

# Towards Fair Licenses for Data from the DGS Corpus Project

Thomas Hanke, Sung-Eun Hong, Susanne König, Gabriele Langer, Rie Nishio, Christian Rathmann  
University of Hamburg, Institute of German Sign Language and Communication of the Deaf

### GENERAL INFORMATION

#### Data Collection

- 2009-2011
- 12 locations (mobile studio)
- 328 informants balanced for sex, age, region
- Studio session:
  - 2 informants (peer-to-peer situation)
  - Moderated by Deaf contact person
  - Duration: one day (approximately 5.5 hours of filming plus breaks)
- Elicitation tasks (about 20 different tasks)

#### The Public Corpus

- Selected parts of the reference corpus (approx. 50 hrs) will be made publicly accessible (including English translation and basic transcription/annotation)

### Informants Recruiting Events

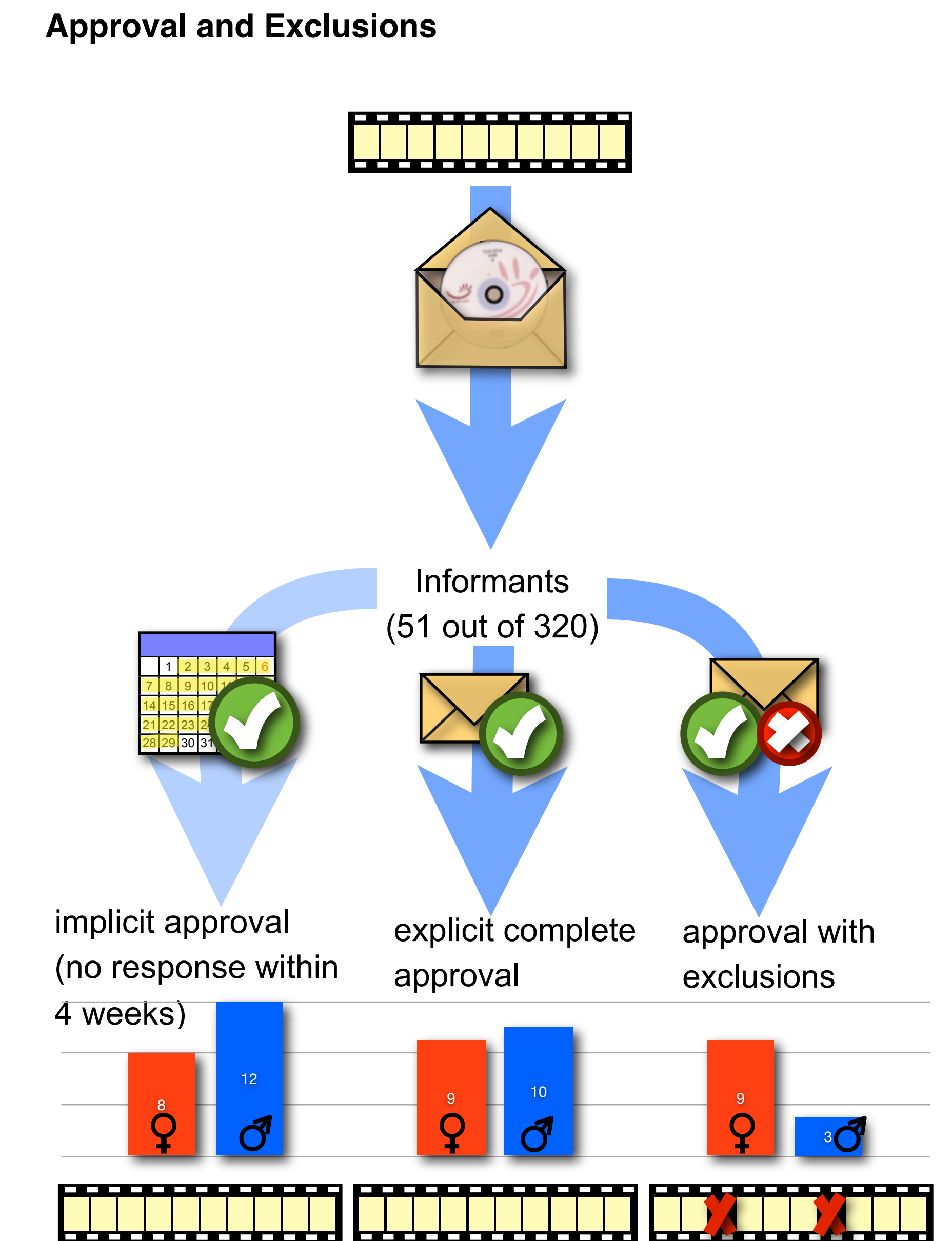
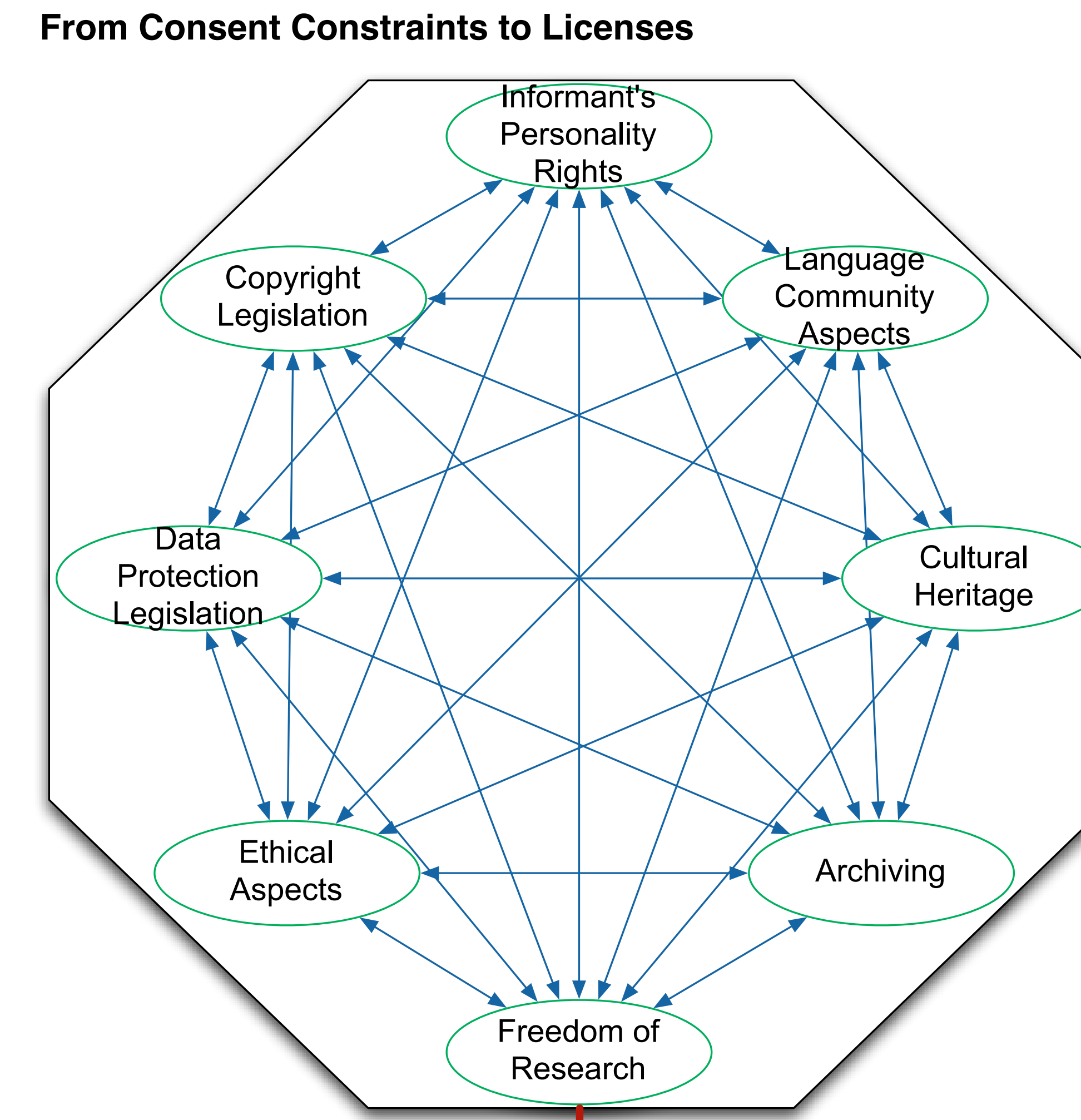


