

Will there ever be a market for signing avatars? Some observations on the past and future of our field

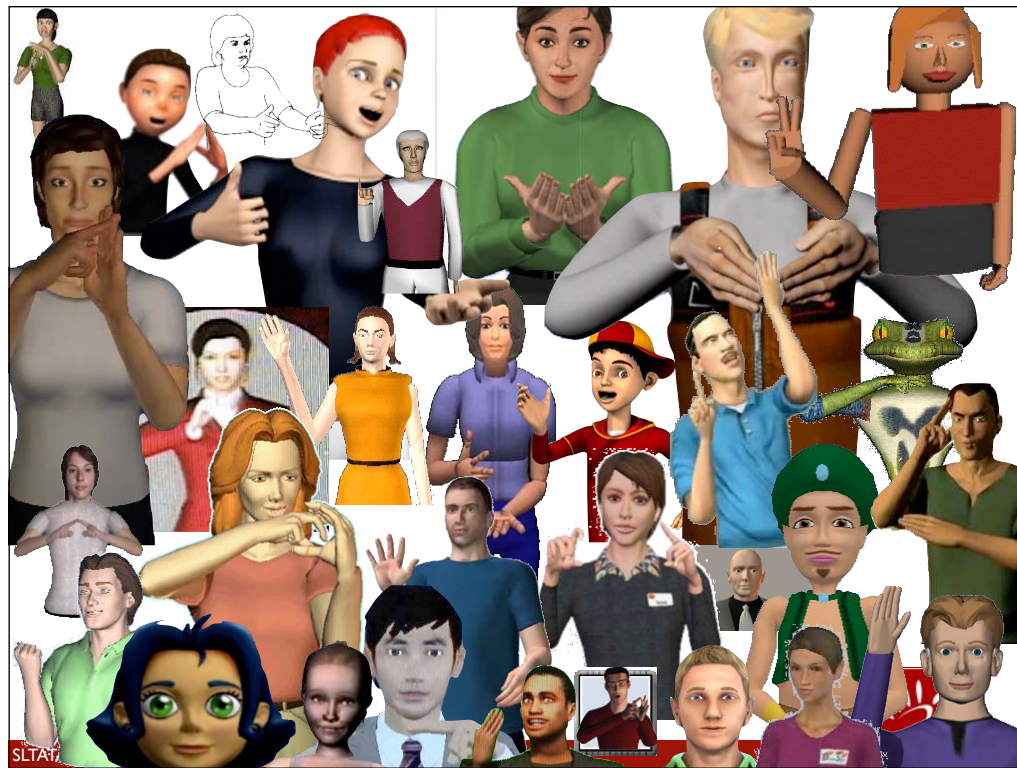
Thomas Hanke
University of Hamburg



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

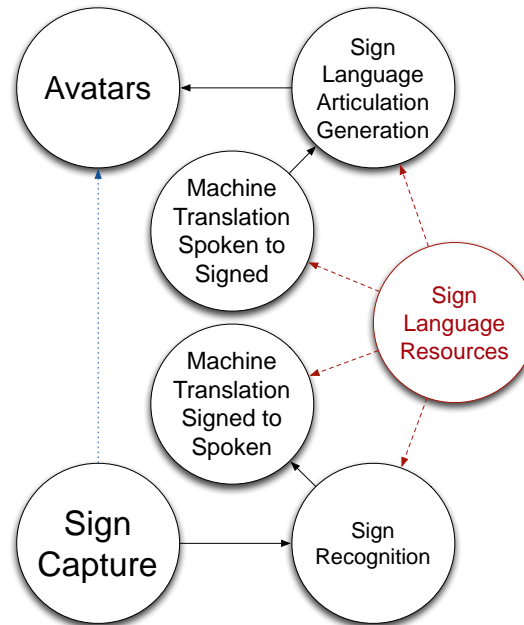


State of the union instead of a research paper. More to initiate a discussion than anything else



Babel problem

What is our field?



SLTAT Chicago 2013



So the avatar is the frontman of a whole bunch of technologies all of which are in their infancy.

The old question: Why avatars and not video?

- Economical reasons: cheaper to produce
- Ethical reasons: Anonymization possible
- Technical reasons: Glued videos look ugly



Cheaper to produce?

