

ANOPRESS, spol. s r.o.

Jan Kozel, Anopress Praha

Informační agentura Anopress shromažďuje a zpracovává plné texty novinových a časopiseckých článků, doslovné přepisy televizních a rozhlasových relací a vybrané knižní a jiné soubory encyklopedického a referenčního charakteru. Ve spolupráci s jednotlivými vydavateli vytváří agentura unikátní databanku plných textů českých tištěných i elektronických médií, doplněnou o další informační zdroje. Současná databáze Anopressu je tvořena více než 160 zdroji, z nichž 152 je denně doplňováno a aktualizováno. Dnes jsou v databázi zařazeny mj. plné texty všech celostátních a regionálních deníků. Objem dat dosahuje řádu desítek GB.

TECHNOLOGIE

Pro práci s daty je používána vysoce výkonná vyhledávací technologie společnosti VERITY, kterou využívají například zpravodajské služby celé Evropy a USA, armády NATO, banky a velké společnosti, ale i tisková agentura Reuter. V současné době představuje tato technologie nejvýkonnější prohledávací systém. K vyhledávání využívá tzv. tématické stromy – TOPIC, kterými je možno popsat jakýkoli jev, činnost, organizační strukturu, odlišit jednotlivá témata, jména apod.

Při zpracování většiny zdrojů používá Anopress vlastní extrakční program, který umožňuje vybírat textové informace z PostScriptových souborů (tj. grafického formátu pro řídicí systémy tiskáren), které získává od vydavatelů. Zdroje, které jsou k dispozici pouze v tištěné podobě, se zpracovávají na technologické OCR lince. Ta umožňuje zpracovávat i tiskové formáty až do velikosti A3.

VYHLEDÁVÁNÍ

Přes jednotlicí prostředí prohlížeče WWW služby sítě Internet může uživatel aktivně vyhledávat, analyzovat a ukládat vybraná data. Kromě jednoduchého a přehledného výstupu vyhledaných informací může navíc využívat unikátních funkcí – shlukování a sumarizace. Tyto funkce umožňují vytvářet anotace vybraných zpráv a nacházet dopředu neznáme souvislosti mezi vyhledanými informacemi. Systém vytváří automaticky hlavičku každé zprávy, která obsahuje název, zdroj, datum, ročník, autora, rubriku, stránku a identifikační číslo. Systém rovněž umožňuje tuto hlavičku doplnit o další údaje. Díky těmto a dalším funkcím je toto fulltextové vyhledávání nejdokonalejším systémem v České republice. Podle dostupných pramenů neexistuje podobný systém ani nikde jinde v Evropě.

Anopress rovněž vyvinul vlastní aplikační software pro zpracování nových titulů, zejména encyklopedií, do elektronické podoby v podobě CD ROM. Tento software splňuje všechny požadavky jako plnohodnotná multimediální aplikace s důrazem na vyhledávání a seskupování informací – viz např. CD ROM Kronika lidstva.

MONITORING

Pro pravidelné sledování a vyhodnocování jednotlivých témat, oblastí, titulů, autorů apod. je možné využívat tzv. profilace zpráv pomocí tématických stromů – TOPICů. Po každém zařazení nové zprávy do databáze, systém ihned informaci, která odpovídá předem zadanému dotazu, vybere a odešle příslušnému uživateli. Ten může dále informace vyhodnocovat, připojovat k nim vlastní údaje, komentář, nebo je exportovat do různých formátů, či vytvářet z nich přímo stránky WWW. Jedinečnou funkcí je okamžitý grafický výstup hodnocených zpráv podle zdrojů, oblastí, témat a autorů.

Společnost Anopress spolupracuje s Národní knihovnou v Praze na propojení bibliografických záznamů uložených a zpracovávaných v Národní knihovně s plnými texty článků v databance Anopressu. Rovněž společně s Národní knihovnou připravuje Anopress sdružení uživatelů, které by umožnilo jednotlivým knihovnám přístup do celé databáze Anopressu za velmi výhodných finančních podmínek.

Veškeré údaje o databance Anopressu, jednotlivých produktech, službách a zdrojích jsou přístupné na Internetu na adrese www.Anopress.cz. Jakékoli další informace lze získat i v sídle firmy.

Kontakt:

Anopress, spol. s r.o.

Gorazdova 24, 120 00 Praha 2

Tel.: 02 / 249 07 561, 569, 568

Fax: 02/ 249 13 817

e-mail: marketing@anopress.cz