the following:

- Lectures are given by experts who know how to explain things, and who ground the contents in relevant research questions and well-motivated methodology.
- Demos starting from research questions, going step by step, using CLARIN resources and leading to the creation and deposit of results.
- Students do exercises (possibly in groups where they can support each other), using resources given by the teachers (preferably reusing data from CLARIN repositories) and tools that are pre-installed or online, avoiding technical bottlenecks.
- In their assignments, master's theses, etc., students ideally create new resources which they deposit in CLARIN.
- Programs should include internships at other institutions, with adequate scientific and technical supervision.
- Teachers (in groups of two or three) may exploit the same resource (e.g. a textual or multimedia corpus) from different perspectives (e.g. public history, digital philology, and text encoding).

Next steps

Day 2

Invited talk

Stefania Scagliola

- Slides:
 - Add url
- Teaching platform:
 - <https://ranke2.uni.lu>
- Cartoon:
 - http://2019.dhbenelux.org/wp-content/uploads/sites/13/2019/08/DH_Benelux_2019_paper_77.pdf

Breakout session 2

- What should national consortia do to improve the integration of CLARIN content to university curricula?
- What should CLARIN ERIC do to improve the integration of CLARIN content to university curricula?

Feedback from working group 1

The role of CLARIN ERIC

- Archive of digital courses
 - A kind of VLO for collecting information about digital teaching materials
 - Whole courses (complete)
 - Teaching materials related to different problems and tools
 - Different descriptions, at least different meta-data for different disciplines
 - E.g. a tool for: corpus annotation vs coding vs tagging vs ???
 - Harvesting and integrating data from platforms for learning content that are set up and maintained by NCs
 - For instance, <https://aoe.fi/>, example course: <https://aoe.fi/#/demo/materiaali/14/copyrights-in-education>
 - But also a central archive for 'homeless' teaching materials
- Collection of descriptions of basic, typical tasks and relevant tools, e.g. CMDI editing (Cheatsheets: http://dobes.mpi.nl/dobesprogramme/training_courses/)
 - Potential role of KSI and K-centres: identification of typical tasks, problems, basic tools, user guides and description of processes
 - dynamic, continuous process

The role of CLARIN ERIC

- Development of teaching materials for central functions of CLARIN ERIC
 - VLO, Switchboard, CLARIN Resource Families
 - Materials written from the perspective of different disciplines
- Provide teacher training
 - Definitely in relation to the teaching materials prepared by CLARIN ERIC (see above)
 - Support teacher training e.g. via mobility grants (see below)
- Support teacher exchange and teacher collaboration
 - Help from CLARIN ERIC to efficient use of ERASMUS grants for the exchange of teachers

Feedback from working group 1

Responsibility and tasks of the National Consortia

- Provide/improve documentation
 - Definitely the responsibility of NCs: if something is meant to be used, it must be well described from the perspective of users
 - The problem of the users' language and task and goal driven perception
 - Different description of tools for different disciplines
- Improve integration of resources and tools
 - Exemplars of processes based on tools and resources (close to the resources used by users)
- Intensify dissemination
 - Definitely
 - How: conferences, CLARIN training workshops focused on teaching how to use tools
 - Training workshops co-organised by users - based on resources and tasks from users
 - Direct collaboration and support for 'key' teachers
 - Looking for some channels of dissemination: Twitter, Facebook, and ???

