

KA 4.7 Inovace nástroje na odhalování plagiátů v IS

Pracoviště: FI – Centrum výpočetní techniky

Nová generace nástroje na odhalování plagiátů

Problematika vizuální kontroly textů je poměrně složitá a potřeba ověřování originality závěrečných a seminárních prací na Masarykově univerzitě trvá. Díky získané zpětné vazbě jsme se mohli zaměřit na ulehčení této činnosti, což vedlo k vývoji nové generace uživatelského rozhraní v aplikaci pro kontrolu podobností. Cílem úprav bylo zvýšit uživatelskou přívětivost, přispět k zefektivnění procesu kontroly a sjednotit dosud oddělené aplikace pro práci s podobnostmi.

Úpravy vycházely z grafického návrhu uživatelského prototypu, v některých případech bylo navrhované řešení mírně upraveno, s ohledem na technické možnosti nebo na základě průběžného testování.

Mezi hlavní realizované inovace patří sjednocení aplikací do jednoho prostředí. Došlo k propojení aplikací „Podobnosti“ a „Porovnej dva“, což umožňuje plynulý přechod mezi nalezenými podobnostmi a detailním porovnáním dvou dokumentů v rámci jedné obrazovky.

Byl maximalizován prostor pro kontrolu, nové rozhraní efektivně využívá šířku monitoru, omezuje rušivé prvky a zajišťuje responsivitu. Přibyla možnost vizualizace podobností pomocí lineární nebo 2D mapy. Uživatel si může nastavit, které znázornění mu více vyhovuje. Celkové procento shody bylo potlačeno, důraz je kladen na analýzu jednotlivých podobností. Uživatel má možnost přepínat citlivost, přičemž snížená citlivost vyhledá méně přeformulací. Segmenty textu jsou barevně odlišeny podle míry shody – od žluté (výrazné přeformulování) po červenou (téměř doslovná shoda).

Došlo také k zjednodušení práce s podobnými dokumenty. Uživatel může snadno vyřazovat nebo vracet nalezené podobné dokumenty a pracovat s aktualizovanými výsledky. Nové grafické prvky umožňují zaměřit se pouze na relevantní podobnosti, a to s možností interaktivního přepočítání výsledků po úpravě výběru dokumentů.

Výsledné rozhraní je navrženo tak, aby podporovalo efektivní spolupráci mezi studentem a vedoucím práce při hodnocení originality textu, a reflektovalo aktuální výzvy v oblasti detekce plagiátů.

Obr. 1: Nalezené podobnosti, vč. 2D grafu.

The screenshot shows the MUNI Similar Documents interface. On the left, a document titled "Weblog jako elektronický informač..." is displayed, with a similarity score of 26%. The main area shows three similar documents with their respective 2D similarity graphs:

- Document 1:** "weblog-jako-elektronicky-informacni-zdroj-vyzkum-informac..." from muni.cz, similarity 12%.
- Document 2:** "1374187-Siri-nakazu-skobrtnuti-o-flow-rocenka-inflow-2009.h..." from docplayer.cz, similarity 8.9%.
- Document 3:** "skobrtnuti-o-flow-rocenka-inflow-2009.pdf" from knihovna.cz, similarity 8.6%.

The 2D graphs show the distribution of similarity scores across the documents, with red lines indicating the overall similarity trend.

Obr. 2: Porovnání dvou dokumentů bez zobrazení 2D grafu podobnosti.

The screenshot shows the MUNI Similar Documents interface, but with the 2D similarity graph disabled. The main area displays the same three similar documents as in the previous screenshot, but without the 2D graphs:

- Document 1:** "weblog-jako-elektronicky-informacni-zdroj-v..." from muni.cz, similarity 12%.
- Document 2:** "1374187-Siri-nakazu-skobrtnuti-o-flow-rocenka-inflow-2009.h..." from docplayer.cz, similarity 8.9%.
- Document 3:** "skobrtnuti-o-flow-rocenka-inflow-2009.pdf" from knihovna.cz, similarity 8.6%.

The interface also shows a "Zavřít toto porovnání" button and a "Mapy podobnosti" toggle switch.