



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Département fédéral de l'intérieur DFI  
Office fédéral de la statistique OFS

# Problématique des jeux de caractères

Claude Gisiger, Office fédéral de la statistique

**17<sup>e</sup> séminaire d'informatique juridique de Macolin**

22.5.2017 Macolin

**DAS MAGAZIN**

LETZTE AUSGABEN

Septembre 2016

## Ić bin kein Schweizer

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



# Interpellation Wermuth (SP, AG)

**16.3717** Naturalisations. Liberté d'orthographier son nom.

Extraits du développement :

*«Selon les médias, près de 100 000 personnes, essentiellement d'origine ex-yougoslave, ont dû modifier la graphie de leur nom lors de leur naturalisation parce que les systèmes informatiques (Infostar) ne disposent pas des caractères spéciaux nécessaires. Cette situation est choquante. (...) Il est grand temps que la Suisse prenne acte sur le plan institutionnel de la réalité migratoire. Les personnes dont le nom comporte, par exemple, un caron croate sont aussi suisses que Madame Porchet ou Monsieur Chollet.»*



# Règles de saisie dans Infostar

- A l'état civil, les noms sont en principe saisis comme ils sont écrits dans les actes ou pièces d'identité présentés, **du moins tant que le jeu de caractères standard** le permet (art. 24 OEC).
- Les données sont saisies selon le jeu de caractères standard déposé (norme ISO 8859-15). Avec ce codage ouest-européen, tous les caractères spéciaux des quatre langues nationales peuvent être représentés.
- Cette norme est aussi reprise comme directive pour les registres des habitants des communes et des cantons dans le Catalogue officiel des caractères de l'OFS (harmonisation des registres).



# Règles dans SYMIC et au DFAE

- Dans les applications du SEM, seuls les caractères latins sont autorisés; les autres alphabets (cyrillique, grec, arabe, etc.) sont transcrits.
- SYMIC soutient tous les caractères latins selon la norme ISO 8859 Latin-1-10 (voir la page suivante).
- Au DFAE (registres Ordipro et VERA), le jeu de caractères ISO 8859 Latin-1-10 est également en usage.





# Translittération et transcription

- Translittération: les caractères spéciaux qui ne sont pas contenus dans le jeu standard sont saisis à l'aide d'une table de translittération.

Par exemple:  $\hat{S} \rightarrow S$ ,  $\check{E} \rightarrow E$ ,  $\acute{C} \rightarrow C$ , etc.

*Ainsi, un prénom comme «Pětr» (serbe) est saisi «Petr»; «Pēteris» (letton), «Peteris»; «Péter» (hongrois) est repris tel quel (le «é» existe dans le jeu de caractères standard).*

- Transcription: les alphabets non latins sont transcrits sur la base de leur prononciation.

*Exemples pour le prénom «Peter»:*

*Пётр (polonais/russe) → Piotr oder Pjotr*

*بتروس (arabe) → Bitrus*



# Le diable est dans les détails !

- **La transcription dépend de la langue parlée (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 ...**

Selon la norme, les glyphes sont identiques mais l'objet peut être différent !

Exemple avec la majuscule Đ: reproduit en minuscule, le Đ peut être un ð (eth) islandais ou un đ (dj) serbe.

Đoković ≠ Đoković !!!

ISO-8859-15

ISO-8859-2



## En résumé

- De nombreux services ne savent plus exactement comment saisir les noms et prénoms étrangers dans leur registre. Les normes et les règles ne sont pas toujours cohérentes. Techniquement, ils sont pourtant souvent en mesure de générer n'importe quel signe spécial.
- Avec l'extension des échanges automatiques de données, cette situation est problématique. Le nombre d'incohérences entre les registres devient toujours plus grand.
- Les chiffres croissants de la migration renforcent le problème.

**Il est donc temps d'agir pour trouver une solution !**