Feedback from working group 1

Responsibility and tasks of the National Consortia

- Offer small funding for
 - Preparation of special versions of tools, adaptation of tools to users' materials
 - Support in running tools during the teaching hours (e.g. a laboratory) for students
 - Preparation of workshops
 - Development of teaching materials which can be shared on the ERIC level

Feedback from working group 2

- What should *national consortia* do to improve integration of CLARIN to university curricula?
 - Using CLARIN materials in existing curricula + create new courses, programs + how to teach our students to use CLARIN
 - The goal would be to address researchers in Humanities, identify potential projects, offer budget
 - CLARIAH experience in The Netherlands:
 - pilot projects applications had to work with tools and staff of the research environment - application had to specify all information: content, methods, some aimed primarily at making content/collections available others addressed more advanced research questions
 - The research projects organically led to teaching activities/tutorials though this was not initially the objective

Note: identify what exists, teaching materials, methods

This would benefit from the international information (question 2)

- Distribute teaching materials in the national infrastructure

Feedback from working group 2

- What should CLARIN ERIC do to improve integration of CLARIN to university curricula?
 - CLARIN ERIC should identify research questions and provide teaching materials
 - Present the information in terms of basic research questions that are addressed by an historian, a social scientist...
 - Suggest also the basic-level courses and/or course building blocks adapted to certain research questions in specific areas in the Humanities, e.g.
 - Basic digital document encoding and organisation (TEI, XML etc.)
 - Basic LT processing (tokenisation, lemmatisation, NERC, semantic nets, SRL etc.)
 - Advanced LT and KT-processing (LOD linking, NE-networks, semantic graphs, etc.)
 - Have them translated/subtitled to other languages
 - Use the DH Course Registry, possibly harvest the information, have links to VLO, also links or content linked to a resource, e.g. “register variation” (additional label “Education”)

Feedback from working group 3

National Consortia

- Different capacities:
- Typology of capacities → models → offer models to choose from CLARIN ERIC
- Monitor the uptake of model
- Teaching → course (usual environment; new environment)
- Dissemination → (seminar ; workshops)
- Identifying key venues for expanding audience→ (teacher association; historians associations; conferences)

CLARIN ERIC

- Monitor national effort
- Identify ideal/typical cooperation with help of national coordinators → offer support for developing user case→ publish user case
- The intersection between language and text corpora
 - Language and social change
 - Micro, meso, macro level
 - Short term, mid term, long term

Take away:

- Eric —> push the agenda of re-use in universities
 - Methods of research
 - Data management
 - Preparation for bachelor and master thesis
 - PhD training
- Publish dissemination material with pedagogical consideration
 - Who
 - What should the audience do
 - In which context

Who goes to the clarin website? Who would you like to reach?

Next steps Group 1

- Collect needs and requirements of lecturers from different fields such as Linguistics, Political Sciences, Literary Studies, Sociology, Anthropology, Social Psychology
Approaches can be: Expert interviews, Survey, Open vote where you can add topics (e.g. presemo: voting at <https://presemo.helsinki.fi/clarin>; result screen at <https://presemo.helsinki.fi/clarin/screen>)
- Organize trainings on a local level as a response to the evaluated requirements with the corresponding financial support
- Setting up a collection point (“VLO”) of existing teaching materials
- Development of teaching materials/ primers for central functions of CLARIN ERIC
- Encourage NCs to improve documentation and user guides of tools, resources, services; e.g., make sure that VLO encourages the user to add a link to a user guide

Next steps Group 2

Survey

- We feel there is no need for a new survey.

Content of CLARIN curricula

- Digital literacy course ideally at the earliest stage possible - course structure would cover key aspects of digital methods (it would resemble an introductory methodology or philosophy course) (See also “scholarly primitives” (Unsworth, 2010) , inspirational for CLARIAH: <https://at.virginia.edu/32nkRSh>)
- Some lessons in a methodology course targeting Humanities
- Content: concept of data set, tools, citation, replicability, legal aspects
- So generic that it enables collaboration in CLARIN ERIC
- More specific blocks would follow from research

Next steps Group 2

Role of National Consortia

- Using CLARIN materials in existing curricula + create new courses, programs + how to teach our students to use CLARIN (for existing material, benefit from information in CLARIN ERIC (question 2)
- Address researchers in Humanities, identify potential projects, offer national budget if possible - identify a potential “CLARIN ambassador” for a specific area in the Humanities
- Distribute teaching materials in the national infrastructure