- Recycling what is already there, ideally a full dictionary and phrases



SLTAT Chicago 2013

In the beginning, we sold our approach to funding bodies via the cost reduction promised: No question that there is a need for signed content, e.g. on the web, but keep costs lower than with video. We are not there yet.

Anonymization

- DictaSign worked with the idea to have Web 2.0 functionality for sign languages
- Wiki



Anonymization (2)



SLTAT Chicago 2013

As an anecdote

Anonymization(2)



Anonymization (2)

DIE LINKE.



SLTAT Chicago 2013

As a side remark: All parties that did NOT have a signed version of their programme online, did not make it into the new Bundestag.

Technical reasons: Glued videos look ugly?

- Is it really that bad when gluing sentences?
- More of an issue in sign-by-sign generation.



SLTAT Chicago 2013

Compare to speech technology. Slow progress, but there is progress. For sign language video, we did not really try - in video. People did try with mocap data.

Driving forces on the market are slow

- Web technologies recommendations like Web Accessibility Guidelines
- Legislation implementing UN Conventions and precursors like ADA
- So far, we did not succeed in making signed content hip for every website owner.
- Signed content does not pay off economically.



An Example: BITV 2.0

- German barrier-free information technology act from 2011
 - Binding only for federal authorities
 - Covers:
 - Information on what a website is about
 - Information on how to navigate on that website
 - Information on what parts of the website are available in sign or easy-to-read language



SLTAT Chicago 2013

1 can be brief, or very brief. In any case, it does not make the contents of the site accessible.
2 is most boring for deaf people, taken over from needs of blind people without too much thinking.
3 can be brief if you want: Say none and you are set.

BITV Navigation

- Almost, but not exactly the same from site to site
- Obviously a field for some building blocks
- Consequently, there was a tender of the Federal Ministry of Finances to make the necessary signs available to all federal agencies.
- Does the market collapse?



Technologies & Applications

	video	avatars		
		mocap	animated	synthetic
fixed contents	✓	✓	✓	(✓)
parametrized contents	(✓)	✓	✓	✓
machine ? translation output	?	(✓)	(✓)	✓



SLTAT Chicago 2013

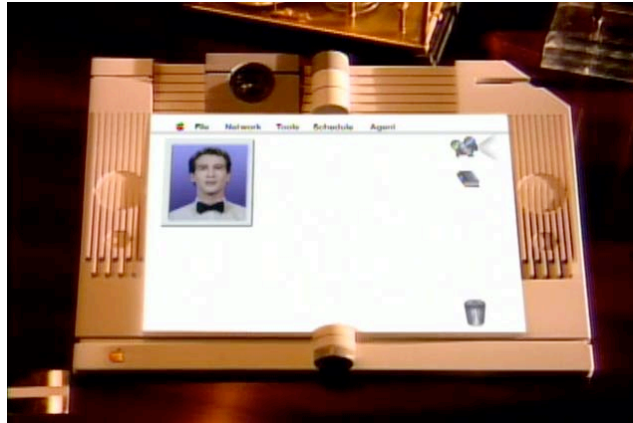
But there is not any machine translation output.

Natural Language Interfaces

- Should standard computer interfaces move away from WIMP towards NLI, sign language users would be disadvantaged once again unless NLI also means sign.

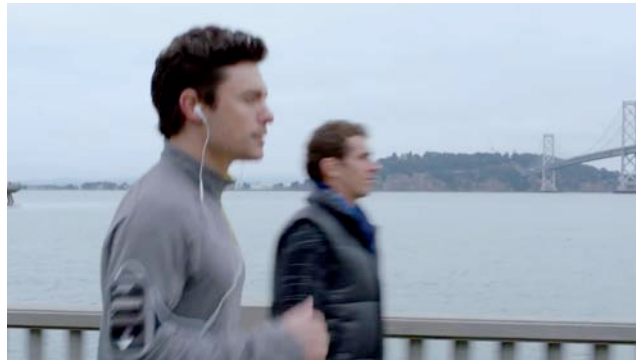


NLI Visions: Knowledge Navigator from 1987



Will NLI ever become a reality?

- At least the idea is not dead:



Remember New Economy?