**What will happen to the recorded data?**  
...and who will be able to see them? (1/2)

- ▶ Different steps of processing
- ▶ Translation → interpreters
- ▶ Transcription → student workers and deaf project researchers

**What will happen to the recorded data?**  
...and who will be able to see them? (2/2)

- ▶ Linguistic analysis → researchers of the project
- ▶ Feedback on the internet → public
- ▶ Isolated signs and example sentences in the electronic dictionary → all users
- ▶ Selected part of the corpus (part of the signed movies and their transcriptions) → all interested people



### Informed Consent

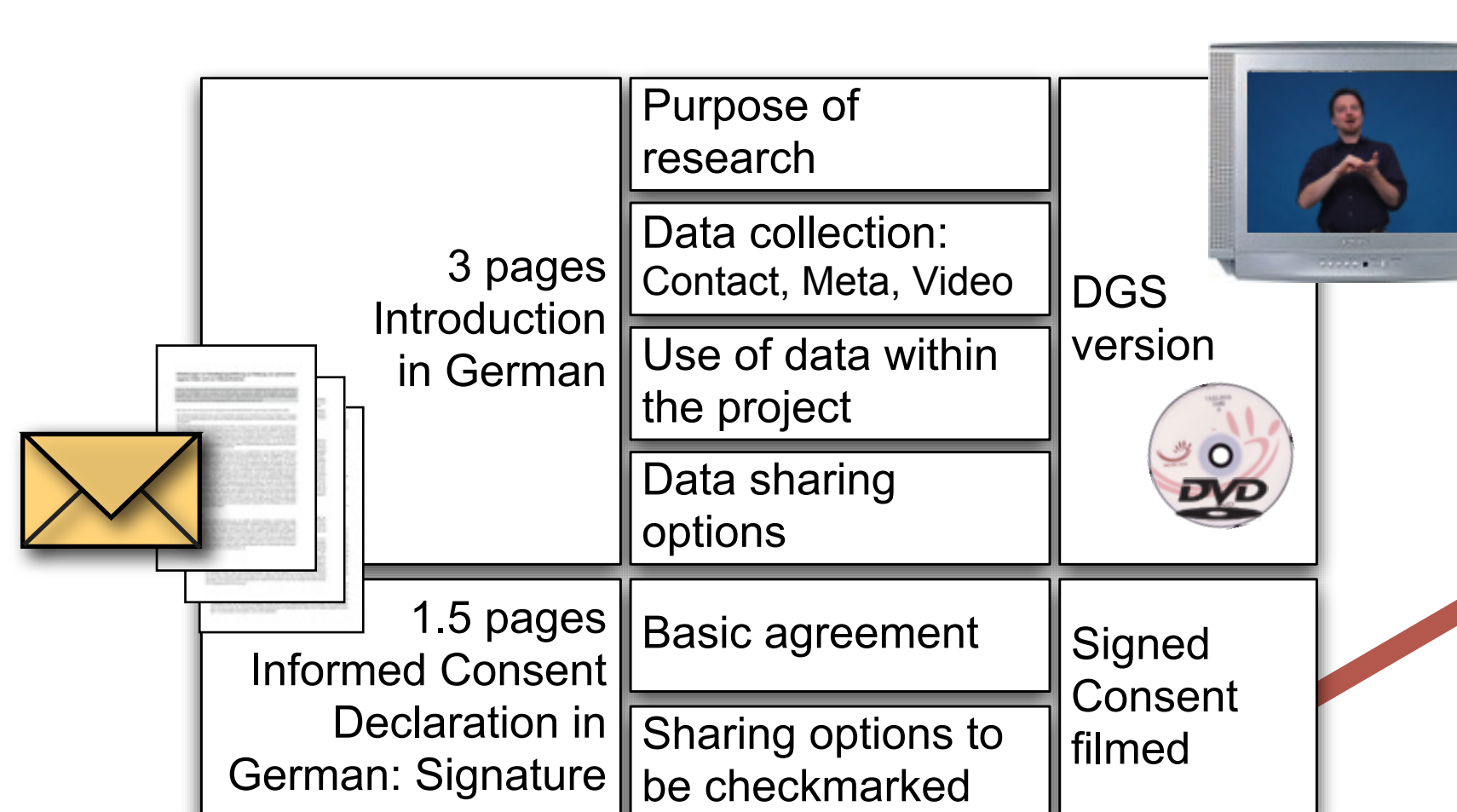
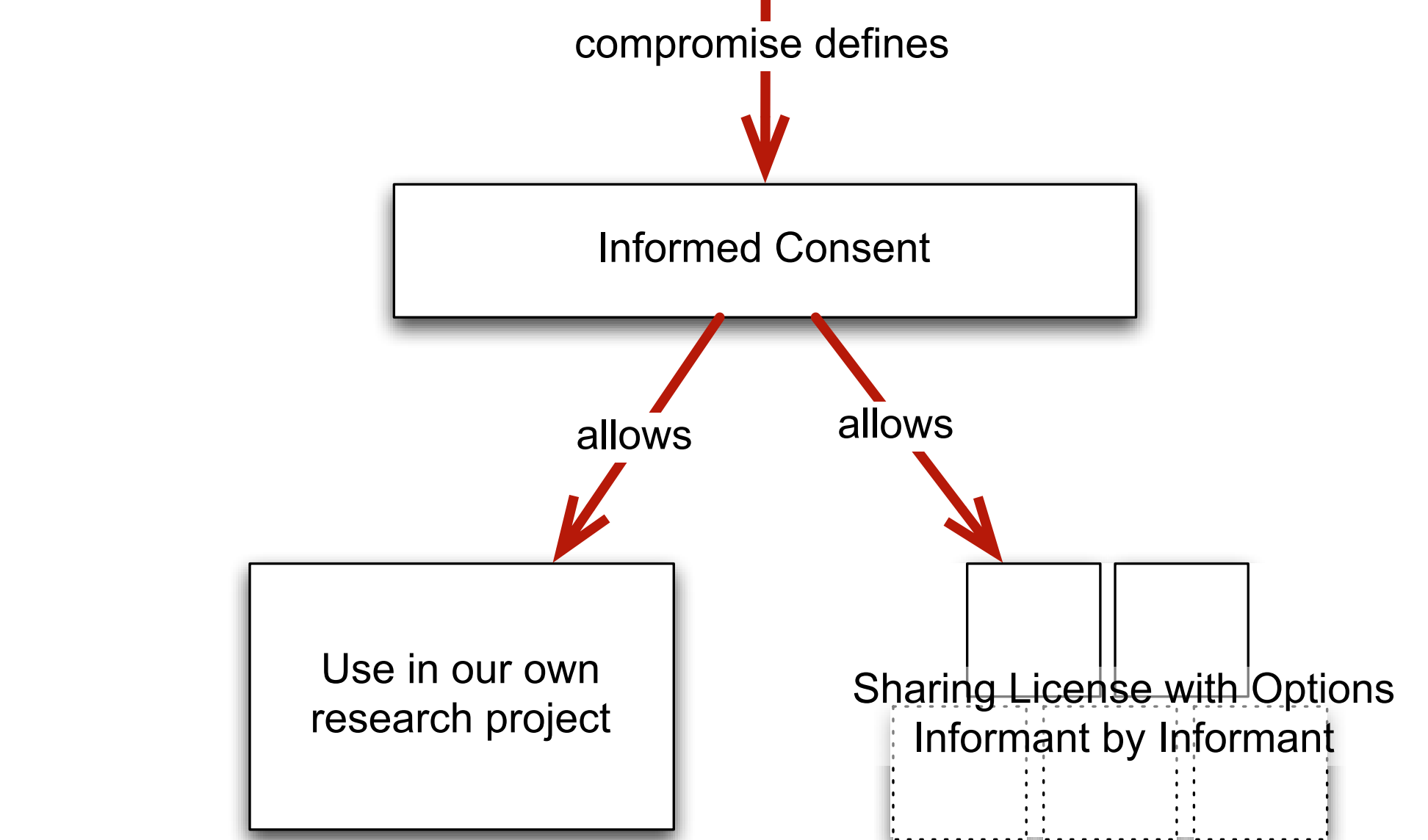
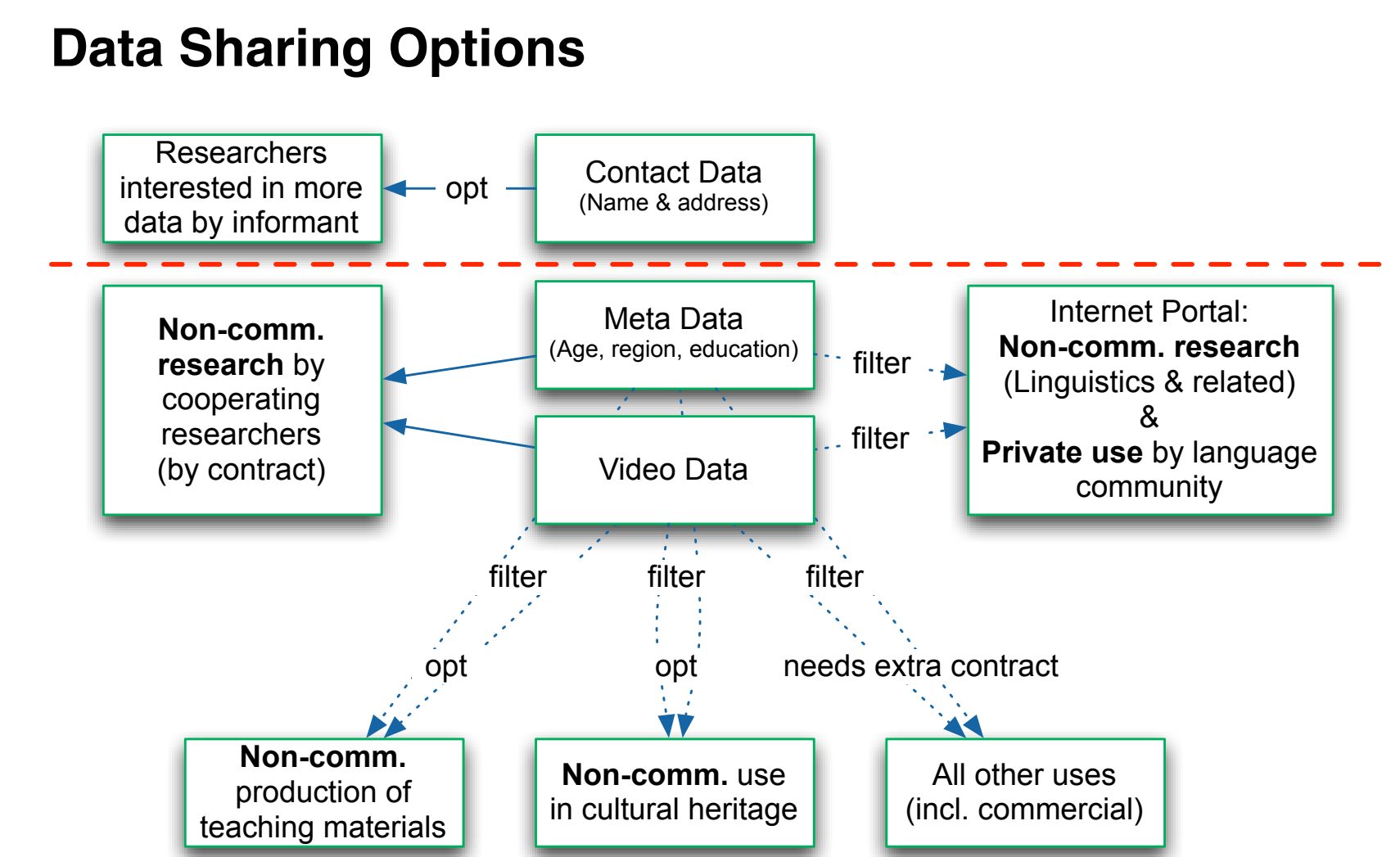


