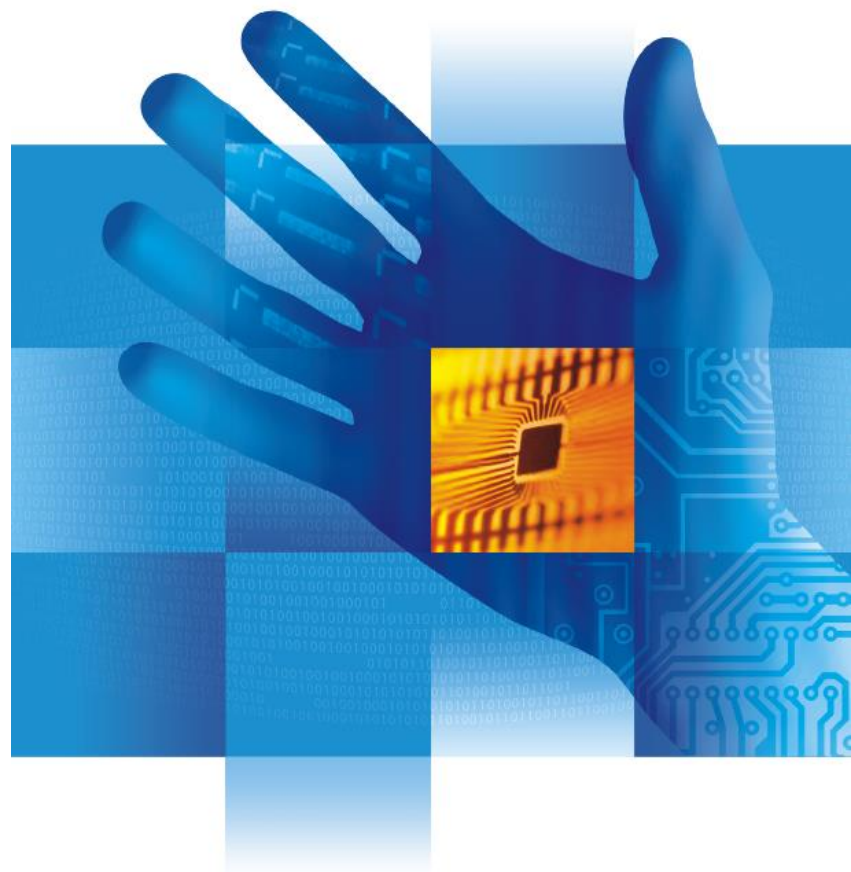


Doporučovací systémy založené na obsahu

Petr Šaloun

VŠB-Technická univerzita Ostrava

FEI, katedra informatiky



Reprezentace položek

- Položky, které mohou být doporučeny uživateli, jsou často uloženy v databázových tabulkách.

ID	Name	Cuisine	Service	Cost
10001	Mike's Pizza	Italian	Counter	Low
10002	Chris's Cafe	French	Table	Medium
10003	Jacques Bistro	French	Table	High



Reprezentace položek

- Tabulka zobrazuje jednoduchou databázi se záznamy (tj. řádky), které popisují tři restaurace.
- Název sloupce označují vlastnosti (též atributy, charakteristiky, pole nebo proměnné) restaurace.
- Každý záznam obsahuje hodnotu pro každý atribut.
- Jednoznačný identifikátor (atribut ID) umožňuje jednoznačně určit jednotlivé položky, i když mají stejné jméno.



Uživatelské profily

- Profil zájmů uživatele je používán většinou doporučovacích systémů.
- Profil se může skládat z mnoha různých typů informací.

Hlavní dva druhy informací jsou:

- model předvoleb uživatele
- historie interakcí uživatele s doporučovacím systémem



Model předvoleb uživatel

- popis typů položek, které zajímají uživatele.
- společná reprezentace je funkce, která pro jakoukoliv položku předpoví pravděpodobnost, že uživatele bude tato položka zajímat.
- Z důvodu efektivity může být tato funkce použita pro získání n položek, které jsou nejpravděpodobněji v oblasti zájmu uživatele.

