

Redukované texty

aneb jak na notace, abstrakt, extrakt, referát,...

Martin Krčál

Redukovaný text - definice TDKIV

Text, který vznikne na základě **obsahové analýzy dokumentu z plného textu** dokumentu **sémantickou redukcí informací** obsažených v dokumentu, tj. výběrem z hlediska obsahu podstatných informací. Na **sémantickou redukcí** zpravidla navazuje **komprimace vybraných informací** a jejich **vyjádření v přirozeném nebo umělém jazyce** formou vět, klíčových slov nebo znaků. Redukovaný text dokumentu je vždy podstatně **kratší** než plný text a naplňuje funkci signální (upozornění na nový dokument), substituční (náhrada plného textu dokumentu) nebo selekční (umožnění výběru dokumentů). Hlavní typy redukovaného textu dokumentu jsou referát, anotace, extrakt, resumé a synopse; někdy se mezi ně řadí i soubory věcných selekčních údajů (např. deskriptorů, předmětových hesel, klasifikačních znaků apod.), které však vyjadřují spíše tematiku než vlastní obsah dokumentu.

Redukované texty

- obsahová analýza
- z plného textu
- sémantické redukce obsahu (informací)
- komprimace vybraných info
- přirozený nebo umělý jazyk
 - ❖ věty, klíčová slova nebo znaky
- kratší
- vzniká **nové dílo** ⇒ vychází z originálu

Funkce RT

■ signální

- ❖ existence textu

■ substituční

- ❖ nahrazení čtení celého textu

■ selekční

- ❖ pro vyhledávání literatury
- ❖ RT v katalogu ⇒ ano nebo ne???

■ ostatní

- ❖ tématické shluky (podobné dokumenty)

Historie RT

- redukce = běžná součást života
 - ❖ pravěké malby, záznamy bitev, beletrie
- od 16.stol. – zprávy o dílech
 - ❖ mezi učenici (humanismus, šíření info)
- 19. stol. – rozvoj vědeckých časopisů
 - ❖ indexové, abstraktové, referátové časopisy, oborová specializace
- 2/2 20. stol. – rozvoj PC a ICT
 - ❖ katalogizace, automatická indexace a abstrahování, abstraktové DB, online rozhraní, propojení s FT, RSS, alerts,...

Dvě cesty RT

- odborná literatura
- beletrie
- **Rozdíly???**



Typologie RT

Extrakt = výťah (extract)

- věty z dokumentu
- nevznikají nové formulace
- cíl: **vystihnout obsah dokumentu**
- zásada neutrálnosti
 - ❖ nemá se manipulovat s obsahem

Resumé (resume, summary)

- shrnutí, souhrn
- rekapitulace přečteného textu
 - ❖ nemá funkci signální a substituční
 - ❖ na konci dokumentu
 - ❖ shrnutí nejdůležitějších poznatů
 - ❖ pro ty, co už text četli (doplnění orientace, správné pochopení)
- druhy
 - ❖ autorské
 - ❖ redakční

Abstrakt (abstract)

- zaměřuje se jen na text
 - ❖ nehodnotí, chybí doplňkové informace
- rozsáhlejší než anotace
- části abstraktu
 - ❖ cíle
 - ❖ použitá metodologie (v rozsahu nezbytně nutném pro pochopení)
 - ❖ výsledky a závěry
 - ❖ vedlejší informace – to co nemá souvislosti s cílem dokumentu, ale je důležité

Umístění abstraktu

- co nejbližší začátku dokumentu
- časopis
 - ❖ pod nadpis a autora
- samostatná publikovaná zpráva
 - ❖ titulní list, dokumentační list, první lichá strana před obsahem
- monografie, VŠKP
 - ❖ rub titulního listu nebo následující lichá stránka
- samostatné abstrakty kapitol
 - ❖ na první stránku kapitoly nebo před ní

Abstrakt (abstract)

■ druhy (dle počtu dokumentů)

- ❖ monografický - 1 dokument
- ❖ přehledový - dokumenty na vybrané téma

■ druhy (dle obsáhlosti)

- ❖ informativní - i substituční funkce, podrobnější, technické a přírodní obory
- ❖ indikativní - signální funkce, kratší, společenské a humanitní vědy, častější

■ druhy (dle autora)

- ❖ autorský - píše sám autor
- ❖ kritický - odborník na téma, text by neměl být hodnocen (např. knihovník)

Abstrakt (abstract)

■ druhy (dle uspořádání témat)

- ❖ analytický - dle věcných hledisek, nemusí odpovídat uspořádání původního dokumentu
- ❖ výběrový - dle určitého hlediska (např. skupina čtenářů, požadavky instituce apod.)
- ❖ strukturovaný - ne text, ale formou bodů (metody, cíle, východiska,...) a k nim jsou navázána fakta, přehledné

