

Title (en)
MACHINE LEARNING FOR TRANSLITERATION

Title (de)
MASCHINELLES LERNEN FÜR TRANSKRIPTION

Title (fr)
APPRENTISSAGE AUTOMATIQUE POUR TRANSLITTÉRATION

Publication
EP 2132657 A4 20180103 (EN)

Application
EP 08731575 A 20080306

Priority
• US 2008056087 W 20080306
• US 89337207 P 20070306

Abstract (en)
[origin: WO2008109769A1] Methods, systems, and apparatus, including computer program products, for performing transliteration between text in different scripts. In one aspect, a method includes generating a transliteration model based on statistical information derived from parallel text having first text in an input script and corresponding second text in an output script; and using the transliteration model to transliterate input characters in the input script to output characters in the output script. In another aspect, a method includes performing word level transliterations. In another aspect, a method includes using an entry-aligned dictionary of source and target script pairs, in which, whenever a particular source word is mapped to multiple target words, the dictionary includes an entry for each target word including the same source word repeated in each entry. In another aspect, a method includes using phonetic scores of words in different scripts to identify corresponding parallel text.

IPC 8 full level
G06F 17/28 (2006.01)

CPC (source: EP)
G06F 40/126 (2020.01); **G06F 40/16** (2020.01)

Citation (search report)
• [I] US 2005182616 A1 20050818 - KOTIPALLI KRISHNA V [US]
• [I] SLAVEN BILAC ET AL: "EXTRACTING TRANSLITERATION PAIRS FROM COMPARABLE CORPORA", 15 March 2005 (2005-03-15), XP055427367, Retrieved from the Internet <URL:http://anlp.jp/proceedings/annual_meeting/2005/pdf_dir/A1-1.pdf> [retrieved on 20171121]
• [I] JONG-HOON OH ET AL: "A Hybrid Model for Extracting Transliteration Equivalents from Parallel Corpora", 1 January 2006, TEXT, SPEECH AND DIALOGUE LECTURE NOTES IN COMPUTER SCIENCE;LECTURE NOTES IN ARTIFICIAL INTELLIGENCE;LNCS, SPRINGER, BERLIN, DE, PAGE(S) 119 - 126, ISBN: 978-3-540-39090-9, XP019040091
• [I] CHOI K ET AL: "A Comparison of Different Machine Transliteration Models", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 6 October 2006 (2006-10-06), XP080531384, DOI: 10.1613/JAIR.1999
• See references of WO 2008109769A1

Citation (examination)
BROWN P F ET AL: "THE MATHEMATICS OF STATISTICAL MACHINE TRANSLATION: PARAMETER ESTIMATION", COMPUTATIONAL LINGUISTICS, CAMBRIDGE, MA, US, vol. 19, no. 2, 1 June 1993 (1993-06-01), pages 263 - 311, XP008022787

Designated contracting state (EPC)
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MT NL NO PL PT RO SE SI SK TR

DOCDB simple family (publication)
WO 2008109769 A1 20080912; EP 2132657 A1 20091216; EP 2132657 A4 20180103

DOCDB simple family (application)
US 2008056087 W 20080306; EP 08731575 A 20080306