

Při ručním překládání krátkých textů, jaké se vyskytují například na sociálních sítích, či mikroblozích (Twitter a podobně), je překladatel často nucen dohledávat doplňující informace v různých zdrojích. Může se jednat o méně běžná slova, o specifické termíny z neznámé domény, či o různé zkratky.

V této práci se zabýváme návrhem a implementací systému, který pro danou krátkou textovou zprávu automaticky sestaví minimální kontextově závislý slovník. Systém v překládaném textu vybírá vhodná hesla do slovníku a vyhledává k nim definice, překlady a příklady v otevřených zdrojích, či je automaticky extrahuje z paralelního korpusu. Získaný slovníček v ideálním případě bude pro překladatele již dostačujícím podkladem, aby překládanou zprávu s jistotou pochopil a zvolil odpovídající překladové ekvivalenty včetně odborných termínů. Empirické vyhodnocení se opírá o statistiky sledující, jak často byli uživatelé s navrženými hesly spokojeni, jak často byla hesla chybná a do jaké míry systém správně určil relevanci pro daný vstupní text.