Jost Gippert, Henning Lobin
D-SPIN WP 6
"Training and Education"

Educational Aspects: Teaching Materials and D-SPIN Summer School

Meeting of the D-SPIN Scientific Advisory Board

Munich, 17./18. Oct 2009

Outline

- D-SPIN WP 6 Introduction

- Part 1: Course Concept & Learning Scenarios

- Part 2: Technical Aspects, Standards & Desiderata

- Part 3: An Example Course (Giessen 2009)

- Part 4: D-SPIN Summer School 2010

- Final Picture: LRT Training and Dissemination

D-SPIN WP 6 Introduction

- Goals
  - Repository of modular course materials as reusable learning objects
  - Creation of e-learning materials
  - Implementation of e-learning materials in the presence and online phases
  - Development of a help desk concept for D-SPIN
- Focus
  - Corpus linguistics, empirical linguistics
- Staff
  - Svetlana Ahlborn, Frankfurt University
  - Frank Binder/ Cafer Travaci, Giessen University
- Supervisors
  - Jost Gippert, Frankfurt University
  - Henning Lobin, Giessen University
  - Erhard Hinrichs/ Kathrin Beck, Tübingen University

Part 1:
Course Concept & Learning Scenarios

Course Concept (Frankfurt 2009/10)

Learning Scenarios – The POP-Model

Module Access

Course concept

module	Frankfurt University, fall/winter 2009/2010 **Introductory course "Corpus linguistics"** **(level: beginners; module access: successive)**
1	Development of corpora
2	Types of corpora
3	Metadata: usage and creation
4	Corpus annotation: usage and creation
5	Markup languages
6	Corpus search and retrieval software
7	Concordancer: application
8	Applications of corpus-based analysis
9	TITUS
10	COSMAS
11	TiGerSearch und TiGerKorpus
12	DoBeS

Learning scenarios: Learning unit

■ One learning unit consists of a combination of presence and online phases[1]:

presence phase	online phase	presence phase	*learning phases*
P	O	P	*learning unit (POP$_n$)*

---

[1] Kerres, M. Multimediale und telemediale Lernumgebungen : Konzeption und Entwicklung 2., vollst. überarb. Aufl., München [u.a.] : Oldenbourg, 2001
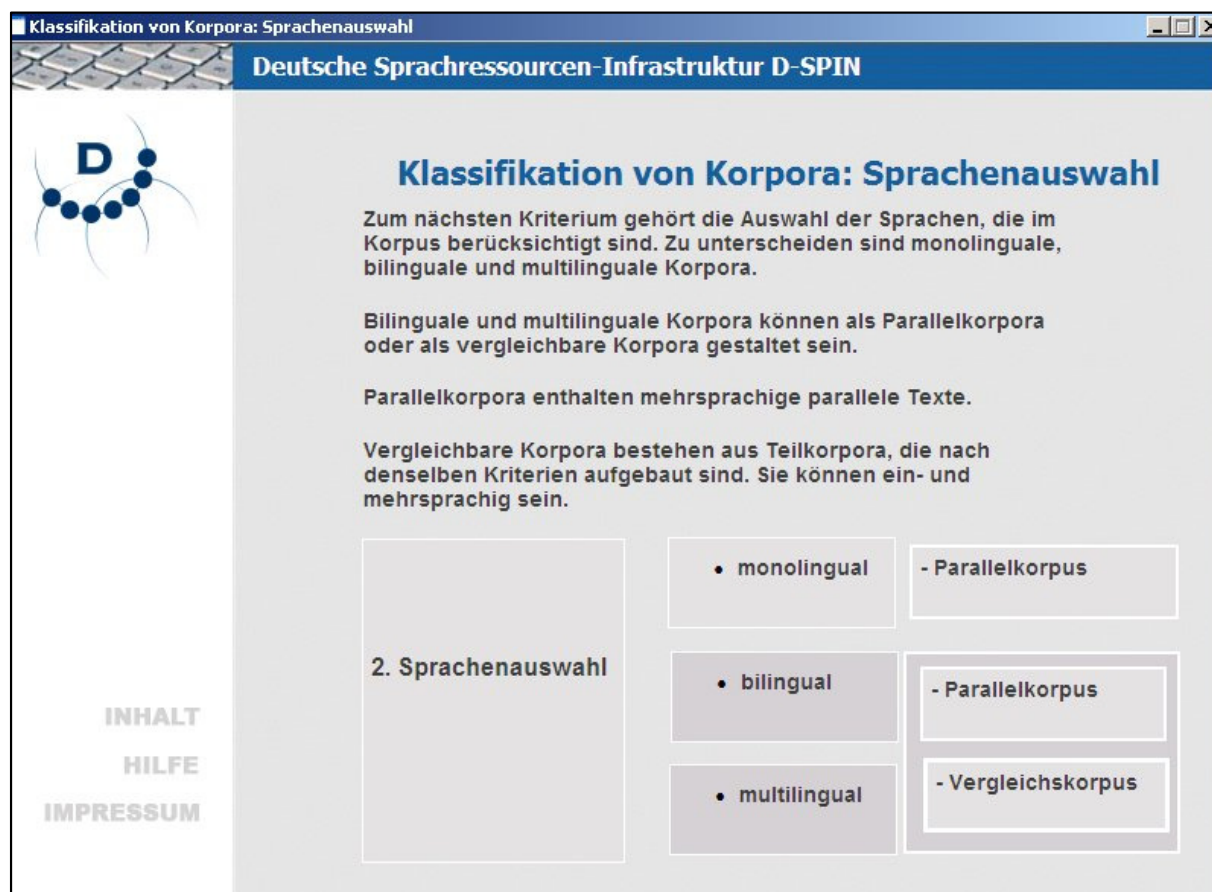
Learning scenarios: The POP model

- The online phase is combined with the previous and the following face-to-face lesson in the classroom. The preceding presence lecture should include feedback in order to support the usage of online materials:

presence phase	online phase	presence phase	*learning phases*

P	O	P	*learning unit (POPn)*
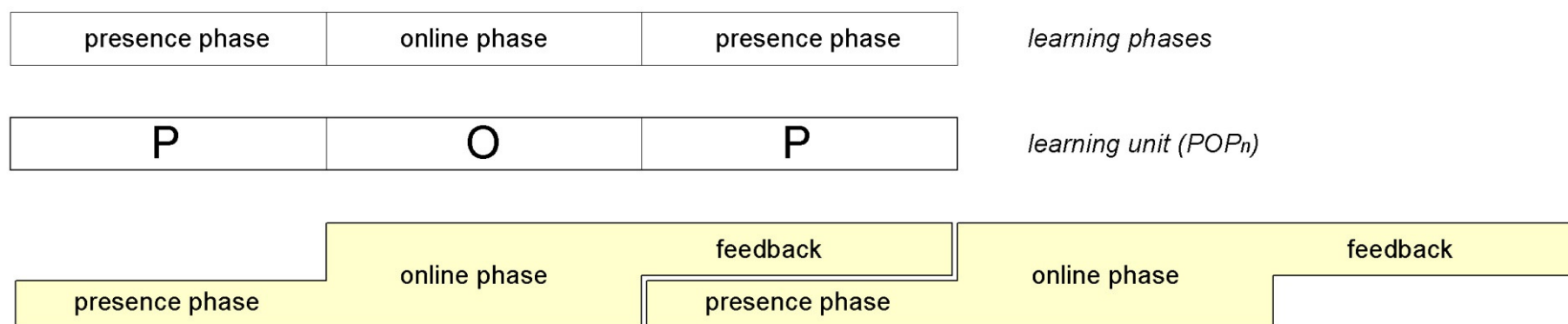
Learning scenarios: Phases (1)