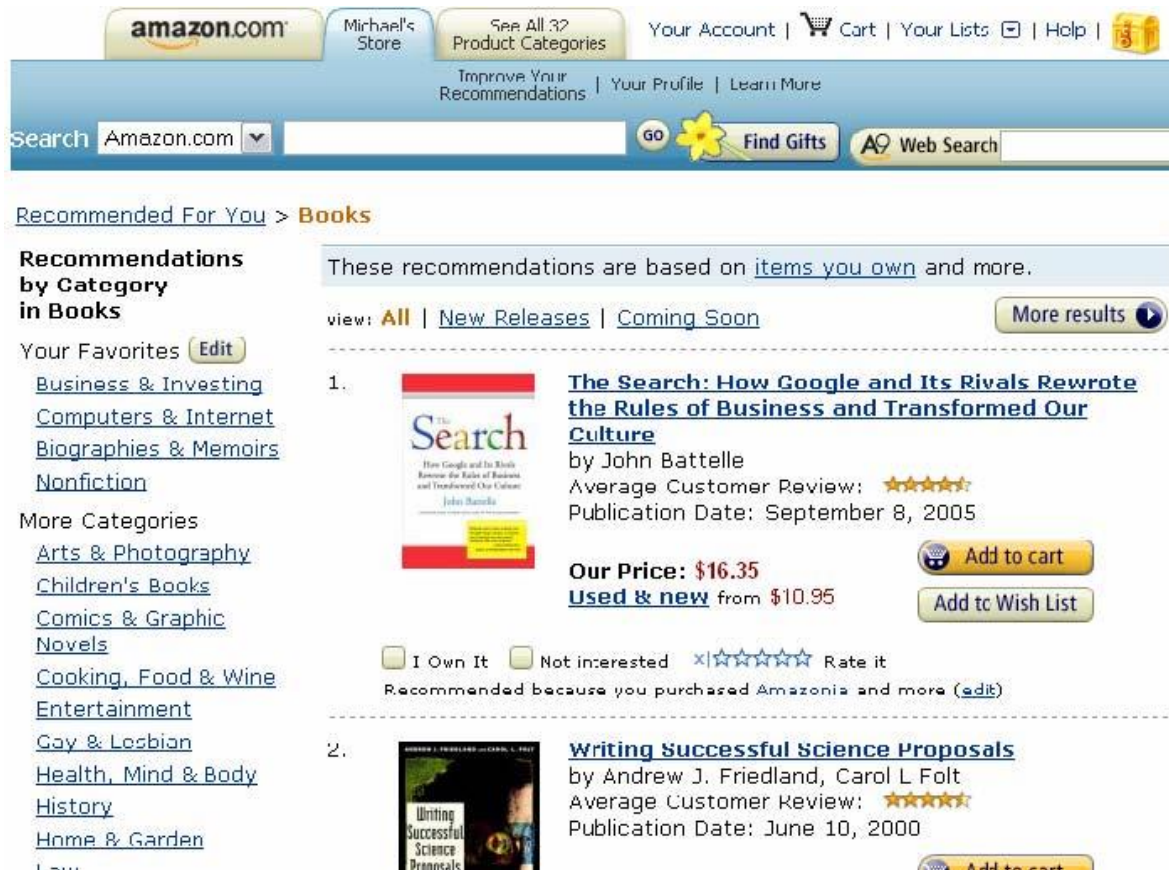


Historie interakcí uživatele s doporučovacím systémem

- Zahrnuje uložení položek, které uživatel viděl, spolu s informacemi o dalších interakcích uživatele.
- Jiný typ historie zahrnuje ukládání dotazů napsaných uživatelem (např.: „uživatel hledal italské restaurace v oblasti s PSČ 90210“).



Příklad 1- doporučování knih na serveru Amazon.com



The screenshot shows the Amazon.com interface. At the top, there's a navigation bar with the Amazon logo, a user profile for 'Michael's Store', and links for 'Your Account', 'Cart', 'Your Lists', and 'Help'. Below this is a search bar with 'Amazon.com' entered, and buttons for 'GO', 'Find Gifts', and 'Web Search'. The main content area is titled 'Recommended For You > Books'. It features a sidebar on the left with 'Recommendations by Category in Books' and 'Your Favorites' (with an 'Edit' button). The main area displays two book recommendations. The first is 'The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture' by John Battelle, with a 5-star average customer review and a price of \$16.35 (used & new from \$10.95). The second is 'Writing Successful Science Proposals' by Andrew J. Friedland and Carol L. Folt, with a 5-star average customer review and a price of \$16.35. Both items have 'Add to cart' and 'Add to Wish List' buttons. A 'More results' button is also present.

