

Počítačová lexikografie

Tvorba slovníku

Adam Rambousek

Tvorba slovníku

Tvorba slovníku, plánování

- tvorba slovníků je drak, náročná a tvrdá práce
- konkurence
- granty nebo se musí vyplatit
- co říkáš? – jaký slovník a jeho uživateli
- vzápětí a zároveň
- uživatelské profily, style guide
- editiční a technická práce, náročná a samostatná
- logika (vývoj, publikování)
- procesy

Tvorba slovníku, výroba

- pilná práce
- editiční práce (tvůrčí, digitální)
- výroba
- propagace
- prodej

Parametry

Rozsah

- oběh v reprezentaci – vzhled – délka
- oběh – počet – kvalita informací a funkcí
- reprezentace – výpočet informací
- n. a. rozsah
- aktualizace
- aktualizace – kvalita informací a funkcí
- vzhled
- přímá – funkce
- rozšíření
- aktualizace – kvalita informací a funkcí

Uživatelské profily

- proč bych slovník připravoval?
- typičtější – jaká slovník používá?
- děti, studenti, dospělí
- typičtější – jaká slovník používá?
- proč bych slovník připravoval?
- typičtější – jaká slovník používá?
- děti, studenti, dospělí
- typičtější – jaká slovník používá?
- proč bych slovník připravoval?
- typičtější – jaká slovník používá?
- děti, studenti, dospělí
- typičtější – jaká slovník používá?

Uživatelské profily

- zvolit uživatele – co si myslí?
- typičtější – jaká slovník používá?
- proč bych slovník připravoval?
- typičtější – jaká slovník používá?
- děti, studenti, dospělí
- typičtější – jaká slovník používá?
- proč bych slovník připravoval?
- typičtější – jaká slovník používá?
- děti, studenti, dospělí
- typičtější – jaká slovník používá?

Uživatelské profily

- údaje o uživateli a vyhledávání
- funkce a služby
- rozložení – funkce a služby
- rozložení – funkce a služby
- rozložení – funkce a služby
- rozložení – funkce a služby
- rozložení – funkce a služby
- rozložení – funkce a služby
- rozložení – funkce a služby
- rozložení – funkce a služby

Výstup

Style Guide

Style Guide

- standardizace – pravidla pro vývoj a implementaci, která se ve slovníku opakují často
- pravidla pro používání slov
- jak používat
- jak používat
- jak používat
- jak používat
- jak používat
- jak používat
- jak používat
- jak používat

Style Guide

- pravidla (např. pravopis, druzky, kategorie)
- data – jak a kde používat
- data – jak a kde používat
- data – jak a kde používat
- data – jak a kde používat
- data – jak a kde používat
- data – jak a kde používat
- data – jak a kde používat
- data – jak a kde používat
- data – jak a kde používat

Style Guide, ukážka

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 slang, or humorous are placed in parentheses when they follow other labels, although not when they are used on their own. Where possible to use two labels from the same category, e.g. two status labels, they are joined using a roman numeral, see 101, unless there is actual data, for example Physics: see 545. 1000, 1000: see 1000