■ Learning phase: theory
(presence and online phases)

JUSTUS-LIEBIG-
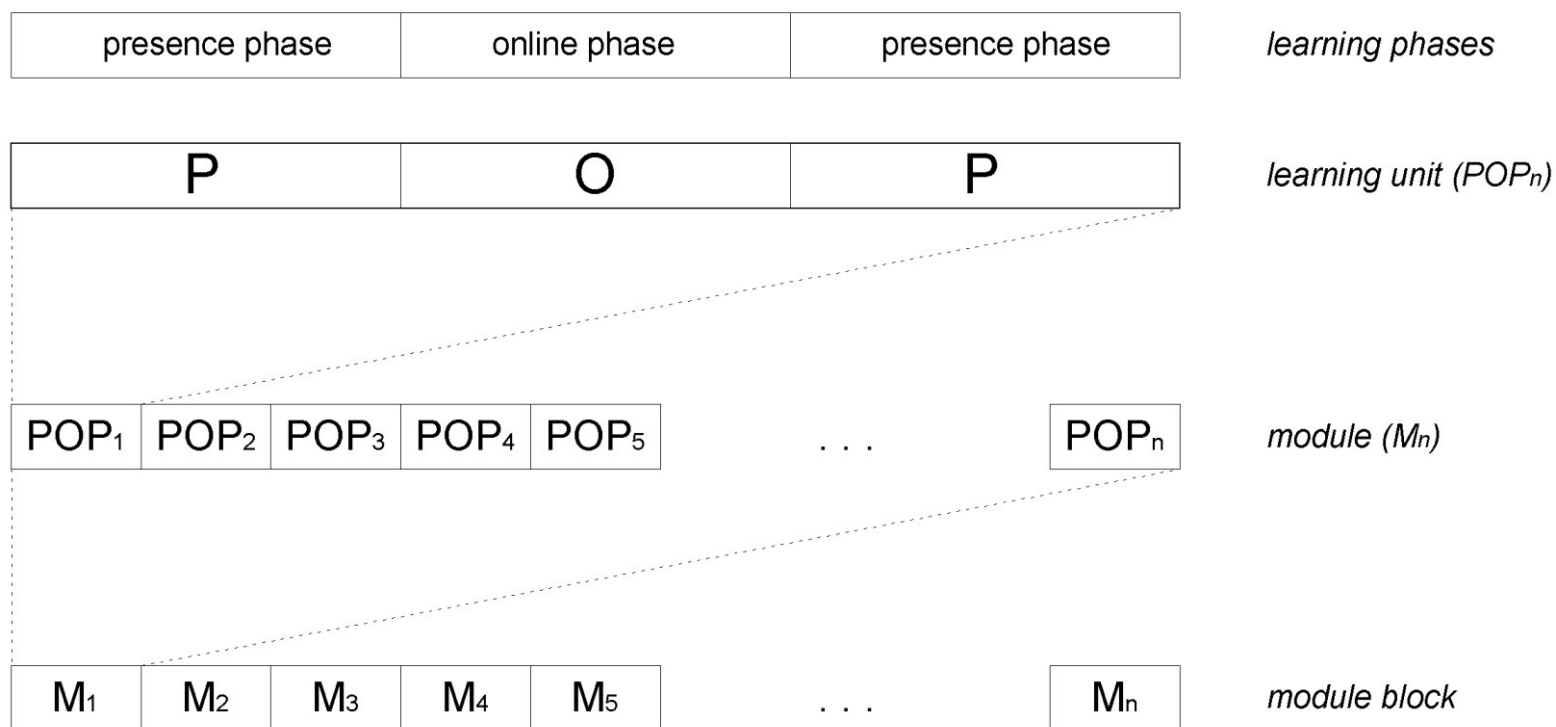UNIVERSITÄT
GIESSEN

D-SPIN
CLARIN

GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN

Learning scenarios: Phases (2)

- Online Phase

- Feedback to the learner

"Learning unit" and "module"

- One module consists of several dynamically growing learning units

presence phase	online phase	presence phase	*learning phases*
P	O	P	*learning unit (POP_n)*

| POP_1 | POP_2 | POP_3 | POP_4 | POP_5 | . . . | POP_n | *module (M_n)* |

| M_1 | M_2 | M_3 | M_4 | M_5 | . . . | M_n | *module block* |

Module access approaches (1)

- Successive access approach

JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

D-SPIN
CLARIN

GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN
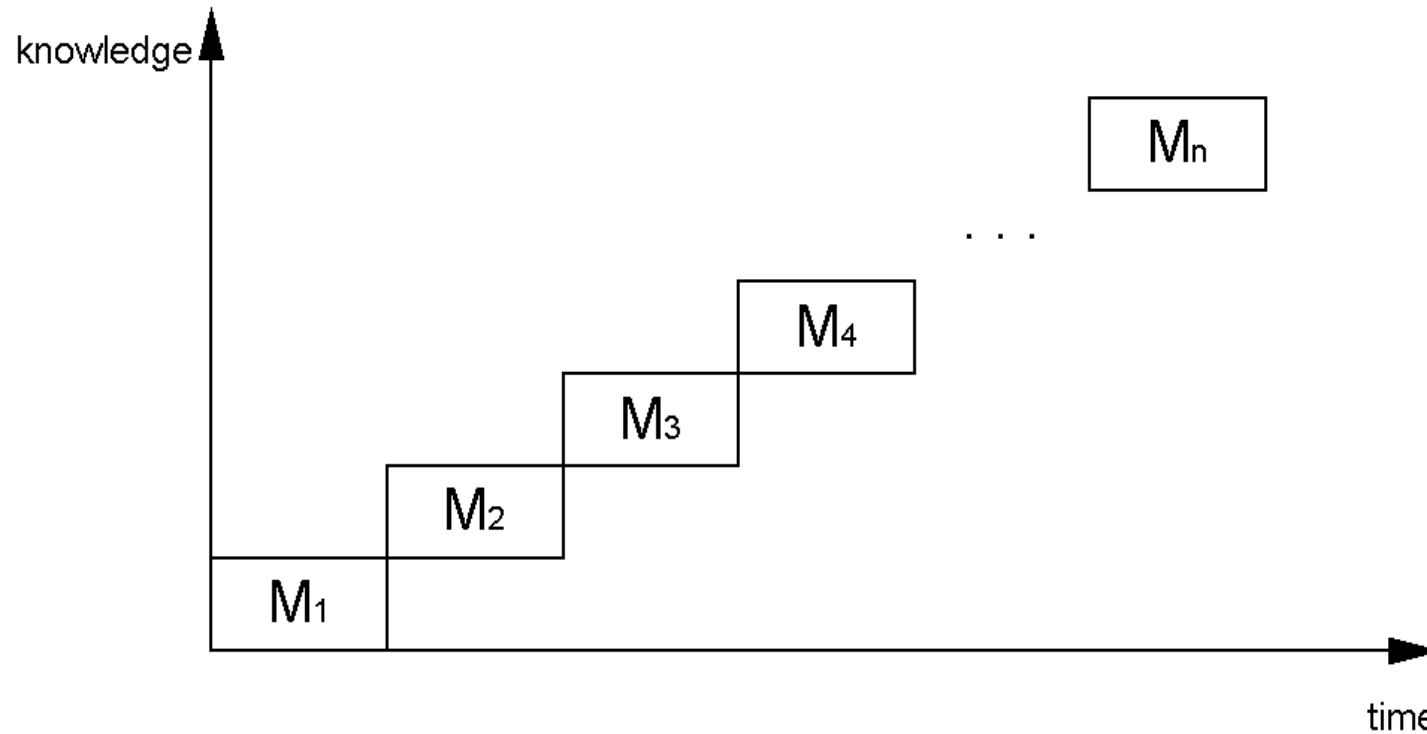
Module access approaches (2)

