

Patent Translation Resource (PatTR)

Fachsprachliches Parallel-Korpus, bestehend aus übersetzten Patent-Texten des European Patent Office (EPO) und der World Intellectual Property Organization (WIPO).

Sprache	Deutsch, Englisch, Französisch
Sprachliche Realisierung	schriftlich
Umfang	22 Mio. Sätze für das Sprachenpaar Deutsch-Englisch, 18 Mio. Sätze für das Sprachenpaar Französisch-Englisch, 5 Mio. Sätze für das Sprachenpaar Deutsch-Französisch
Medium	Mehrsprachige Patentbeschreibungen, satzweise aligniert
Zeitliche Einordnung	aktuell
Form der Daten	vergleichende Textsammlungen mit Metadaten, nach Sprachpaaren sortiert und zum Download verfügbar
Format	XML
Annotation	Satzgrenzen
Quelle /Herausgeber	Universität Heidelberg, Institut für Computerlinguistik
Nutzungsvoraussetzungen	Zugang frei
Link	http://www.cl.uni-heidelberg.de/statnlpgroup/pattr/
Literatur	Wäsche, K., Simianer, P., Bertoldi, N., Riezler, S., and Federico, M. (2013). "Generative and Discriminative Methods for Online Adaptation in SMT". In: Proceedings of Machine Translation Summit XIV. Nice, France. Download