Počítačová lexikografie

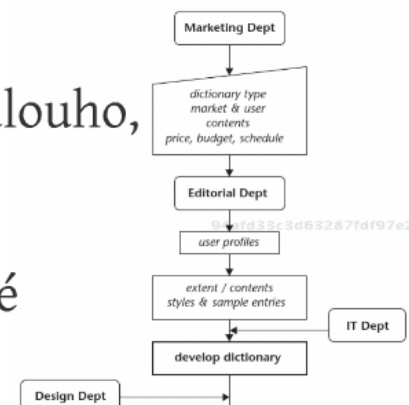
Tvorba slovníku

Adam Rambousek

Tvorba slovníku

Tvorba slovníku, plánování

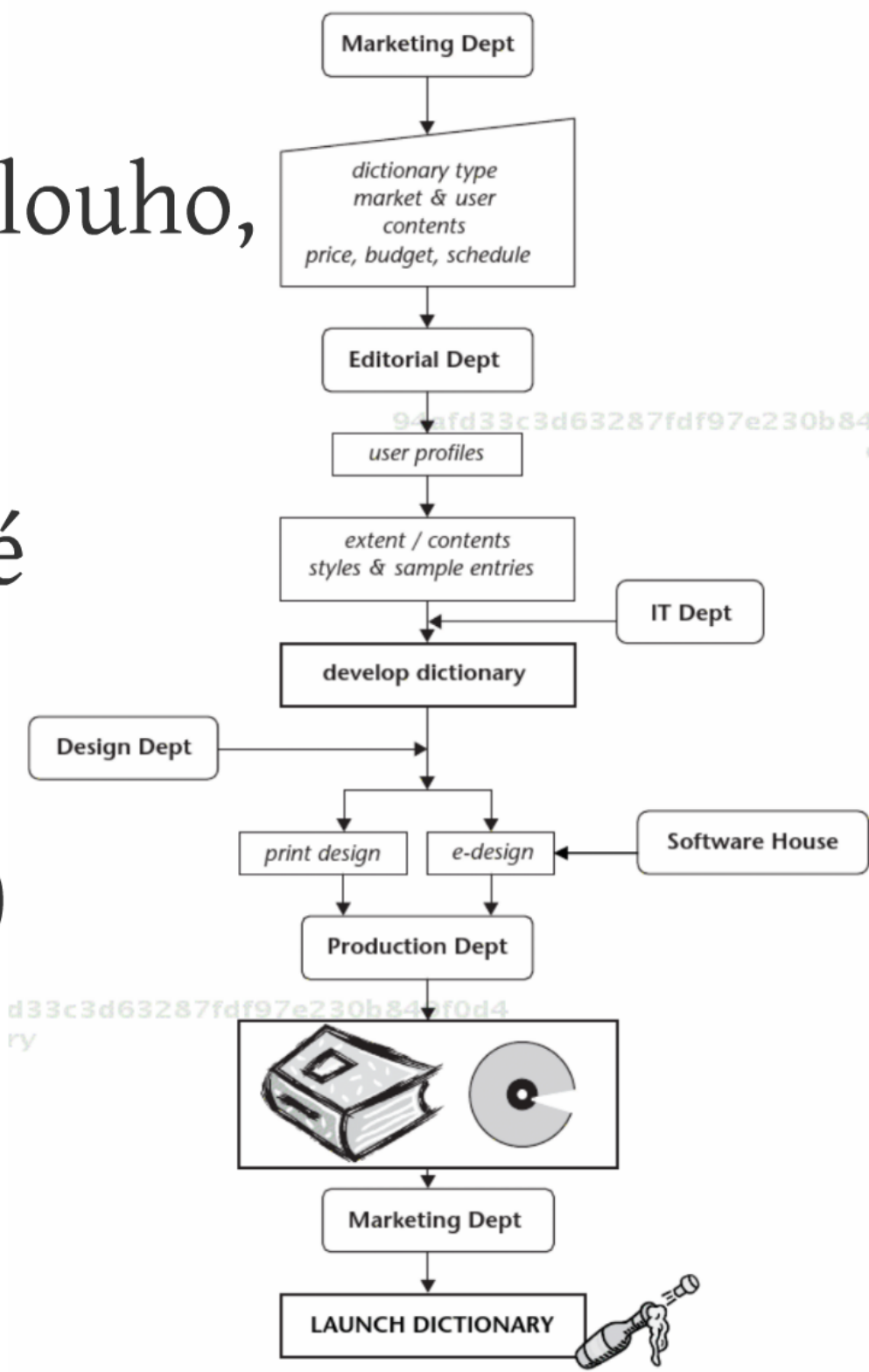
- tvorba slovníků je drahá, náročná a trvá dlouho, konkurence
- grant nebo se musí vyplatit
- co chybí? -> druh slovníku a jeho uživatelé
- rozpočet a časový plán



