

Evaluationsbericht über das Bayerische Archiv für Sprachsignale (BAS)

5.5.2000, München

Für die künftige Wissensgesellschaft ist die Sprachtechnologie eine der Schlüsseltechnologien. Die Basis jeglicher Sprachtechnologie sind Datensammlungen gesprochener Äußerungen in sehr großem Umfang. Da wir heute erst ganz am Anfang der sprachtechnologischen Entwicklung stehen, ist für die nächsten Jahre ein steigender Bedarf immer umfangreicherer und verschiedenartiger Sprachressourcen vorauszusehen.

Am BAS wurden in den letzten fünf Jahren seit der Gründung 68 GByte Sprachdaten digital erfasst, transliteriert und validiert. Dies entspricht 590 Stunden gesprochener Sprache von 2600 Sprechern. Diese Ressource stellt einen unschätzbaren Wert nicht nur für die deutsche Sprachforschung, sondern auch für die derzeit stark expandierende Sprachindustrie dar. Außerdem wird dadurch die Synergie zwischen Informationstechnik und Kulturwissenschaften nachdrücklich gefördert. Mit der Einbeziehung multimodaler Daten werden jetzt noch ehrgeizigere Ziele verfolgt.

Die Gutachter waren nicht nur beeindruckt von der extrem hohen Qualität und der Quantität der Datensammlungen, sondern auch von der hohen Effizienz und Motivation, mit der ein kleiner Stab hochkarätiger Mitarbeiter diese hervorragenden Resultate erarbeitet hat. Die interdisziplinäre Zusammensetzung der Arbeitsgruppe mit international anerkannten Wissenschaftlern mit einem Hintergrund in Elektrotechnik, Informatik und Phonetik ist in Deutschland einmalig.

Sehr positiv fällt auch die hohe Drittmittelquote auf, die sogar anwendungsnahe Forschungseinrichtungen wie die Fraunhofer-Gesellschaft übertrifft. Die wissenschaftliche Weitsicht von Prof. Tillmann, gepaart mit einem realistischen Blick für die Anforderungen der sprachtechnologischen Anwendungen, ist ein Garant des weiteren Erfolgs für BAS.

Die internationale Nachfrage nach den Sprachressourcen von BAS ist der beste Beleg für den Bedarf und die Akzeptanz der Arbeiten von BAS. Heute ist BAS das deutsche Center of Excellence für alle Fragen der Standardisierung, internationalen Kooperation auf dem Gebiet der Sprachdatensammlung.

Das BAS hat durch seine Pionierarbeit für die deutsche Sprachsignalarchivierung heute eine Know-How- und Entwicklungsvorsprung, der in Deutschland einmalig ist. Außerdem wird die Ausbildung des dringend benötigten Nachwuchses durch BAS sehr gefördert.

Auch international ist das BAS absolut konkurrenzfähig, was durch die zahlreichen eingeladenen Vorträge und Publikationen eindrucksvoll unterstrichen wird. Darüber hinaus sind die Mitarbeiter des BAS an europäischen Aktivitäten wie ELRA im Rahmen etablierter Zusammenarbeit an maßgeblicher Stelle beteiligt.

Softwarewerkzeuge wie MAUS sind nicht nur wesentlich für die effiziente Transliteration großer Sprachsignalsammlungen, sondern dienen auch darüber hinaus als eine Art

defacto–Standard. Auch die grundlegenden Beiträge zur instrumentellen Erforschung der Artikulation verbunden mit der Erhebung multimodaler Daten stellen eine wesentliche Voraussetzung für den Fortschritt in Phonetik und Linguistik allgemein und für die deutsche Sprache im Besonderen dar.

Die Gutachter befürworten nachdrücklich den Plan, auch auf dem Gebiet der Evaluierung von Anwendungssystemen tätig zu werden. Dazu wäre aber ein weiterer Personalaufbau vorzusehen.

Unbedingt anzustreben ist eine längerfristig angelegte institutionelle Förderung, da die Persistenz der Daten unabdingbar ist. Mittelfristig bietet sich für diese Daueraufgabe eine Eingliederung in die Blaue Liste an. Dies sollte aber nach Gutachtermeinung mit einer regelmäßigen Evaluation im 5–Jahresrhythmus verbunden sein.

Insgesamt unterstützen die Gutachter mit großem Nachdruck die weitere, hochprioritäre Förderung des BAS und seinen strategischen Ausbau zu einer Deutschen Signaldatenbank.

Paul Dalsgaard (Dänemark)

Hiroya Fujisaki (Japan)

Wolfgang Hess (Deutschland)

Ron Kay (USA)

Joseph Mariani (Frankreich)

Wolfgang Wahlster (Deutschland)