Diagram illustrating the informed consent process. It shows the flow from **3 pages Introduction in German** and **1.5 pages Informed Consent Declaration in German: Signature** to **Basic agreement** and **Signed Consent filmed**. The process involves **Data collection: Contact, Meta, Video** and **Use of data within the project**.

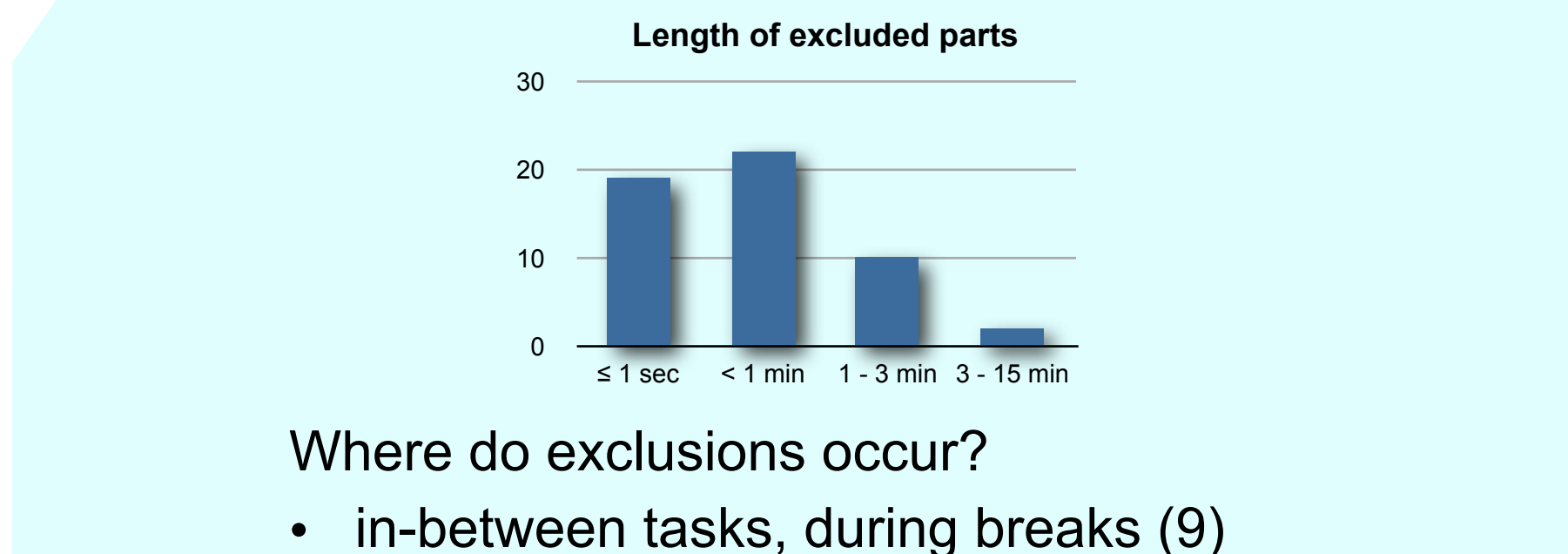
### DGS Consent Procedure



Diagram illustrating the DGS Consent Procedure. It shows the flow from **On the day of filming...** to **Do you have any questions?**, **We also would like to film you sign the consent.**, and **Please watch the movie clip.** The process involves **1. I've read and understood the consent form.**, **2. I agree to be filmed.**, and **3. I agree that my film data will be published.**



Total number of exclusions: 60  
Total time of excluded material: 0:48 hrs of estimated 125 hrs material (0.64 %)



Bar chart showing the length of excluded parts. The x-axis represents the length of excluded parts in minutes, and the y-axis represents the number of excluded parts.

Where do exclusions occur?

- in-between tasks, during breaks (9)
- within "free conversation task" (23)
- within "Deaf events task" (8)
- other tasks (20)

Reasons for exclusion:

- Content-related:
  - Names, name signs (22)
  - Opinions on other persons, groups, topics
  - Information about other persons, institutions
  - Information about own person, workplace
  - Private conversation in between tasks
  - other
- Related to studio situation (e.g. questions on task, feedback comments)
- Non-linguistic reasons (e.g. blowing nose, fixing clothing)

Poster presented at the Workshop on Exploitation organized by the Sign Linguistics Corpora Network, held in Berlin, Germany Dec 3.-4., 2010.  
The research leading to these results has received funding from the German Academies of Science Programme.

### License Options Chosen (122 out of 320)

