



Die Präsenz kleiner Sprachen in der digitalen Welt

Jan Budar



Sorbisch digital – Geschichte

- 1990 ...
 - Entwicklung Lausitz-Schriften und Tastatortreiber
 - erster Versuch einer Rechtschreibprüfung („SorbOrtho“)
 - erste sorbische Webseiten (Institutionen, Gemeinden, Vereine, sorbspirit.de)
 - sorbisches Cyberdorf internecy.de (vor Boom sozialer Netzwerke!)
- 2000 ...
 - erste Sprachressourcen (boehmak.de, dolnoserbski.de, prawopis.de, korpus)
 - sorbische Schrift wird durch Unicode-Schriften unterstützt
 - sorbische Tastatortreiber werden in Windows integriert
 - Webseiten für Unterstützung sorbischer Nutzer (radyserb.net, praskot.de)
 - erster sorbisches Computerspiel („Krabat je so nawrócił“)



Sorbisch digital – Geschichte

- 2010 ...
 - Präsenz der sorbischen Sprache in freien Plattformen (OpenStreetMap, Wikipedia, Mozilla Firefox, LibreOffice, WordPress etc.)
 - Entstehung starker sorbischer Kanäle in sozialen Netzwerken (Facebook: MDR Serbja)
 - Aufbau von soblex.de
 - erste sorbische Apps
 - erste sorbische e-Shops
 - digitale Angebote sorbischer Printmedien
 - Angebote im Zusammenhang mit der zusätzlichen Förderung der Projektes „Sorbische Sprache in den neuen elektronischen Medien“ (ab 2015)



Sorbisch im Vergleich mit anderen kleinen Sprachen

- die meisten Sprachen sind inzwischen in ISO 639 registriert
 - Code für Ober-/Niedersorbisch: hsb/dsb
 - Grundlage für die Einbeziehung von Sprachdaten in Standards (Unicode etc.)
- Auswahl von 5 Sprachen zum Vergleich
 - Walisisch (cym, 750.000 Sprecher)
 - Westfriesisch (fry, 500.000 Sprecher)
 - Rhetoromanisch (roh, 60.000 Sprecher)
 - Ladinisch (lld, 30.000 Sprecher)
 - Obersorbisch (hsb, 20.000 Sprecher)



Sprachen in Angeboten globaler Konzerne

- Facebook
 - 111 Oberflächensprachen -> cym, fry („aus dem Polnischen übersetzen?“)
- Microsoft Windows 10
 - 110 Oberflächensprachen (LIP, LEP) -> cym
- Microsoft Office 2016
 - 103 Sprachen, aber nur teilweise voll unterstützt (LAP) -> cym, fry, roh
 - hsb/dsb nur, wenn Rechtschreibprüfung installiert -> nur Wörterbuch
- Google
 - 150 Oberflächensprachen für eigene Produkte -> cym, fry, roh
 - 46 Sprachen zur Sucheingrenzung -> žana z wuběra pódla



Weitere Beispiele digitaler Präsenz

- Walisisch kann als Sprache in Geldautomaten ausgewählt werden!
 - Wieso sind Parkautomaten in Siedlungsgebiet nicht sorbisch?
- Ansprachen in Bahnhöfen und öffentlichen Verkehrsmitteln sind im Kanton Graubünden auch Rhetoromanisch!
 - Hoyerswerde ist gutes Beispiel, digitale Systeme erleichtern Mehrsprachigkeit
- Initiative „Coding for Language Communities“ führt Sprecher, Linguisten a Programmierer zusammen zur Sprachverarbeitung.
 - FOSDEM-Konferenz 01.02.2020 in Brüssel
- Die Nationalbibliothek in Wales füllt seit 2015 die walizische Wikipedia und führt Weiterbildungen für Autoren durch.



Und was ist mit gesprochenener Sprache?

- kommerzielle Sprachassistenten gehen gerade durch die Decke
 - Microsoft Cortana, Apple Siri, Amazon Alexa, Google Home
- keine Unterstützung kleiner Sprachen (große Spr. teilw. mit Dialekten)
 - Alexa: 7 Sprachen -> jendź., franc., něm., ital., jap., port., špan.
 - Google: 13 Sprachen -> zusätzl. dan., nižoz., hindi, kor., norw., šwed.
 - Apple: 21 Sprachen -> zusätzl. arab., chin., fin., hebr., malay, rusce, thai, turk.
- Prognose: bis 2025 wird sich der Markt versiebenfachen
 - Schlüssel zur Akzeptanz ist die Sprache, die zu Hause gesprochen wird
 - genügend Daten erhöhen die Chance, von Konzernen integriert zu werden
- Alternative: freie Lösungen (Mycroft, Mozilla Deep Speech)



Herausforderungen und Chancen

- kein kommerzielles Interesse von globalen Konzernen
 - Integration vielleicht durch politischen Druck möglich
- wenig Ressourcen
 - wenig Sprachdaten (Qualität umso wichtiger)
 - dies relativiert sich aber gerade durch die enormen Fortschritte bei neuronalen Netzen
 - wenig Personalressourcen (Koordination umso wichtiger)
- freie Angebote und Lösungen unterstützen
- Synergien mit anderen kleinen Sprachen aufbauen
 - kleine Sprachen sind z.B. auch Tschechisch



Fragen für die Zukunft ...

- In welche Bereiche sollten wir investieren?
- Welche Rolle spielen die bestehenden Institutionen, brauchen wir eine zusätzliche Institution?
- Wie können aktive Sprecher und Enthusiasten einbezogen werden?
- Welche strategischen Partner brauchen wir in Zukunft?
- Wie vereinen wir die wissenschaftliche Perspektive mit den Erwartungen der Nutzer?



In diesem Sinne ...

... wünsche ich eine konstruktive Konferenz!