amazon.com Michael's Store See All 32 Product Categories Your Account | Cart | Your Lists | Help |

Improve Your Recommendations | Your Profile | Learn More

Search Amazon.com GO Find Gifts Web Search

Recommended For You > Books

Recommendations by Category in Books

Your Favorites [Edit](#)

- [Business & Investing](#)
- [Computers & Internet](#)
- [Biographies & Memoirs](#)
- [Nonfiction](#)

More Categories

- [Arts & Photography](#)
- [Children's Books](#)
- [Comics & Graphic](#)
- [Novels](#)
- [Cooking, Food & Wine](#)
- [Entertainment](#)
- [Gay & Lesbian](#)
- [Health, Mind & Body](#)
- [History](#)
- [Home & Garden](#)

These recommendations are based on [items you own](#) and more.


view: **All** | [New Releases](#) | [Coming Soon](#) [More results](#)

- 

The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture
by John Battelle
Average Customer Review: ★★★★★
Publication Date: September 8, 2005

Our Price: \$16.35
Used & new from \$10.95

[Add to cart](#)
[Add to Wish List](#)

I Own It Not interested x [★ ★ ★ ★ ★](#) Rate it
Recommended because you purchased Amazonia and more ([edit](#))
- 

Writing Successful Science Proposals
by Andrew J. Friedland, Carol L. Folt
Average Customer Review: ★★★★★
Publication Date: June 10, 2000

[Add to cart](#)



Příklad 1 - vysvětlení

- Amazon obsahuje vlastnost nazvanou oblíbené (favorites), která představuje kategorie položek, které uživatel preferuje. Oblíbené položky jsou buď vypočítány sledováním položek, které si uživatel koupil, nebo mohou být vybrány ručně uživatelem.



Příklad 2- Uživatelské přizpůsobení na serveru Amazon.com



Edit Favorites

Mark the categories that interest you the most.

Books

Submit

Your Books Favorites

Categories

- | | |
|---|--|
| <input checked="" type="checkbox"/> Biographies & Memoirs | <input checked="" type="checkbox"/> Nonfiction |
| <input checked="" type="checkbox"/> Business & Investing | |
| <input checked="" type="checkbox"/> Computers & Internet | |

Add to Your Favorites

- | | |
|--|---|
| <input type="checkbox"/> Arts & Photography | <input type="checkbox"/> Outdoors & Nature |
| <input type="checkbox"/> Children's Books | <input type="checkbox"/> Parenting & Families |
| <input type="checkbox"/> Comics & Graphic Novels | <input type="checkbox"/> Professional & Technical |
| <input type="checkbox"/> Cooking, Food & Wine | <input type="checkbox"/> Reference |
| <input type="checkbox"/> Entertainment | <input type="checkbox"/> Religion & Spirituality |

Příklad 2 - vysvětlení

- Příklad uživatelského rozhraní pro přizpůsobení, ve kterém může uživatel vybrat kategorie.



