Corpus à la carte

Improving Access to the Public DGS Corpus

Reiner Konrad, Thomas Hanke, Amy Isard, Marc Schulder, Lutz König, Julian Bleicken, Oliver Böse
Institute of German Sign Language and Communication of the Deaf, University of Hamburg, Germany

• Antipasti •

Public DGS Corpus

50+ hours of data from the DGS Corpus, curated for public release. Contains videos, annotations, translations and pose information.

Release 4

- New portal
- New data
- New features
- Interconnectivity

NEW

4 public portals

Each portal meets different needs. Release 4 provides direct connections between the portals.

- *MY DGS:* Community Portal
- *MY DGS – annotated:* Research portal
- *MY DGS – ANNIS:* Web-based corpus search
- *MY DGS – iLex:* DB for use with iLex NEW

• Primi •

New Resources

Transcripts

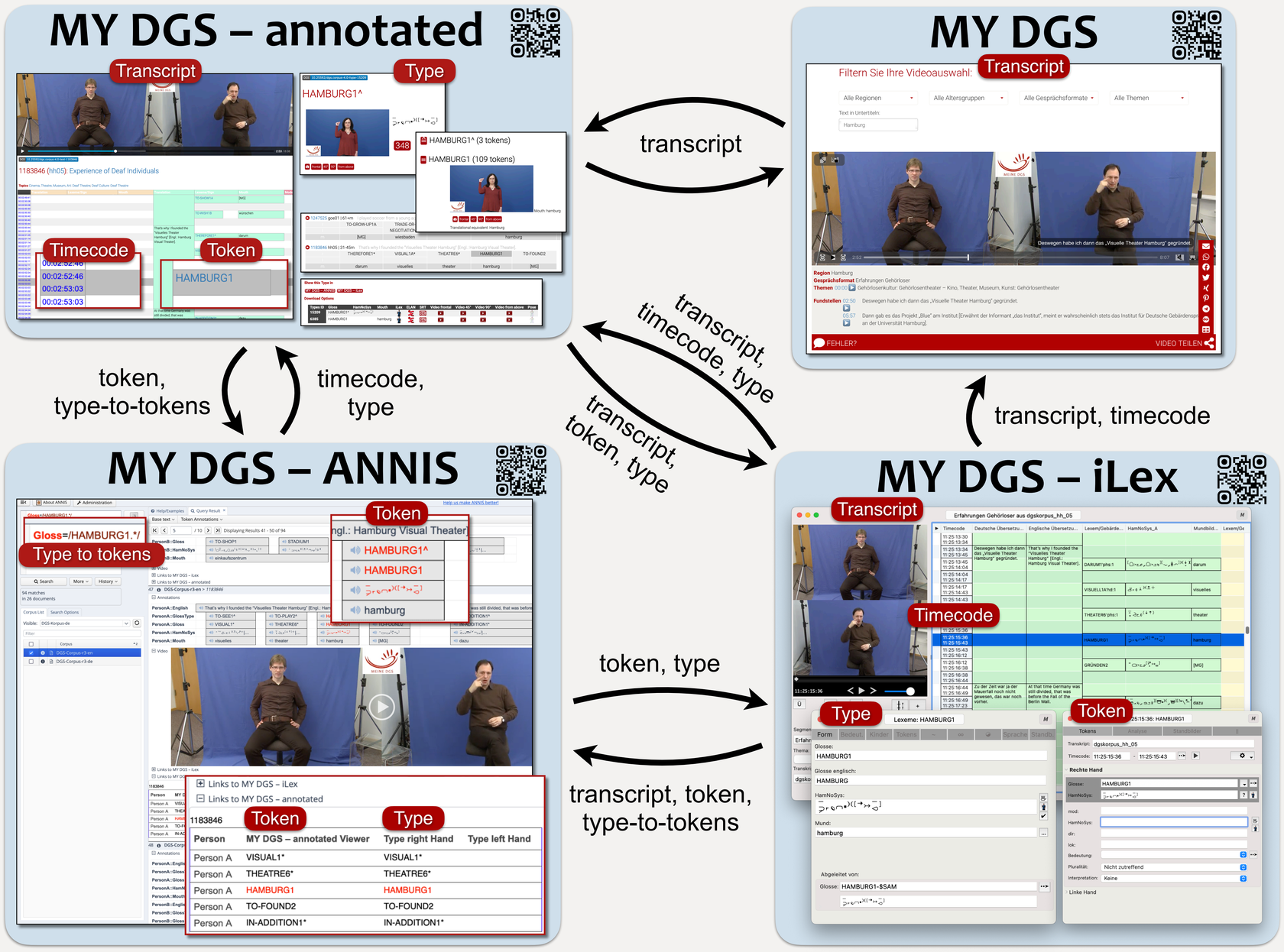
One hour of new recordings: 18 process descriptions (53 min) and 4 retellings (7 min)

Type information

- Translational equivalents for subtypes
- Videos and HamNoSys for subtypes
- Download files for types and subtypes

Pose information

In addition to OpenPose now also provides model estimates by MediaPipe, Apple Vision and Ivashechkin et al. (2023).



