



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Statistik BFS

Zeichensatzproblematik

Claude Gisiger, Bundesamt für Statistik

17. Magglinger Rechtsinformatikseminar

22.5.2017 Magglingen

DAS MAGAZIN

LETZTE AUSGABEN

September 2016

Ić bin kein Schweizer

Robert Matešić muss seinen Namen falsch schreiben, wenn er sich einbürgern lassen will.



Interpellation Wermuth (SP, AG)

16.3717 Wer Schweizerin oder Schweizer wird, soll die Schreibweise seines Namens frei wählen können

Auszug aus der Begründung:

«Gemäss Medienberichten mussten insgesamt bis zu 100 000 Personen vornehmlich mit Migrationshintergrund aus den ex-jugoslawischen Staaten die Schreibweise ihres Namens bei der Einbürgerung ändern, weil die entsprechenden Informatiksysteme nicht über die notwendigen Sonderzeichen verfügen (Infostar). Dieser Umstand ist stossend. (...) Die Schweiz muss ihre migrantische Realität endlich auch auf institutioneller Ebene ernst nehmen. Menschen und Namen mit z. B. kroatischem Akut sind genauso vollwertig «schweizerisch» wie Frau Müller oder Herr Meier.»



Erfassungsgrundsätze im Infostar

- Beim Zivilstand werden die Namen so erfasst, wie sie in den Urkunden oder massgebenden Ausweisen geschrieben sind, **soweit es der Standardzeichensatz** erlaubt (Art. 24 ZStV).
- Die Daten werden nach dem hinterlegten Standardzeichensatz erfasst (ISO-Norm 8859-15). Mit dieser westeuropäischen Kodierung, lassen sich alle Sonderzeichen der vier Landessprachen der Schweiz abbilden.
- Diese Norm wird auch als Richtlinie für die Einwohnerregister der Gemeinden und Kantone im Merkmalskatalog des BFS deklariert (Registerharmonisierung).



Grundsätze im ZEMIS oder beim EDA

- In den Anwendungen des SEM sind nur lateinische Zeichen zugelassen, andere Schriftsysteme (kyrillisch, griechisch, arabisch usw.) werden transkribiert.
- ZEMIS unterstützt sämtliche lateinische Zeichen gemäss ISO 8859 Latin-1-10 (siehe folgende Folie).
- Beim EDA (Register Ordipro und VERA) wird auch der Zeichensatz ISO 8859 Latin-1-10 berücksichtigt.



Transliteration und Transkription

- Transliteration: Nicht im Standardzeichensatz enthaltene lateinische Sonderzeichen werden gestützt auf eine Tabelle erfasst. Zum Beispiel: $\hat{S} \rightarrow S$, $\check{E} \rightarrow E$, $\acute{C} \rightarrow C$ etc.
Somit werden Vornamen wie «Pětr» (serbisch) als «Petr» oder «Pēteris» (lettisch) als «Peteris» erfasst; «Péter» (ungarisch) wird aber 1:1 übernommen («é» existiert im Zeichensatz).
- Transkription: nichtlateinische Alphabeten werden aussprachebasiert transkribiert.
Beispiele betreffend Vornamen «Peter»:
Пөmp (polnisch/russisch) → Piotr oder Pjotr
بٹروس (arabisch) → Bitrus



Der Teufel liegt im Detail !

- **Die Transkription ist sprachabhängig** (d ≠ f ≠ i ...).

Boris Eltsine	[transcription française]
Boris Jeltzin	[deutsche Transkription]
Boris El'cin	[trascrizione italiana]
Boris Yeltsin	[english transcription]

- **What you see is not always what you get ...**

Je nach Norm sind die Glyphen identisch aber das Objekt verschieden!

Beispiel mit Grossbuchstabe Đ: Klein geschrieben kann das Đ ein isländisches **ð** (eth) oder ein serbisches **đ** (dj) bedeuten.

Đoković ≠ Đoković !!!

ISO-8859-15

ISO-8859-2



Fazit

- Viele Amtsstellen wissen nicht mehr genau, wie sie ausländische Namen und Vornamen in ihren Registern erfassen müssen. Die Normen und die Regeln stimmen nicht immer überein. Technisch sind sie aber vielfach im Stande, jedes beliebige Sonderzeichen zu generieren.
- Mit der Verbreitung der automatischen Datenaustausche ist diese Situation problematisch. Die Anzahl der Inkohärenzen zwischen den Registern wird immer grösser.
- Die steigenden Zahlen der Migrationen verschärfen die Problematik.

Es herrscht also Handlungsbedarf !