

Odkazy

 Generování popisů obrázků - problémy, poznámky, prompty

1. Schválené úkoly

- ☐ Filip upravit prompty na základě diskuse a otestovat je na existujících příkladech
- ☐ Filip poslat upravené prompty skupině nebo Martinovi před další schůzkou
- ☐ Připravit větší datovou sadu s novějšími dokumenty pro testování
- ☒ Změnit klíčová slova na množné číslo ve všech kategoriích
- ☒ Omezit počet klíčových slov z 10-15 na nižší počet a sledovat kvalitu
- ☒ Zakázat AI psát popisy v archaickém stylu podle období obrázku
- ☒ Omezit určování výtvarného stylu pouze na případy, kdy si je AI jistá
- ☒ Zakázat zmínku jednopanelového formátu u komiksů bez více panelů
- ☒ Přidat přepínač pro zobrazení všech obrázků v knize do rozhraní
- ☒ Omezit klíčová slova u grafů pouze na obsah, ne na jednotlivé hodnoty
- ☐ Zakázat používání slova infografika u starších dokumentů
- ☒ Omezit popis rukopisných poznámek bez hodnocení čitelnosti
- ☐ Dokončit anotaci face detectoru a vyřešit zbývající neověřená pole

2. Témata vyžadující další testování

- Komiksy potřebují testování na novějších dokumentech s více panely
- Mapy vyžadují větší vzorek různorodých map pro vyhodnocení chování
- Grafy potřebují testování na novějších dokumentech pro ověření kvality
- Unifikovat měřítka map podle katalogizačních pravidel a vyhledávat podle nich

3. Diskutované problémy

- Klíčová slova obsahují příliš obecné a nepřesné pojmy
- Popisy v češtině jsou někdy méně kvalitní než anglické verze
- AI píše popisy příliš strojově bez přirozeného jazyka
- AI někdy špatně detekuje mapy, zaměřuje se na části místo celku
- Rukopisné poznámky jsou často špatně rozpoznány a nesprávně hodnoceny
- Iniciály v textu nejsou propojeny s ALTA polem v OCR
- Celkově funguje systém dobře, ale potřebuje testování na více modelech

4. Otevřené otázky a návrhy

- Zvážit použití jiné AI pro kontrolu popisků místo manuálního procházení
- Nabídnout kontrolu popisků jako praxi pro studenty

- Zjistit, zda existuje slovník pro unifikované hodnoty měřítek map - <https://invenio.nusl.cz/record/188281>
- Kontaktovat kartografy v MZK ohledně standardizace měřítek