Serving suggestion

• Secondi •

New Features

Interconnectivity

- Easily jump to the same transcript/token/type in other MY DGS portals (see diagram above)

MY DGS

- Filter/Search by topics and over 570 keywords
- Full-text search of German/English translations, including wildcard "*"

MY DGS – annotated

- Filter/Search by glosses, translation equivalents, and HamNoSys
- Improved transcript overview shows transcript ID, annotation depth, and participant pair ID
- Data collection formats now link to entries in the *Sign Language Dataset Compendium*

MY DGS - ANNIS

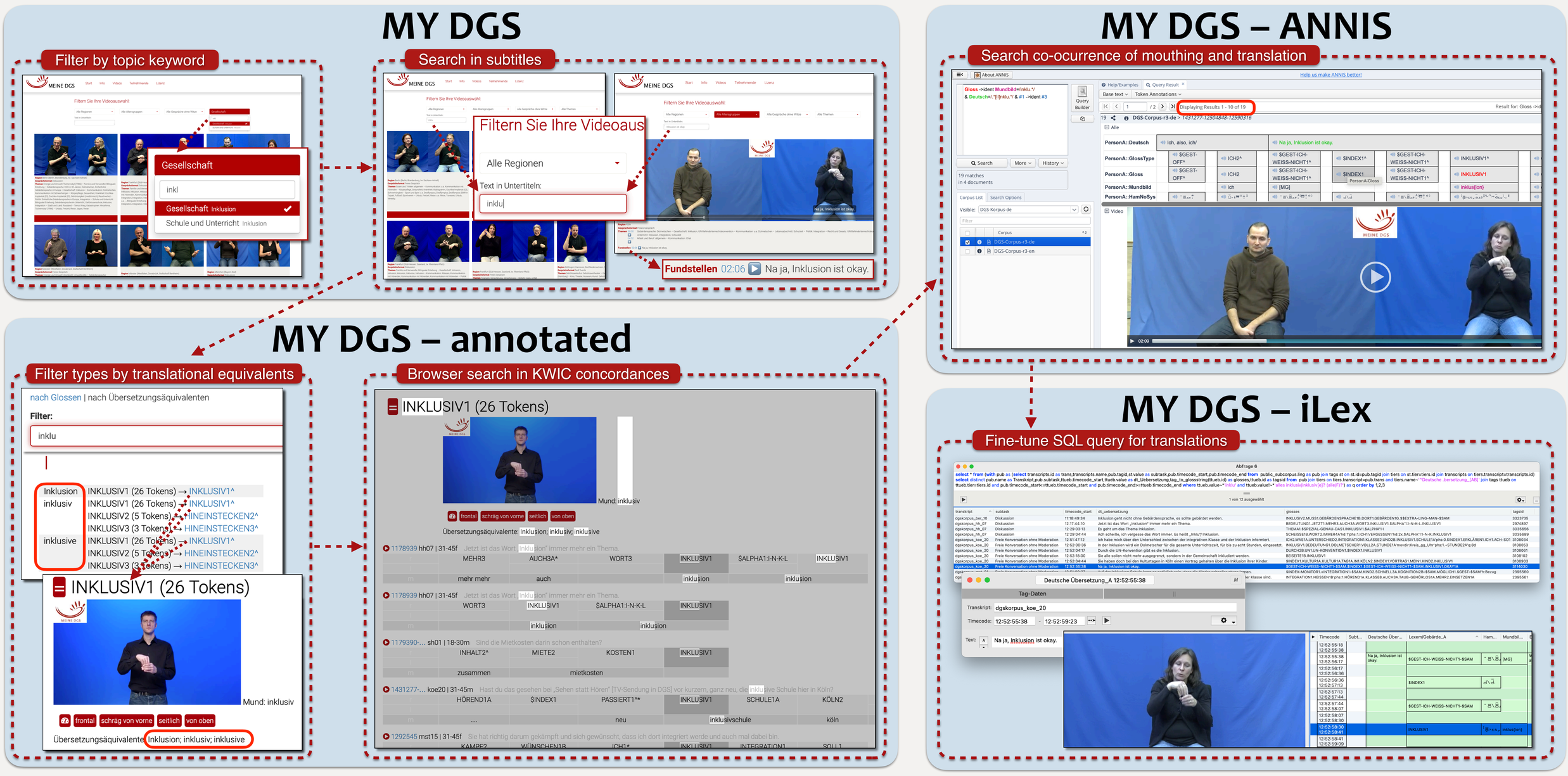
- Adds search by topics and over 570 keywords
- New Query Wizard helps users write complex ANNIS queries

MY DGS - iLex

- View MY DGS in iLex, the annotation environment used for creating the DGS Corpus

• Dolci •

Use Case Example: Finding discussions about inclusivity



Ivashechkin M, Mendez O, Bowden R. **Improving 3D Pose Estimation for Sign Language**. In Proc. of SLTAT 8. 2023 IEEE, Rhodes Island, Greece, 2023, pp. 1-5. DOI: 10.1109/ICASSPW59220.2023.10193629

Poster presented at the *11th Workshop on the Representation and Processing of Sign Languages (sign-lang@LREC-COLING 2024),* Turin, Italy. 25 May 2024.

This publication has been produced in the context of the joint research funding of the German Federal Government and Federal States in the Academies' Programme, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies' Programme is coordinated by the Union of the German Academies of Sciences and Humanities.