čná a trvá dlouho,

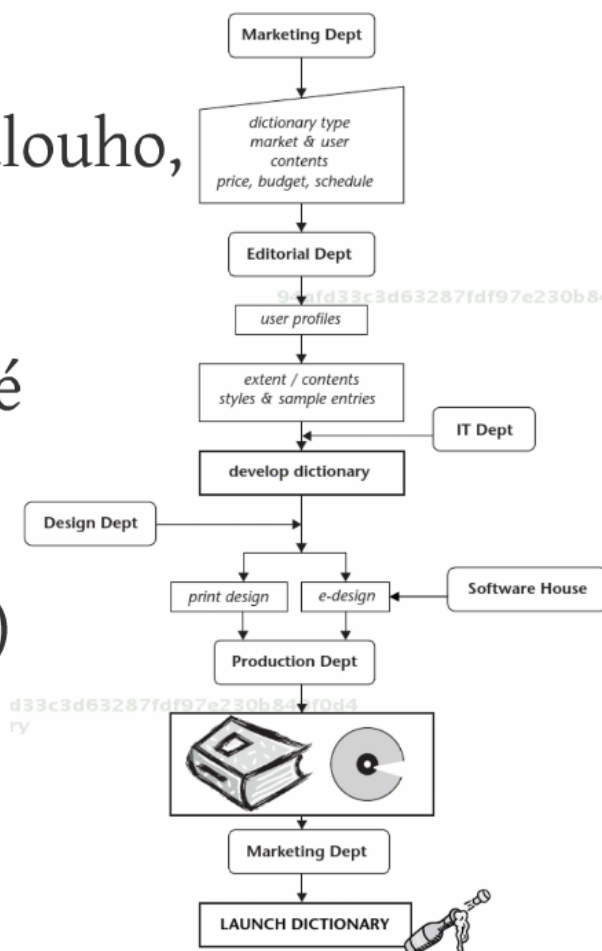
ho uživatelé

e
o nastavení)



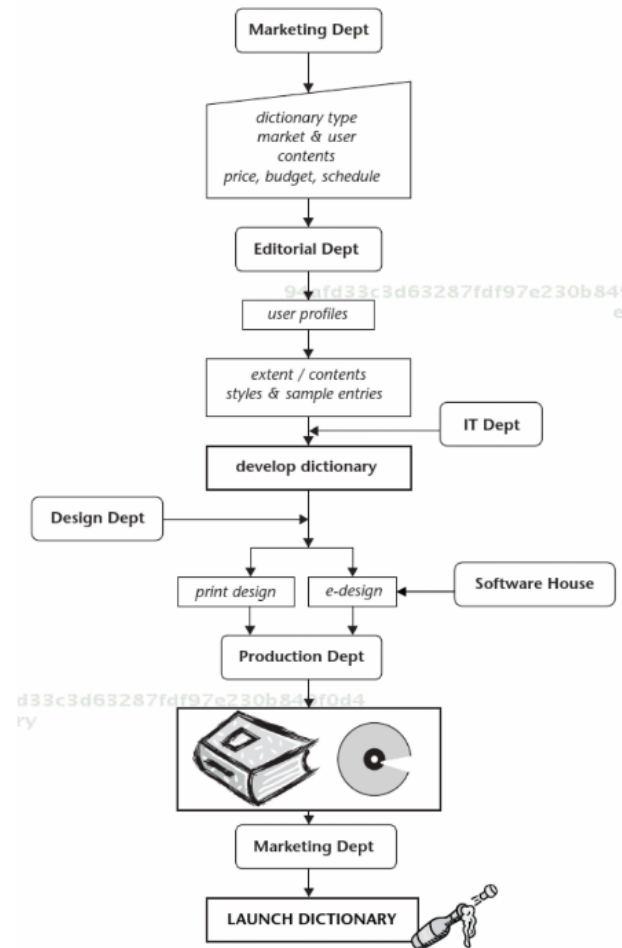
Tvorba slovníku, plánování

- 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, výroba

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





Parametry



Rozsah

- obsah + reprezentace + vzhled \Leftrightarrow
- *obsah*: počet hesel a informace o he



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)

Uživatelské profily

- údaje z webu a vyhledávání
 - *Dictionary.com*
 - následné hledání: bastion, hiatus, enmity, decorous
- zpětná vazba
 - *Merriam-Webster's Advanced Learner's English Dictionary*

Výzkum

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*)

m

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

Výzkum

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



Style Guide

standardize provide proprietary info

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*