- ## Successive access approach
  - determines the basic knowledge
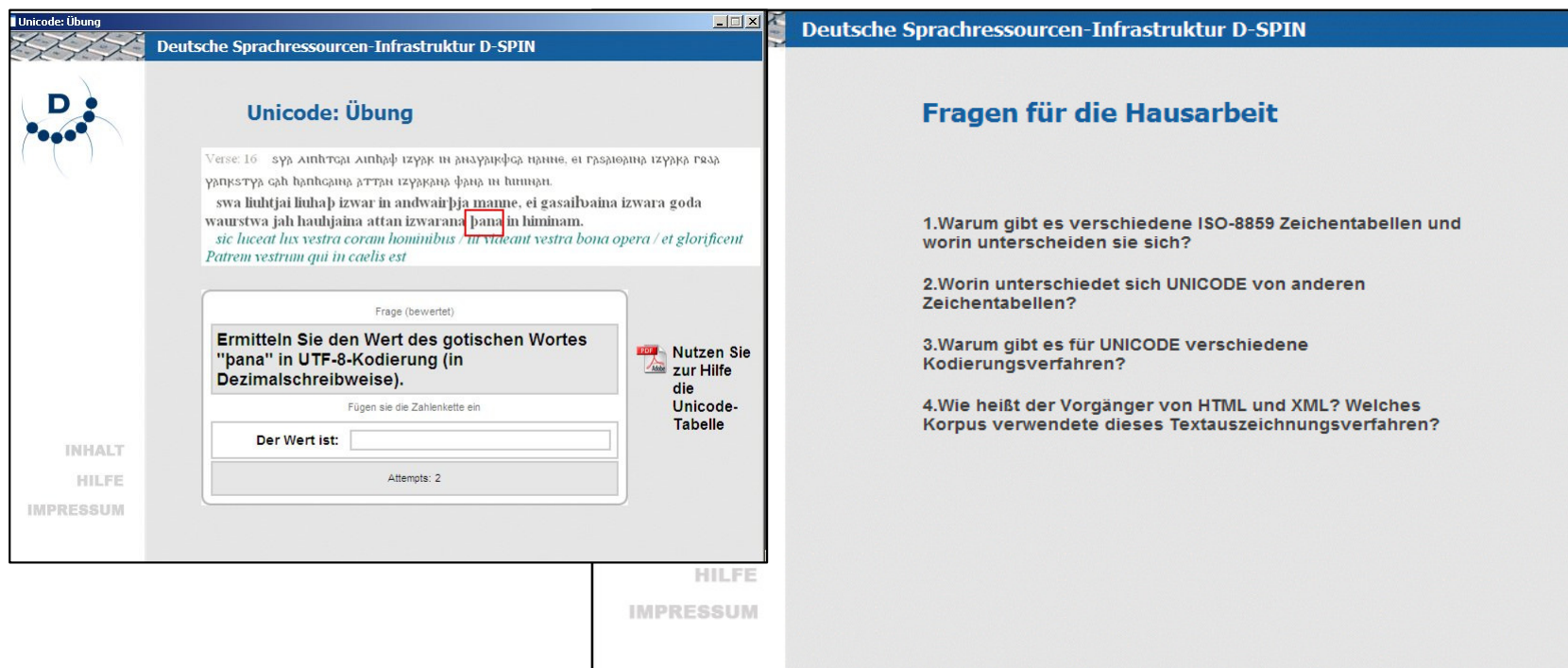
Module access approaches (3)

- ## Successive access approach
  - obligatory examination at the end of each module

Module access approaches (4)

- Random access approach

JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

D-SPIN
CLARIN

GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN
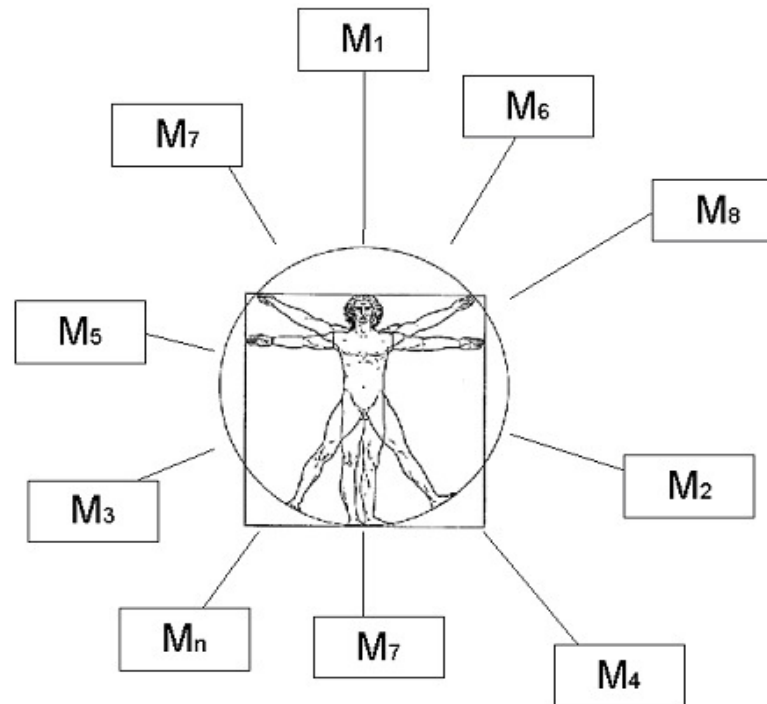
Module access approaches and student levels

student level	module access approach	
beginners	successive access	
advanced students	successive access + random access	
special purpose students	random access	