Abstrakt (abstract)

- modulární abstrakt

- ❖ 4 složky – anotace, indikativní, informativní a kritický abstrakt

Formální a stylistická úprava

■ bibliografické údaje

- ❖ primární dokumenty - stejná stránka
- ❖ sekundární dokumenty + samostatné abstrakty - těsně před nebo za
- ❖ ČSN ISO 690, 690-2

■ délka

- ❖ průměr 250 slov, krátké text 100 slov, max. 500 slov (max. 1 tisková strana)
- ❖ určuje ji obsah než délka textu

■ 1. věta = ústřední myšlenka, u sekundárních druh dokumentu

Formální a stylistická úprava

- členění do odstavců jen u delších abstraktů (ne nadpisy a kapitoly)
- souvislé věty, krátké věty, logicky navazující
- klíčová slova za abstraktem (volitelně)
- slovesa v činném rodě
 - ❖ trpný rod pouze tam, kde má být zdůrazněn účastník děje
- 3. osoba, jiná osoba jen pro snazší pochopení textu

Formální a stylistická úprava

- obecná terminologie
 - ❖ definice v abstraktu ne!!!
- standardizované měrné jednotky, symboly, zkratky
 - ❖ ISO, národní normy
- nepoužívat rovnice, tabulky, vzorce, diagramy,...
- specifikace v normě
 - ❖ **ČSN ISO 214** Dokumentace - Abstrakty pro publikace a dokumentaci (únor 2001)

Anotace (annotation)

- kratší a stručnější než abstrakt
- hodnocení, info o tématu, autorovi, ...
 - ❖ i z jiných dokumentů

Druhy anotace

■ dle autora

- ❖ autorská, nakladatelská, redakční

■ dle hodnocení

- ❖ doporučující (z hlediska skupiny čtenářů)
- ❖ textologická (kvalita textu)

■ další anotace:

- ❖ analytická – strukturovaná RT dle věcných hledisek
- ❖ skupinová – více dok. na společné téma
- ❖ bibliografická – anotace v bibliografii
- ❖ informativní - nehodnotí

Další druhy RT

■ Synopse

- ❖ redukovaný text rozsáhlejších děl
- ❖ nebo RT mezi názvem a textem

■ Regest

- ❖ výtah diplomatického textu

Kdo tvoří nejlépe RT?

- autor
- odborník na danou problematiku
- knihovník/informační specialista



Požadavky na tvůrce RT

- zkušenosti
- přehled
- znalost oboru
- jazykové schopnosti

Tvorba RT

Co ovlivňuje délku RT

- délka originálního textu
- komplexnost řešeného problému
- pokrytí řešených oborů
- důležitost obsahu pro čtenáře
- přístupnost – nedostupný = delší
- účel RT

Proces tvorby

■ čtení textu

- ❖ porozumění obsahu, linking words

■ analýza

- ❖ selekce (nedůležité věci)
- ❖ interpretace (vyjádření obsahu vlastními slovy, subjektivní, myšlenkové mapy)

■ syntéza

- ❖ napsání RT

Automatická RT

- rozdělení textu na věty
- ohodnocení
- složení

Přístupy k ART

■ statistický přístup

- ❖ frekvence slov, stop slova, stemming (varianty slova, slovník), problémem synonyma

■ syntaktický přístup

- ❖ struktura textu, zákonitosti (umístění slov, významovost vět, podřazené věty, vazby, parsery pro rozložení vět)

■ sémantický přístup

- ❖ důležité části spojeny s určitými pojmy (podstatný, záměr, cíl, důležitý), kombinace

Evaluace

- zejména pro automaticky tvořené
- neexistuje univerzální návod na automatickou tvorbu RT
- možnosti srovnání
 - ❖ originální text
 - ❖ gold summary
- posouzení celku nebo jednotlivých částí

Druhy evaluace

- vnější

- ❖ účel textu, proč byl vytvořen

- vnitřní

- ❖ orientace na text a obsah

Problémy evaluace

- nikdy se nevytvoří stejný RT
- nikdy se nevytvoří stejné hodnocení

Zdroje

- *Dokumentace – Abstrakty pro publikace a dokumentaci : ČSN ISO 214*. Praha : Český normalizační institut, 2001. 15 s.
- *Referát a anotace : ČSN 01 0194*. Praha : Vydavatelství Úřadu pro normalizaci a měření, 1983. 7 s.
- MACHÁČKOVÁ, Martina. *Redukované texty v katalozích ČR* [online]. Brno, 2010. 95 s. Vedoucí diplomové práce Mgr. Josef Schwarc. Dostupné z WWW: http://is.muni.cz/th/179890/ff_m/Machackova_diplomova_prace.pdf.

Závěr



Děkuji Vám za pozornost

Martin Krčál
krca1@fss.muni.cz