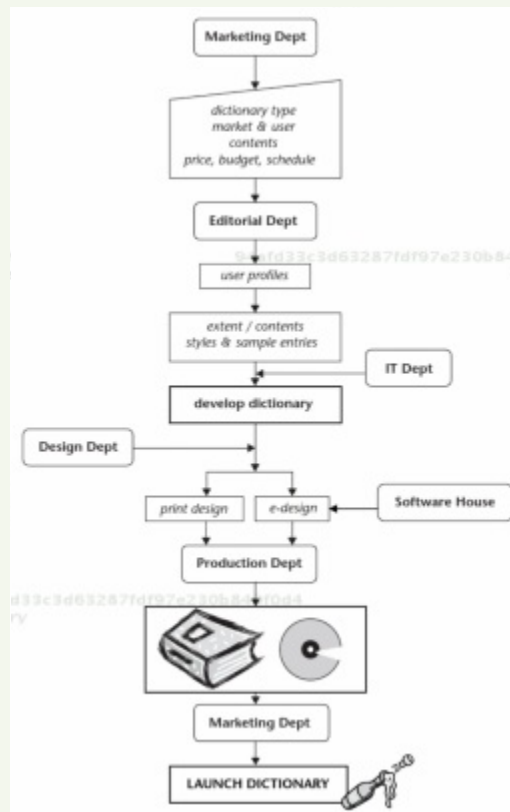


Počítačová lexikografie Tvorba slovníku

Adam Rambousek

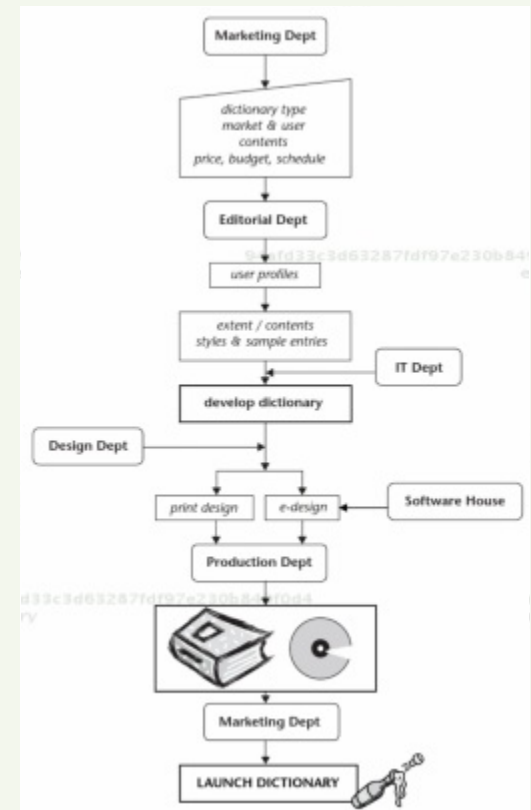
Tvorba slovníku

- tvorba slovníků je drahá, náročná a trvá dlouho, konkurence
- grant nebo se musí vyplatit
- co chybí? > druh slovníku a jeho uživatelé
- rozpočet a časový plán
- uživatelské profily, Style guide
- editační software (výroba nebo nastavení)
- korpus (vývoj, prohledávání)
- procesy



Tvorba slovníku

- píšeme slovník
- vzhled a sazba (tisk, digitální)
- výroba
- propagace
- prodej
- profit



Rozsah

- obsah + reprezentace + vzhled \Leftrightarrow délka
- obsah: počet hesel a informace o heslech
- reprezentace: zápis informací
 - n. x noun
 - skloňování
 - zkracování heslového slova
- vzhled
 - písmo (velikost)
 - nové řádky
- elektronické verze, mobily

Uživatelské profily

- pro koho slovník připravujeme?
- typ uživatele – kdo slovník používá?
 - děti, studenti, dospělí
 - rodilí mluvčí, studenti (úroveň)
 - veřejnost, odborníci
- druh použití – k čemu slovník používá?
 - příručka (neznámá slova, pravopis, křížovky...)
 - studium oblasti, jazyka
 - překlady
 - psaní úkolů, příprava na zkoušky