Part 2:
Technical Aspects, Standards & Desiderata

Standards & Technical Aspects
SCORM
Learning Management Systems (LMS)
Authoring Tools
Desiderata

Standards & Technical Aspects

- Standard-compliant e-learning packages
  - Sharable Content Object Reference Model (SCORM)

- E-Learning Platforms (with SCORM support)
  - ILIAS
  - Moodle

- Authoring-Tools:
  - CourseLab

E-learning standards: SCORM

- SCORM – Sharable Content Object Reference Model – combines several industry standards

- Provides a unified learning content model

- Allows content aggregation

- Addresses high-level requirements:

    - Accessibility from & delivery to multiple locations
    - Adaptability   – to satisfy differing user needs
    - Durability   – based on open standards
    - Interoperability   – develop on one system, run on another
    - Maintainability   – support content evolution
    - Reusability   – use components in multiple contexts

An Introduction to SCORM 2004 3rd Edition
http://projects.aadlcolab.org/scourse/2004_3rdEd/_viewer/index.html

Learning Management Systems (LMS)

- ## E-Learning Platform compatibility with SCORM
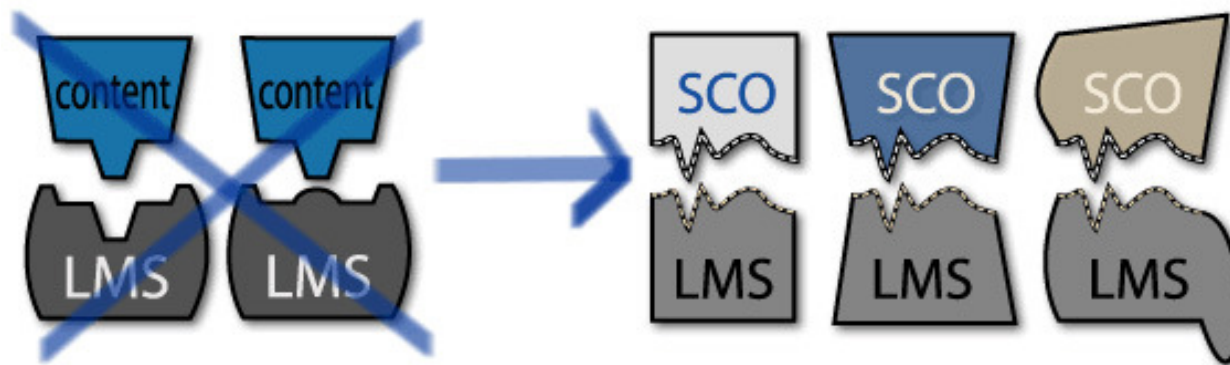
- LMS without SCORM support
- LMS with SCORM support

  - ILIAS
  - Moodle
  - … and many others

An Introduction to SCORM 2004 3rd Edition
http://projects.aadlcolab.org/scourse/2004_3rdEd/_viewer/index.html

Authoring Tools

- Several free & commercial off-line authoring tools were considered

- Currently using **CourseLab 2.4**
  - Free of charge, but closed-source
  - WYSIWYG – easy to use
  - Well documented & intuitive
  - **Supports export as SCORM-packages** (zip-Files)
  - Resulting packages also work off-line (without LMS)

- Some technical issues – but they are solved:
  - Print version of modules –
    now available through a patch by Caver Travaci (Giessen)
  - Cross-Browser compatibility – workarounds available

- Authoring Pipeline: Off-line editing → Module repository → LMS

- Other tools may be used at a later stage

http://www.courselab.com/

Desiderata

- ## Easy online usage of language resources

  - additional software installation (e.g. JRE)

  - access rights for each student

  - problems accessing corpora online

A learning module
http://moodle.studiumdigitale.uni-frankfurt.de/moodle

– Multimodal corpus online
http://corpus1.mpi.nl

Part 3:
Dissemination through university curricula: An example course at Giessen University

Overview
Prerequisites
Course schedule
Elements of the course
Selected LRTs
Résumé & Outlook
Suggestions

An example course: Overview

- Goal: To educate interested potential users

- Ideas: Offer course as optional part in common curricula;
  Use existing resources – as readily available

- Summer term 2009 (Apr-Jul) – F. Binder – Giessen University

- "Einführung in die Arbeit mit computerbasierten Sprachressourcen"

- 12 sessions – 22 students

- 80% of students had no need for/no interest in programming

- Mixed course for several target groups:
  - Bachelor students in German studies (2nd year): 10
  - Pursuing teaching certification (2nd to 3rd year): 6
  - Master students (1st year) / Magister (3rd year): 6

- PC-equipped course room (10 PCs)

An example course: Prerequisites

We are fortunate to have

- Textbook(s) in German – e.g. Lemnitzer & Zinsmeister 2006; Carstensen et al. 2004

- PC-equipped course room with very supportive admins

- Some prototypical language resources of interest that are
  - well documented
  - easy-to-access and
  - easy-to-use (for beginners) & efficient-to-use

- Local E-Learning infrastructure team[1] – they provide:
  - **Stud.IP** for course managment[2], **ILIAS** for hosting SCORM modules[3]
  - Training for course instructors / e-learning authors
  - Evaluation service: Online student questionnaires

An example course: Outline

Session	Session topics – "LRT for linguists" – Giessen, Summer 2009
1	Introduction and overview
2	Empiricalists, Rationalists & Language Data
3	Basic corpus linguistics – search & concordances in remote data
4	Basic corpus linguistics – concordances with local data
5	Quantitative corpus work: Collecting and organizing data
6	Quantitative corpus work: Statistics continued & advancing
7	Annotation: Using Annotation (in remote data)
8	Annotation: Using & Creating Annotation (in local data)
9	Lexical Ressources
10	Ressources for educational purposes & language teaching
11	Legal issues, copyright & licences
12	Final discussion

An example course: Elements

- 5-min presentations of software, tools, resources
  – by students, obligatory for all participants

- 30-min lectures/talks on session topics with group tasks
  – mostly by students, required for graded certificate

- Wide range of topics: Technical, non-technical, theory, …

- Readings – text books, papers/studies

- E-Learning assignments – 3 phases conducted

- Practical tasks for individual and group work

- Preparing group discussions using a WIKI:
  "Role-play" the use of LRT in academics, publishing, industry & education
  to identify relations, issues & potential

- Scientific term papers for graded certificate

An example course: Selected LRTs

- **Selected web resources**
  - Portals and online concordancers / dictionaries
    (e.g. DWDS – BBAW Berlin; Cosmas/DeReKo – IDS Mannheim;
    "Deutscher Wortschatz" – Uni Leipzig;)
  - Downloadable reference corpora (plain text) – Uni Leipzig –
    for use in contrastive analysis, e.g. keywords with AntConc
  - Downloadable texts & corpora to be studied,
    e.g. Project Gutenberg, Deutscher Bundestag, …
  - WordNet, (GermaNet?), Wikipedia, …

- **Selected desktop tools with graphical user interface (Windows XP & Vista)**
  - AntConc – concordances & word lists
  - TreeTagger – as a processing tool for automatic annotation
  - TigerSearch – as a retrieval tool
  - Office & gnuplot – managing data & visualization

