

Detekcia duplicit

- Webová služba – REST-like JSON alebo XML.
- Paralelné vykonávanie s vysokým výkonom.
- Vlastný konvertor z PDF formátu špecializovaný pre slovenčinu a češtinu, vlastný jazykový analyzátor a korektor.
- Vysoko konfigurovateľný a flexibilný systém.
- Možnosť porovnávať s aktuálnymi svetovými encyklopédiami napr. Wikipediou alebo inými zdrojmi textu.
- Porovnávanie na úrovni paragrafov, poradí si aj so zmenami slovosledu, morfológických tvarov a prekrývanie odstavcov.

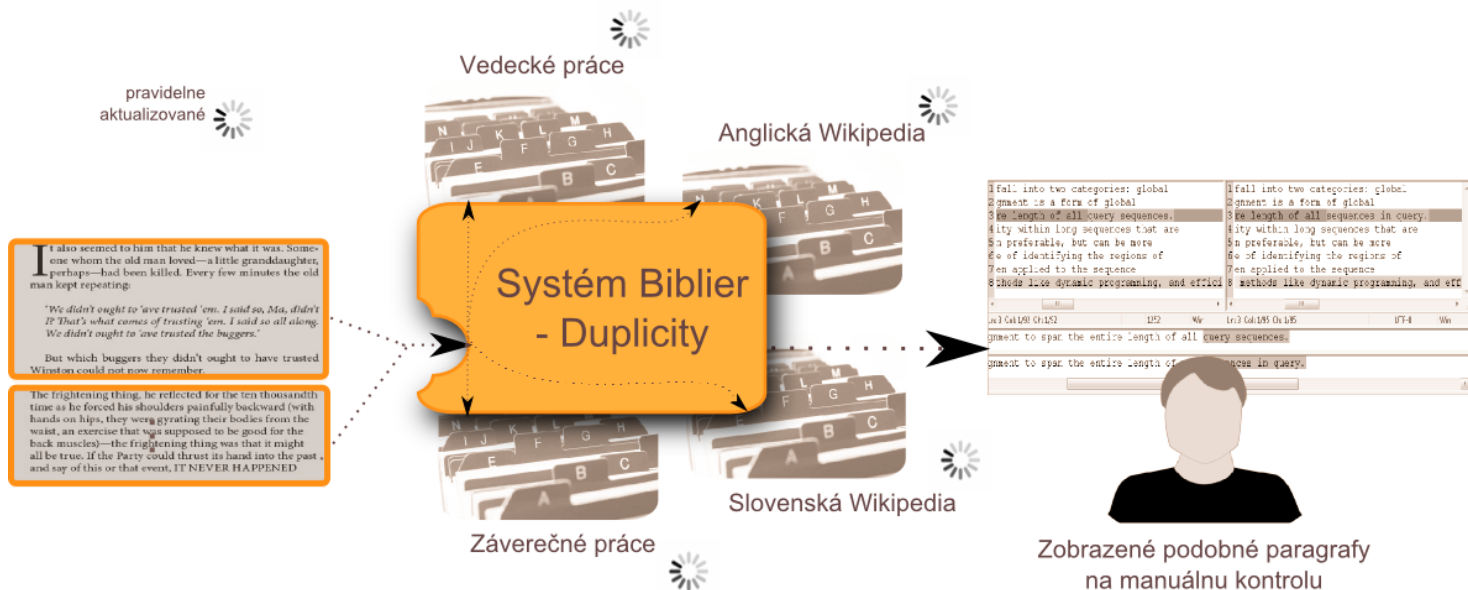
Detekcia duplicit

Mazoea s.r.o.

Hľadanie duplicit na úrovni paragrafov aj so zmeneným slovosledom a s ohľadom na morfológiu jazyka vo Vami definovanej sade dokumentov. Vďaka definovanej podobnosti je možné nájsť aj časti odstavcov, ktoré sú si podobné.

Rozšírenie Systému Biblier

Toto rozšírenie systému vyhľadávania dokáže porovnať prácu s rôznymi sadami dokumentov. Ak je podozrenie, že napr. Wikipedia je veľký zdroj necitovaného textu, môže byť jednoducho pridaná do kolekcie dát. Vo výsledku zistíte kde a ako sa overovaný dokument zhoduje s kolekciou dát. Vďaka paralelnému prístupu je jednoduché pridávať, aktualizovať a meniť kolekcie dát.



Mazoea s.r.o.

<http://www.mazoea.eu>
office@mazoea.eu