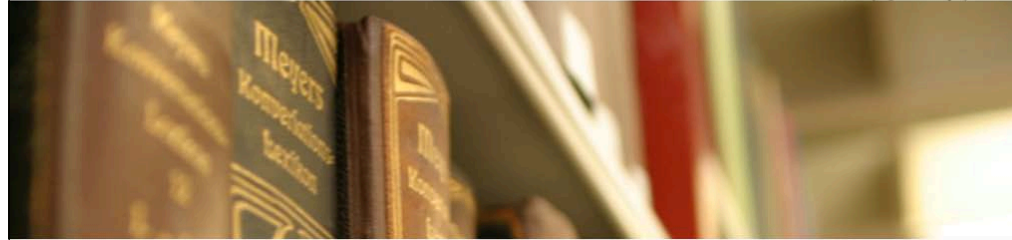
SLTAT Chicago 2013



Back then it seemed most urgent to enable avatars to sign



Guten Morgen! Ich heiße Stella. Kann ich Ihnen Tipps
für Ihre Literaturrecherche geben?



en

hem -

eluga ☐ Webseite

Service

Öffnungszeiten
Wo finde ich ...?
Ausleihe
Fernleihe
Fragen Sie uns!
Lernen und Arbeiten
in der Stabi
Bibliotheksführungen
Medienwerkstatt
Publikationsservice
Bistro

Bibliotheken

Stabi
Portrait
Sammlungen
Projekte
Presse, Ausstellungen und
Veranstaltungen
Spenden und fördern
Ausbildung und Stellenangebote
Für die Fachwelt
Fachbibliotheken
Portraits

Blog

- Stabi: 3. Oktober geschlossen
- Serviceerweiterung – Ausleihzentrum
ab 1.10.2013 länger geöffnet
- Freier Zugang zu Wissen: Open-
Access-Tage vom 1. bis 2. Oktober in
der Stabi
- Bewerbungsfrist Stellenausschreibung
Gruppenleiter/-in verlängert (16.10.)
- Veranstaltungsflyer Oktober
- Zeitgenössische spanische Kunst auf
Hamburg Art Week

Generating Human Movement

- Imitating human movement
 - often with a focus on manual articulation
- Animating human movement exaggerating important elements



Imitating Human Movement

- optical mocap equipment
- camera & depth sensor combinations such as Kinect
- high temporal resolution
- spatial resolution not sufficient to decide on \pm contact
- handshape and facial detail difficult



SLTAT Chicago 2013

While not ok for corpus data collection in a linguistic sense, certainly ok for actors to perform certain utterances.
Kinect skeleton data

Imitating Human Movement

- Frame-by-frame adjustment of a 3D model to match a video recording (“rotoscopy”)
- Interpolation between keyframes as a quality/effort trade-off
- Use multi-cam or 3D cam to disambiguate 2d views without relying on the animator’s intuition

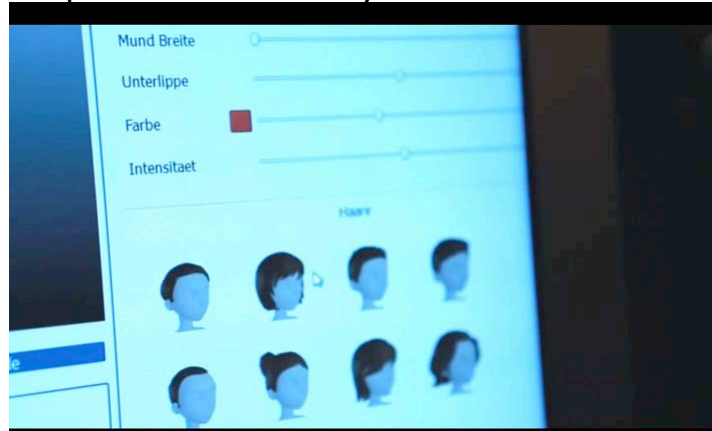


SLTAT Chicago 2013

Kinect skeleton data

Animating Human Movement

- Implement an artistic style



SLTAT Chicago 2013



Kinect skeleton data

Chunking granularity

- synthetic signing: sign level
 - plus some larger structures
- mocap & animated signing: flexible
- video: minimally “paragraphs”
- The lower we go, the less we keep of the original dynamics



SLTAT Chicago 2013

i.e. we need more research about intersign/intrasign movement differentiation
Chunking not only in the temporal domain

Machine Translation

- No large corpora available as training data (as with most languages not having a written form and many other languages as well)
- Not a sequence of symbols: More than one articulator
- Classifier constructions: Not every primitive can be found in the lexicon
- World knowledge about physical shape properties of what you are talking about



SLTAT Chicago 2013

2 articulators

Major implications on resources such as Wordnet.