An example course: Résumé & Outlook

- Received interesting term papers – for example about:
  - Corpus work in the German classroom?
  - Reviews /comparisons of corpus-based studies
  - Legal issues and licensing
  - CL-topics like WSD, annotation formats

- Our e-learning experience – supported by student questionnaires
  - Learning Management Systems are very helpful:
  - WIKIs, mailing lists, and file repositories
    are the most frequently used features
  - Some delay due to technical issues, but overall successful
  - E-Learning materials need to be complemented; Students prefer practical
    tasks – do more individual or group project work

- Next & different course is starting: Winter 09/10 in Frankfurt (S. Ahlborn)

An example course: Suggestions

- Step 1: Offer such elective courses for interested potential LRT *users* at basic levels of common curricula in the "D-SPIN / CLARIN users' community"
  (general linguists, HSS, academic teacher training etc.)
  – to spread the word

- Step 2 – *later*: promote some participants to specialized degree programs
  – to train people who are likely to spread the word

- Our current job: Provide materials for courses at step 1

Part 4:
D-SPIN Summer School 2010

Overview
Placing our offer
Goals and plans
The CLARIN picture

Summer School: Overview (1)

- 30. Aug – 03. Sept 2010 (Mo-Fr)

- Bad Homburg –
  Forschungskolleg Humanwissenschaften

http://www.forschungskolleg-humanwissenschaften.de/
http://www.forschungskolleg-humanwissenschaften.de/downloads/Flyer.pdf

Summer School: Overview (2)

- 35 students in up to 3 parallel sessions

- Low fees for students –
  2 packages conceivable:
  - 120 Euro total for "all incl." – full board, accommodation & some travel support
  - 35-50 Euro total for courses & catering, not incl. accommodation & travel

- Compensations for course instructors
  - Reimburse travel & accommodation expenses
  - Plus some amount paid per session (external experts only)

- Total funding estimate: 11.000 ~ 14.000 Euro (less than 1% of D-SPIN budget)

http://www.forschungskolleg-humanwissenschaften.de/
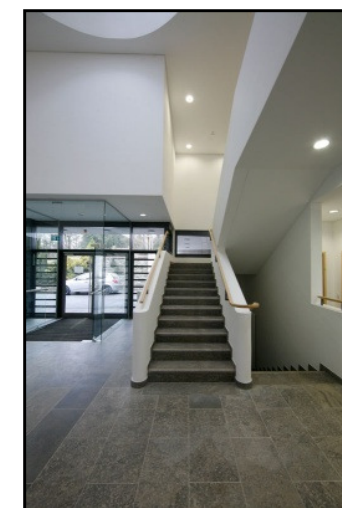
Summer School: Placing our offer (1)

From the students' point of view:

- Consider the following academic offers:

  University programs   **Summer Schools**

- Consider 3 (related) factors: Level of target audience; scope; course language

- Ignore (for now) topical differences

- Focus on potential LRT-users from HSS
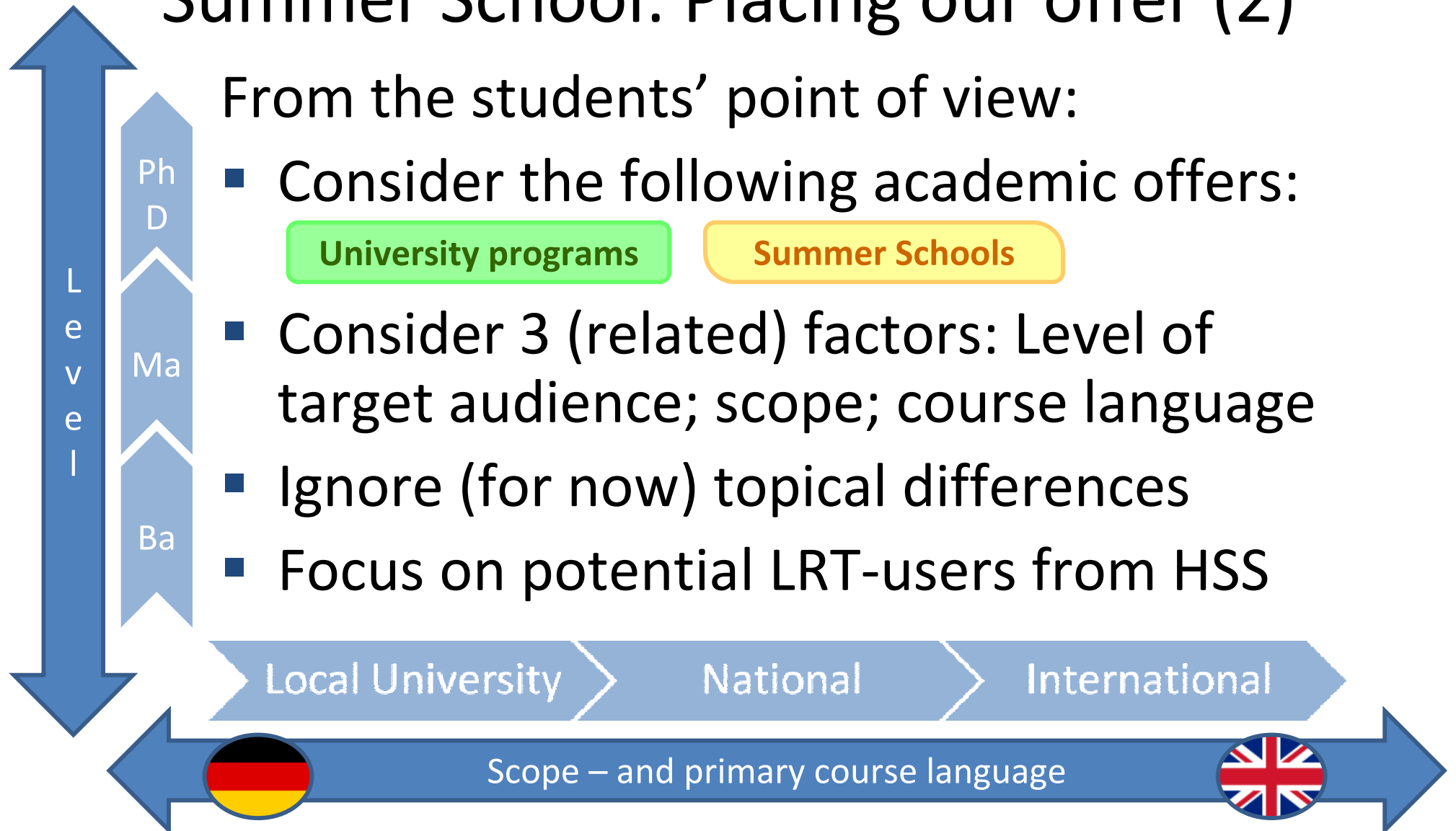
Summer School: Placing our offer (2)

From the students' point of view:

