Title (en)
SYSTEMS AND METHODS FOR LANGUAGE DETECTION

Title (de)
SYSTEME UND VERFAHREN ZUR SPRACHDETEKTION

Title (fr)
SYSTÈMES ET PROCÉDÉS DE DÉTECTION DE LANGUE

Publication
EP 3519984 A1 20190807 (EN)

Application
EP 17788004 A 20171002

Priority
• US 201615283646 A 20161003
• US 2017054722 W 20171002

Abstract (en)
[origin: WO2018067440A1] Implementations of the present disclosure are directed to a method, a system, and a computer program storage device for identifying a language in a message. Non-language characters are removed from a text message to generate a sanitized text message. An alphabet and/or a script are detected in the sanitized text message by performing at least one of (i) an alphabet-based language detection test to determine a first set of scores and (ii) a script-based language detection test to determine a second set of scores. Each score in the first set of scores represents a likelihood that the sanitized text message includes the alphabet for one of a plurality of different languages. Each score in the second set of scores represents a likelihood that the sanitized text message includes the script for one of the plurality of different languages. The language in the sanitized text message is identified based on at least one of the first set of scores, the second set of scores, and a combination of the first and second sets of scores.

IPC 8 full level
G06F 17/27 (2006.01)

CPC (source: EP)
G06F 40/263 (2020.01)

Citation (search report)
See references of WO 2018067440A1

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

Designated extension state (EPC)
BA ME

DOCDB simple family (publication)
WO 2018067440 A1 20180412; AU 2017339433 A1 20190502; CA 3039085 A1 20180412; CN 110023931 A 20190716; EP 3519984 A1 20190807; JP 2019535082 A 20191205

DOCDB simple family (application)
US 2017054722 W 20171002; AU 2017339433 A 20171002; CA 3039085 A 20171002; CN 201780074219 A 20171002; EP 17788004 A 20171002; JP 2019517966 A 20171002