sl translation

- sign-to-spoken
 - statistical
 - symbolic
- spoken-to-sign
 - symbolic
 - statistical
- sign-to-sign
 - symbolic
 - statistical

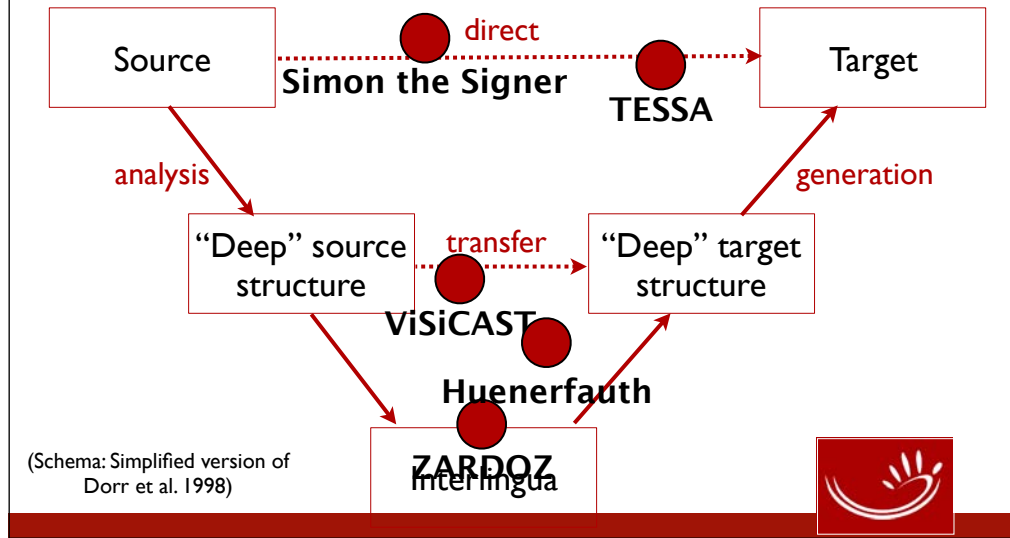


Most approaches targeting speech go thru written as an intermediate step, using standard voice recognisers or generators.

sign-to-sign cheating: gloss-to-gloss

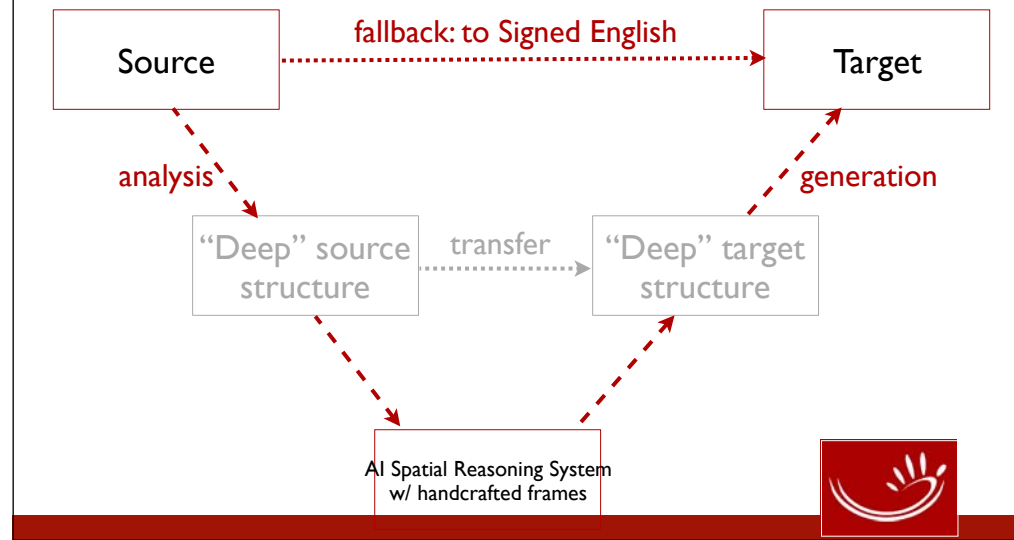
Example-based mt (EBMT) requires parallel corpora