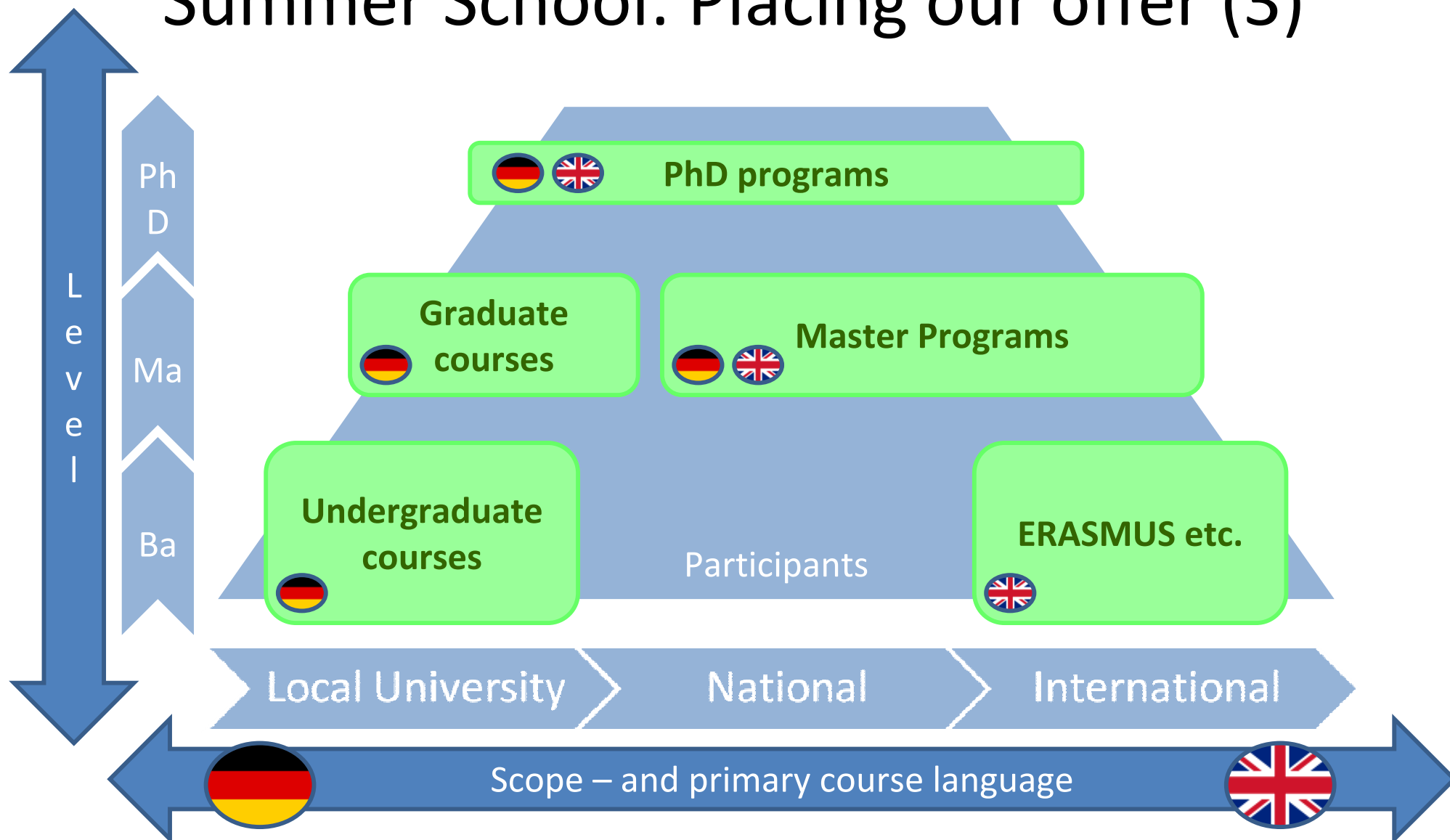- Consider the following academic offers:
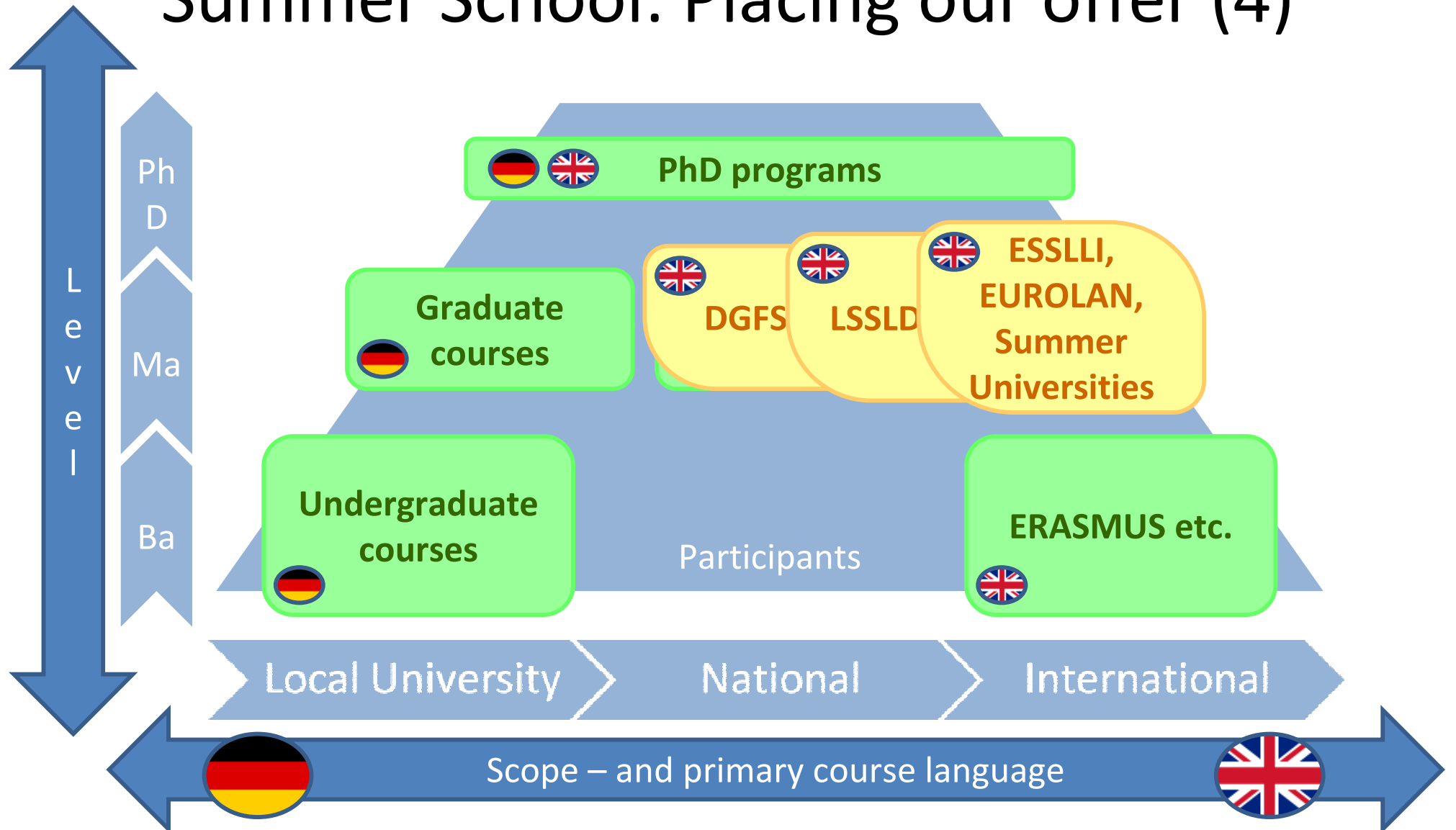
  University programs   **Summer Schools**