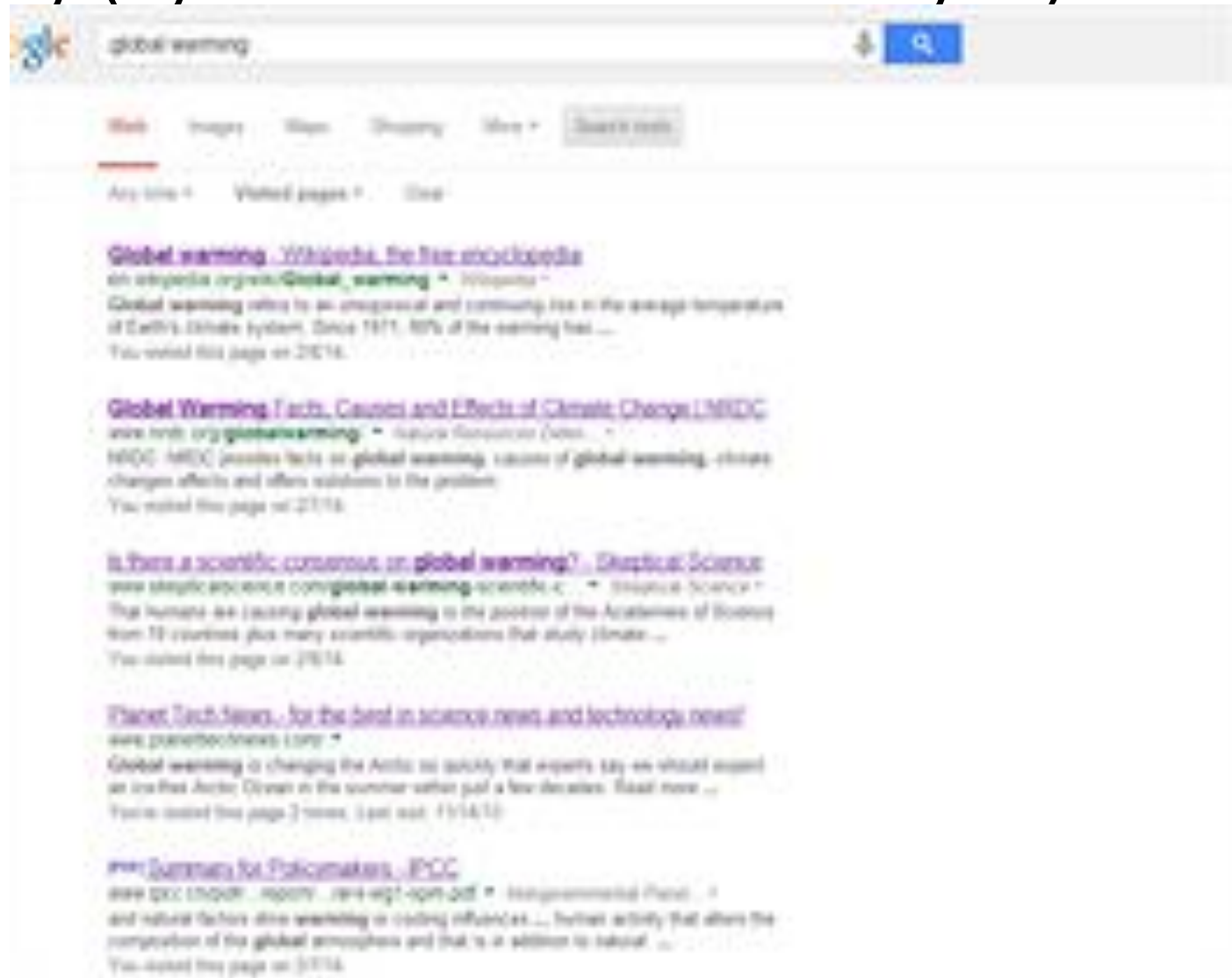
Google vyhledávání (tipy a triky)

- Vyhledávač chápe význam slov a snaží se pochopit uživatelův záměr



Google vyhledávání (tipy a triky)

- Filtry (vyhledávání z navštívených)



Google vyhledávání (tipy a triky)

- www.google.com/ncr

homepage nepřesměruje na stránku podle současné polohy.

NCR = no country redirect



Google vyhledávání (tipy a triky)

- Operátory:
- **OR** - stačí splnění jedné z podmínek
- **Uvozovky** - konkrétní fráze
- **Hvězdička (*)** zastupuje libovolné slovo (slova), např. *"in * invented * phone"*
- **Mínus (-)** odmítnutí
- **AROUND** dva výrazy musejí být blízko (kolik slov)
- .. Interval, např. **2008..2010**.
- ~ **vlnovkovka** = včetně synonym
např. ~tutorial najde i guide a manual.

Příklad: první **AROUND(5)** "počítačová myš"
vynález OR prototyp -krysa



Google vyhledávání (tipy a triky)

Vymeznění

- **Site:** doména
- **Intitle:** v titulku stránky
- **Inurl:** adresa URL.
- **Filetype:** typ souboru.



Google vyhledávání (tipy a triky)

- Kalkulačka
prostě $5 + 6 * 7$
- Převody jednotek
5 km in miles
- Obrázky
stačí přetáhnout



Závěr

- Doporučovací systémy založené na obsahu doporučují uživateli položky na základě popisu položky a profilu uživatelských zájmů.
- I když uživatelský profil může být vytvořen vložením dat uživatelem, většinou je vytvořen učením se z uživatelské zpětné vazby.
- Pro učení uživatelských profilů bylo adaptováno mnoho různých učících algoritmů a výběr učícího algoritmu závisí na reprezentaci obsahu.