Approaches to (Symbolic) Machine Translation



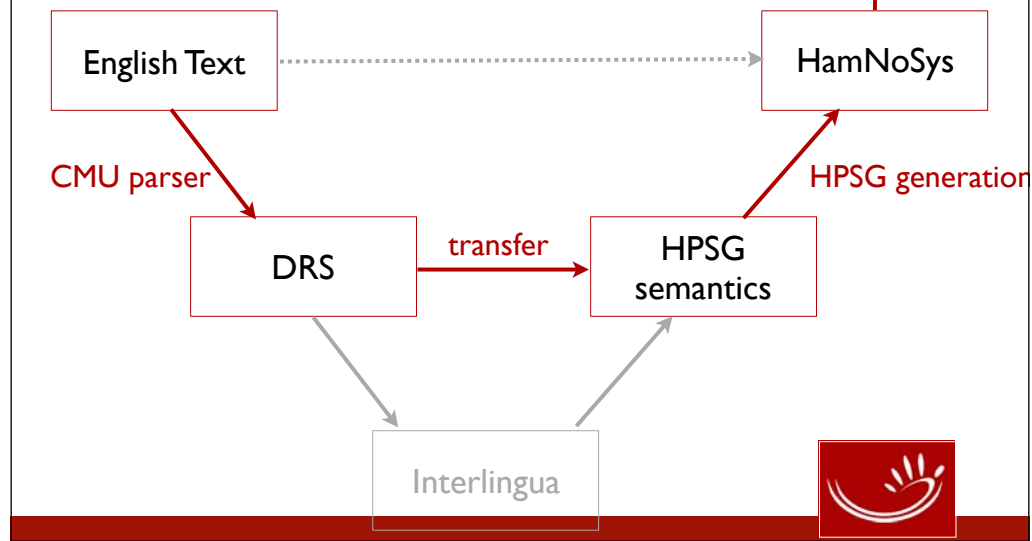
Vauquois diagram
Deep: syntax/semantic

Zardoz



Never fully implemented. Conway/Veale were ahead of their time: When the project was closed down in 1998, the first version of a FrameNet resource was published by Fillmore et al.