- Consider 3 (related) factors: Level of target audience; scope; course language

- Ignore (for now) topical differences

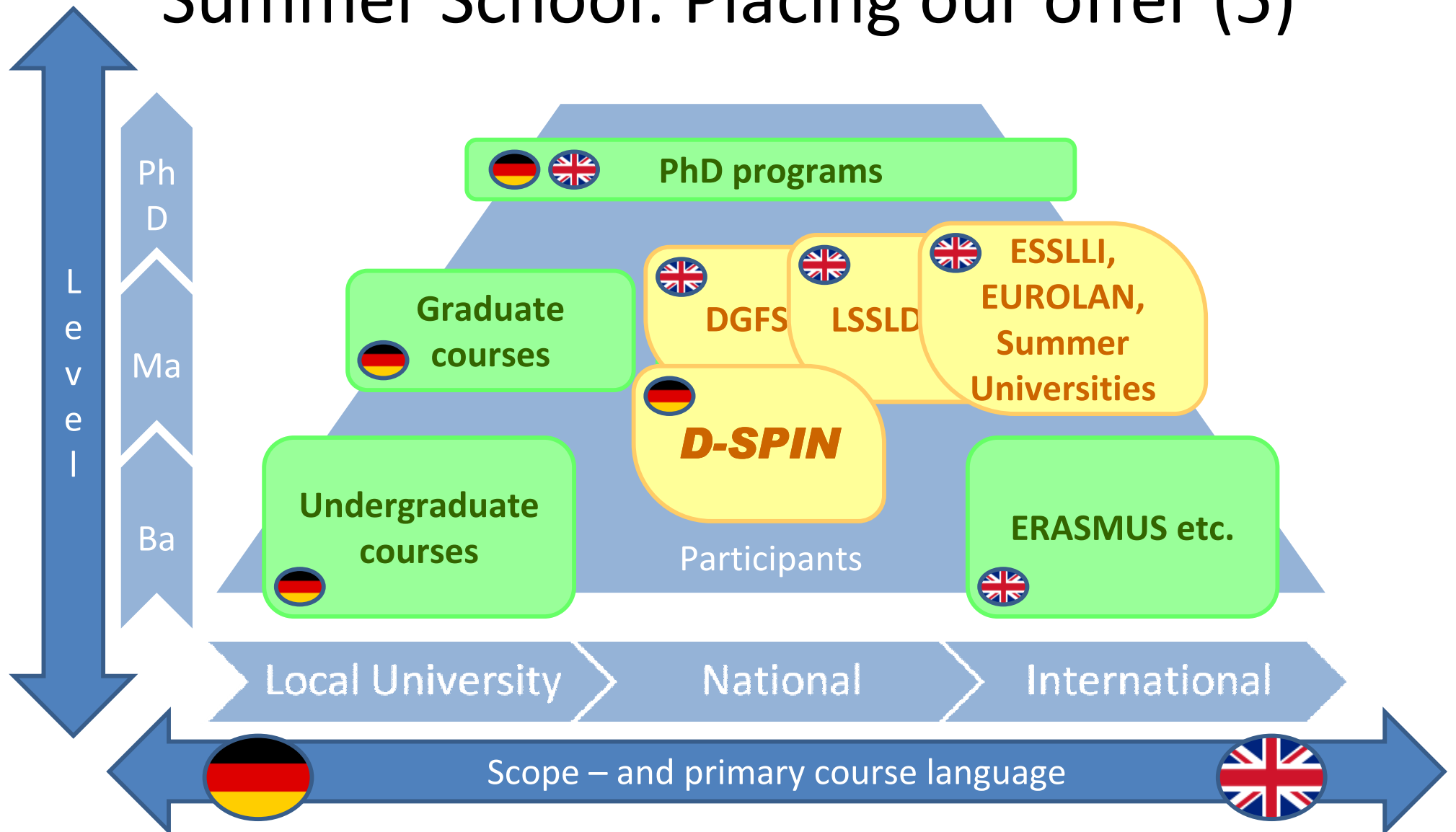- Focus on potential LRT-users from HSS

Level

PhD

Ma

Ba

Local University > National > International

Scope – and primary course language

Summer School: Placing our offer (3)

Summer School: Placing our offer (4)

Summer School: Placing our offer (5)

Summer School: Goals and ideas (1)

With regard to student participants - we plan to

- Require little prior knowledge & no prior experience
- Attract post-2nd-year students (Ba & Ma) – before their projects & theses
- Be interdisciplinary – invite any potential LRT user
- Keep costs & fees for students low

JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

D-SPIN
CLARIN

GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN

Summer School: Goals and ideas (2)

With regard to course offers - we plan to

- Offer (and demand) very practical courses (after some course-specific hands-off intro)

- Be open to the community – invite course instructors from D-SPIN and the community