Uživatelské profily

- znalosti uživatele – co už umí?
 - lingvistické termíny, morfologie
 - použít slovník – značky, zkratky, odkazy
 - výslovnost, IPA
- výzkumy
 - průzkum trhu (vydavatelé) – dotazníky, agentury, online doplňky s registrací... "reagujeme na přání čtenářů"
 - akademické – často využití při vzdělávání
- They that take a dictionary into their hands, have been accustomed to expect from it a solution of almost every difficulty. (Samuel Johnson)

TABLE 15 Success ratings of dictionary look-ups according to specific types of information sought

Answer to Q5: Did they find it?	(Q.4 = "A") Seeking an L2 translation		(Q.4 = "B") Checking word as L2 translation		(Q.4 = "C") Seeking grammar of L2 word		(Q.4 = "D") Seeking collocations of L2 word		(Q.4 = "E") Seeking other information	
Yes	124	31%	127	41%	16	47%	39	38%	9	20%
Yes, but ...	92	23%	78	25%	6	18%	22	22%	6	14%
No	179	45%	102	33%	12	35%	39	38%	28	64%
Unspecified	1	—	2	1%	—	—	2	2%	1	2%
Total	396	100%	309	100%	34	100%	102	100%	44	100%

- We believe that dictionary skills must be taught, carefully and thoroughly, if dictionary users are to extract from their dictionaries the information which lexicographers have put into them. Teachers will be better able to carry out such teaching if they are fully aware of exactly what their students are doing with their dictionaries, what they expect from them, and how easily they are satisfied during the process of consultation. (Sue Atkins, Krista Varantola: Monitoring Dictionary Use)

Style Guide

- standardizace pravidla pro jevy a informace, které se ve slovníku opakují často
- instrukce pro jednotlivé části hesla
 - jak popisovat
 - jak zobrazit
 - pořadí částí hesla
 - zadávání do databáze
- pro lexikografy: jistota, konzistentnost
- pro čtenáře: navigace, důvěra

Style Guide

- pravidla (např. pravopis, zkratky, kategorie)
 - často řeší aplikace (např. výběr z možností)
- doporučení záleží na úsudku editora, kterou z nabízených možností použije
- obecné principy (pro výjimky, které je zbytečné řešit podrobně)
- pokyny pro zadávání v aplikaci
- pilotní fáze skutečná hesla napsaná podle SG, ověření a doladění

Style Guide, ukázka

➤ Order and punctuation of labels

Single labels are followed by a full stop.

Where there is more than one label in an entry, the order is:

regional, subject, register, usage, status ...

Usage labels such as derog. or humorous are placed in parentheses when they follow other labels, although not when they are used on their own.

Where you wish to use two labels from the same category, e.g. two status labels, they are joined using a roman 'and', not 'or', unless there is actual doubt.

Examples:

Physics. rare.

S.Afr. Mining.

N.Amer. Mil. slang

XML, lehký úvod

- značkovací (meta)jazyk
- pravidla, jak má vypadat správně vytvořený dokument snadné strojové zpracování a výměna informací
- konkrétní názvy značek určuje uživatel (standards, vlastní)
- elementy `<značka>obsah</značka>`
- atributy `<značka atribut="hodnota"/>`
- samozřejmě mix `<značka atribut="hodnota">obsah</značka>`
- správné zanoření značek
- popis/kontrola obsahu: DTD, XML Schema

XML, slovník

➤ heslo

- heslové slovo
- gramatika
- výslovnost
- významy
 - definice
 - příklad

```
<heslo>
```

```
< slovo ></ slovo >
```

```
< gramatika ></ gramatika >
```

```
< výslovnost ></ výslovnost >
```

```
< významy >
```

```
< význam >
```

```
< definice ></ definice >
```

```
< příklad ></ příklad >
```

```
</ význam >
```

```
</ významy >
```

```
</ heslo >
```

Píšeme slovník

- Dictionary Writing System
 - <http://tshwanedje.com/tshwanelex/>