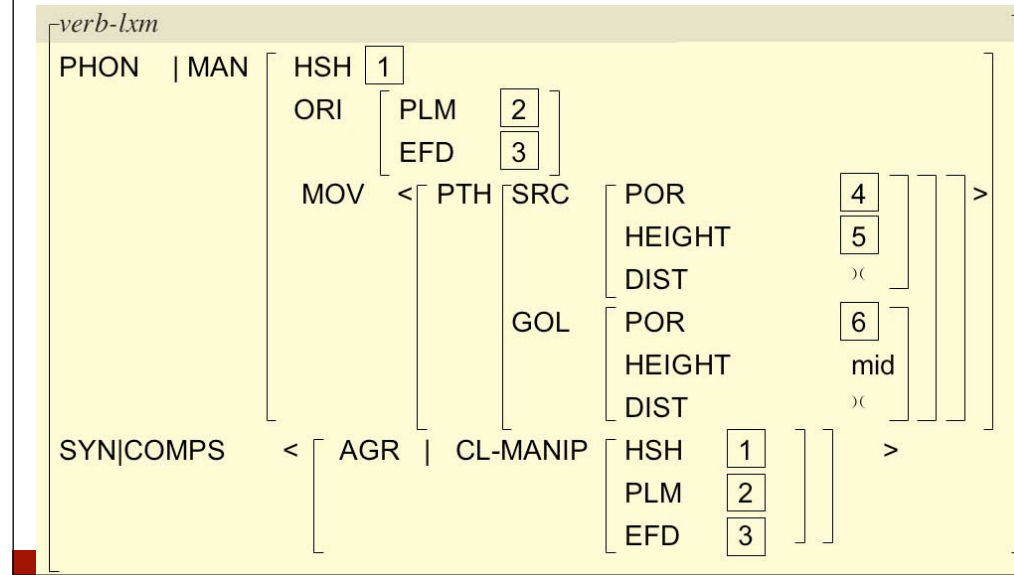
The ViSiCAST Text-to-SL System



HPSG Semantics: Minimal Recursion Semantics

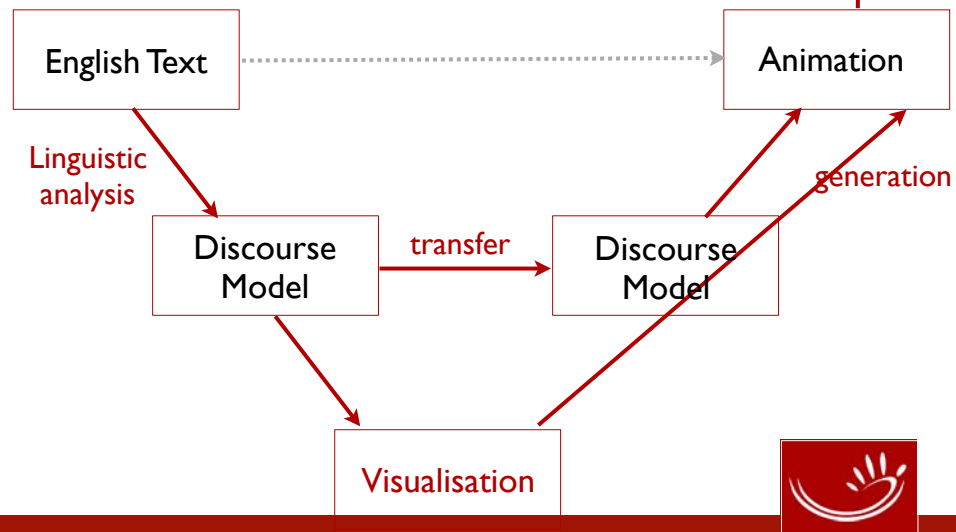
DRS: Discourse Representation Structures (Kamp/Reyle)

Example: Classifier&Directional



simply encoding the consequences of physical properties into the lexicon. Works for small domains, but leads to an explosion of types. Think about the implications for a Wordnet for sls.

Huenerfauth 2006



Huenerfauth 2006



ASL man passes between tent and frog

Machine translation

- Traditional symbolic translation and statistical approaches are still separated in our field (due to project size...)
- “hybrid approaches have become the standard in language processing” (Wahlster, July 2013)



What happened to MPEG-II & Co.?

- In 2002, there were prototype “SNHC” players that could combine avatar performance and “real” video
- Why care?
 - There is no standard way of delivery for avatar content



SLTAT Chicago 2013

Why care? Obviously you can build your own website with an integrated avatar, but: Think about the iPhone receiving an email with signed content.

Corpus linguistics too slow to fully support the field

- The idea of combining mocap data and synthetic signing has been around at least since ViSiCAST times



Language Resources supporting recognition & generation

- Beyond simple glosses: Qualified types (= type + controlled inflection vocabulary) w/ HamNoSys for each form
- Not only natural dialogue, but also competence examples that might be more appropriate for training
- No annotation standards now or in the foreseeable future: Why not define one that would support MT?



Statistical phonological rules

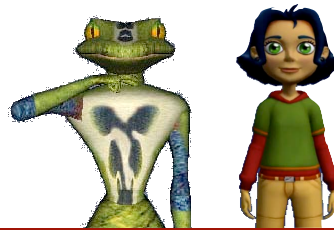
- Apply doubling to one-handed signs between two two-handed signs



Contrary to Filhol and colleagues, we remain in the paradigm of corpus linguistics.

Mission of the field

- Access to information
- Educational content in the preferred language



Mission of the field

- Access to information
- Educational content in the preferred language
- Communication across languages
- Development of sign language as a communications medium beyond face-to-face
- Integrate with future HCI
- Support sign language linguistics



SLTAT Chicago 2013

"Writing"

Lizard

How will the market develop?

- Slowly...
- Increased interest from signed content providers in avatars now that the gold rush on video is coming to an end
- Improvements needed
- More attention to how our field is observed by decision makers



SLTAT Chicago 2013

but once again compare to speech generation

And the users?

- Why is your avatar not like Pedro?
- Who in the hearing world is enthusiastic about automatic translation, speech synthesis or speech recognition as such?
- In games and educational content, this is part of the story, or an enabling technology, or... – and accepted



Cooperate!

- Think about open source, e.g. to allow PhD students to join the field
- Mix approaches
- Join efforts for a virtual larger-scale project
- No more weather forecasts!
- Develop new application areas



SLTAT Chicago 2013

At least 4 projects used this domain that does not really need translation

Thank you very much for your attention!



AKADEMIE DER
WISSENSCHAFTEN
IN HAMBURG



- The work described here is partially supported by
 - German Academies of Science programme (DGS-Korpus)
 - European Commission, 7th framework IST (Dicta-Sign)
 - and predecessors ViSiCAST & eSIGN