- Use German as primary course language, national scope

LRT – Training and Dissemination: Towards a larger picture around CLARIN

LRT – Training & Dissemination:
Towards a larger picture around CLARIN (1)
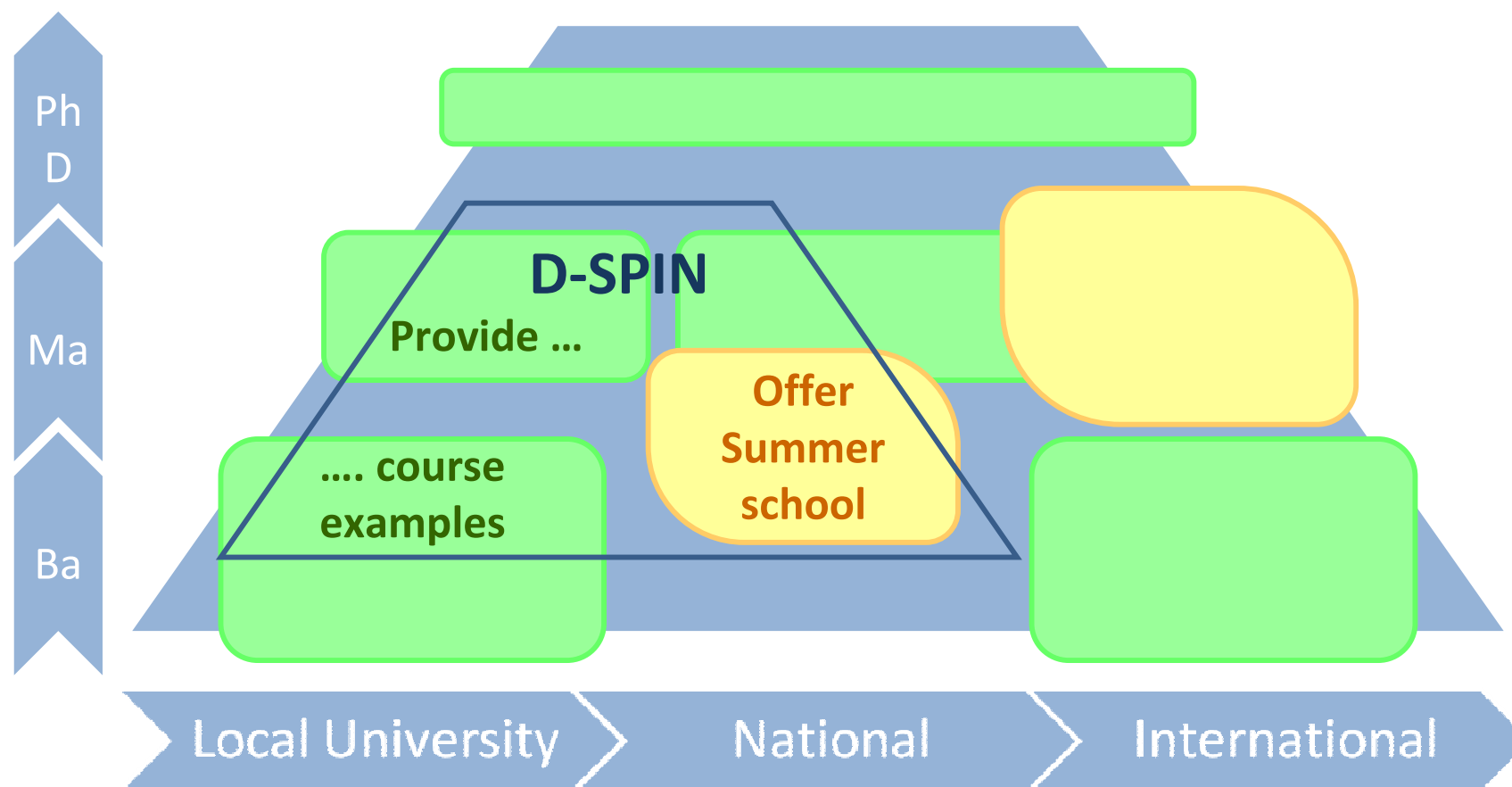
LRT – Training & Dissemination:
Towards a larger picture around CLARIN (2)
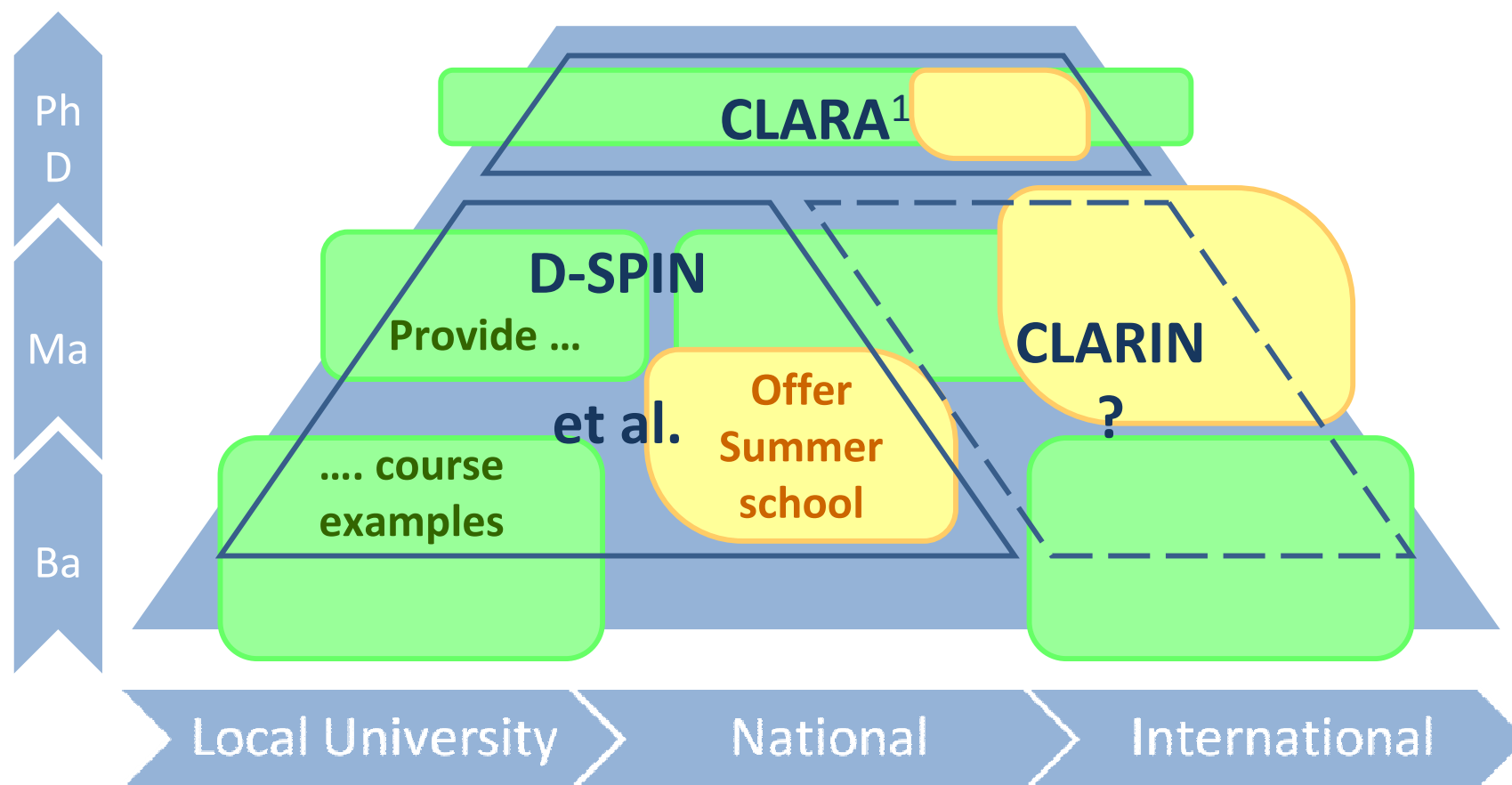
[1] CLARA Training Program – a Marie Curie ITN
http://clara.uib.no/training-program/

WP 6 Summary – Our Focus

- Develop course materials & example courses
  - To **educate potential users** of LRT
  - Integrate LRT dissemination into general university curricula

- Offer e-learning materials
  - As **standard-compliant packages** for use in LMS
  - Easy to maintain and to adopt; also usable offline

- D-SPIN Summer School 2010
  - National scope, courses in German
  - **Open to the community**
  - Attract post-2nd-year students (Ba/Ma)
  - Offer **practical tutorials** with concrete resources & tools
  - As a complement to university programs