Next steps Group 2

Role of CLARIN ERIC

- CLARIN ERIC should provide a repository for teaching materials
- Suggest also the basic-level courses and/or course building blocks adapted to certain research questions in specific areas in the Humanities, e.g.
 - Basic digital document encoding and organisation (TEI, XML etc.)
 - Basic LT processing (tokenisation, lemmatisation, NERC, semantic nets, SRL etc.)
 - Advanced LT and KT-processing (LOD linking, NE-networks, semantic graphs, etc.)
 - Have them translated/subtitled to other languages
- Enhance the DH Course Registry, possibly harvest information from VLO
- Better sharing of teaching material within CLARIN community
- Consider using the CLARIN mobility budget to implement collaborations (in research and/or teaching)

Next steps for NCs (underlined: where CLARIN ERIC can help)

- Understand the educational landscape in your country
 - identify key courses & teachers
 - build a contact list & network of teachers
 - promote & support exchange of experience and materials among teachers
 - support teacher exchange, both nationally and internationally
 - if possible, be part of new curricula development
- Organise training activities for teachers
 - discipline-specific works best
 - training for early-career teachers works best
 - share experience with other NCs as examples
- Provide user guides and training materials
 - improve documentation and user guides for tools, resources and services & share with other NCs
 - develop new course material for your flagship resources and tools and show how to teach it
- Dissemination
 - be visible at all major key national academic events
 - distribute material nationally & internationally (as examples, or as localisable material)
 - collect appealing use cases & share with other NCs

Next steps for CLARIN ERIC

- Technical development
 - add a facet to VLO for resources and tools that are particularly suited to be used in teaching
 - develop a facility to exchange teaching materials
- Development of teaching materials
 - develop teaching materials for general topics on RIs & central services
 - provide central support to prepare training activities for broader cross-border and cross-disciplinary use (e.g. beefing them up, making recordings, subtitling)
 - use mobility funds for collaboration on developing courses or course material
- Development of expertise
 - organize training activities for teachers & disseminate info about training activities offered by national consortia
 - add a track to report on teachers' experience in using CLARIN in their classes at the CLARIN Annual Conference
- Strategic activities
 - appoint a CLARIN Ambassador to promote the educational aspects of CLARIN
 - stimulate submissions for [Erasmus+](#) and [Marie Skodlovska ITN](#) calls

Survey follow-up

The survey was a pilot, should there be a next one?

- Pro: we might still find new useful information or ideas
- Con: not much new to be expected from small countries, too much work in large countries, especially if we go beyond the CLARIN sites
- Suggestions:
 - National consortia carry out the survey as required/feasible
 - Explain better what CLARIN content is
 - Only do it if we have new questions, such as
 - Use of existing e-learning platforms (or should we develop our own)
 - Asking requirements instead of status quo

Recommendations for teachers

General recommendations for the development of courses

- Try to use CLARIN material in your courses (and mark it as such)
- Piecemeal, next step only if previous step has been properly digested and worked with hands-on on relevant and authentic questions/tasks
- Make it manageable and don't underestimate the effort required for the development of teaching materials
- Realize that for every new element in the curriculum something (and often even someone) else has to go out

Open questions and caveats

For all material collections:

- Quality control hard to organise (but maybe not necessary if audience is experienced teachers)
- Obsolescence

Integrate CLARIN in general Digital literacy courses/Research methods corpuses or in specific courses?

- Option1:
 - Pro: Can be generic and easily shareable
 - Pro: Help preparing the grounds for real digital scholarship if organised at a very early stage in the curriculum
 - Con: Not problem driven and therefore less appealing for humanities scholars
- Option2:
 - Pro: Good bypass if option1 can't be achieved for whichever reason
 - Pro: Contextualized repetition of specific issues could be a good thing for the students when coupled to a relevant problem driven teaching unit
 - Con: Only limited attention to selected topics possible, cannot be comprehensive