



Lösung

Der erste Buchstabe ist ein H, also ist das erste Zeichen des Codes ebenfalls ein H. Danach werden alle A, E, I, O, U, H, W und Y gelöscht, nun muss also noch Hlbrt übersetzt werden. Das Ersetzen der Buchstaben durch ihre Entsprechungen ergibt H4163. Es gibt keine nebeneinanderliegenden Buchstaben mit demselben Code, also muss auch nichts gelöscht werden. Nun werden die ersten vier Zeichen aufbewahrt, also ist H416 die richtige Antwort.

Dies ist Informatik!

Das Soundex-Verfahren, genauer gesagt der amerikanische Soundex, wurde bereits vor 100 Jahren von Robert C. Russel und Margaret King Odell entwickelt und patentiert. Es wurde dazu verwendet, ähnlich klingende Wörter in der englischen Sprache insbesondere auch ähnliche Namen von Personen zu finden. Das funktioniert, weil die Gruppen von Buchstaben, die demselben Code zugeordnet werden, ähnlich klingen: B, F, P und V sind Lippenlaute, C, G, J, K, Q, S, X und Z sind Gaumenlaute und Zischlaute, D und T sind Zahnlaute, L ist ein langer Fließlaut, M und N sind Nasenlaute und R ist ein kurzer Fließlaut.

Da es sehr einfach ist und nicht nur in der englischen Sprache relativ gute Resultate gibt, wird es häufig zur phonetischen Suche, also zur Suche nach ähnlich klingenden Wörtern verwendet. Es ist auch als Standard in vielen Datenbanken eingebaut.

Die Beispiele oben stammen von Donald Knuth, einem der ganz grossen Informatiker des 20. Jahrhunderts, der bis heute an seinem Buch „The Art of Computer Programming“ arbeitet. Im Band 3 „Sorting And Searching“ findet sich das beschriebene Verfahren.



Stichwörter und Webseiten

Phonetische Suche, Soundex

- <https://www.functions-online.com/soundex.html>
- <https://de.wikipedia.org/wiki/Soundex>
- <https://de.wikipedia.org/wiki/Laut>
- <https://www-cs-faculty.stanford.edu/~knuth/taocp.html>
- <http://www.highprogrammer.com/alan/numbers/soundex.html>