Článek *lemonUby - velký, propojený, syntakticky bohatý lexikální zdroj pro ontologie*[1] [1], představuje nový zdroj propojených dat lemonUby. LemonUby má dvě hlavní využití, napomáhá k lepšímu lexikálnímu popisu ontologických entit a také slouží jako zdroj pro aplikace pro zpracování přirozeného jazyka.

Úvod mluví o kontextu vzniku lemonUby, za důležité činitele jsou zde označeny sémantický web, Linguistic Linked Open Data Cloud jako integrační platforma pro nové data a rostoucí zájem o nasazení principů otevřených propojených dat v oblasti zpracování přirozeného jazyka a lingvistiky obecně. LemonUby je tedy rozšířením LLOD Cloudu ve kterém vyplnil jakousi mezeru. V LLOD Cloudu, totiž do té doby scházely velkoobjemové zdroje bohaté na lexikální informace o slovesech, jejich syntaktickém chování a o sémantických rolích argumentů sloves.

Z názvu lemonUby lze vyčíst, že jedná o technologii založenou na slovníkovém modelu *lemon*[2] [2] a sítě propojených lexikálně – sémantických zdrojů *UBY* [3]. V principu je lemonUby výsledek převedení zdrojů UBY do modelu lemon. Autoři podotýkají, že spojení těchto dvou technologií znamená spojení dvou odlišně se vyvíjecích oblastí. Zatímco UBY bylo vyvíjeno nezávisle na sémantickém webu, lemon byl vytvořen právě za účelem integrace lexikálních dat na sémantickém webu.

Následující kapitoly článku se zabývají právě těmito dvěma technologiemi způsobu, jak jsou konvertovány do lemonUby. Tato konverze jo možná díky formalizaci obou prvků. UBY zdroje mohou spolupracovat, protože vycházejí ze slovníkového modelu UBY-LMF[3] a jsou roztříděny do kategorií s množiny ISOcat[4]. V článku se však zdůrazňuje, že UBY-LMF nijak nesouvisí s ontologiemi a LMF není otevřeným standardem, ani jedno z těchto tvrzení však neplatí pro lemon. Postup jak konvertovat UBY-LMF data lemonUby spočívá v namapování UBY-LMF elementů na koncepty a vlastnosti sémantického webu. Nezanedbatelnou překážkou v tomto ohledu se stalo modelování datových modelů. Ačkoliv je obsah ISOcat dostupný jako propojená data nelze provést přímé propojení lemonUby na další zdroje LLOD Cloudu, protože ISOcat není formální ontologií, jedná se „pouze“ o semistrukturovanou kolekci definic tedy nemá formální datový model. Pro tento účel došlo k vytvoření ontologie ubyCat definující sémantiku lingvistických termínů. Samotná konverze dat je docílena XML[5] transformací. Do výčtu konvertovaných UBY dat patří: WordNet, FrameNet,VerbNet, anglický a německý Wikislovník spolu s příspěvky na OmegaWiki a také spoje mezi těmito slovníky.

Jako poslední krok jak dosáhnout skutečné integrace na sémantickém webu, je propojení lemonUby na další jazykové zdroje v LLOD cloudu. Propojení referenční modelu OLiA[6] představuje

[1] Z anglického lemonUby – a large, interlinked, syntactically-rich lexical esource for ontologies
[2] Lexicon model for Ontologies
[3] Lexical Markup Framework
[4] Implementace ISO 12620:2009 s několika rozšířeními
[5] Extensible Markup Language
[6] Ontologies if Lingquistic Anotations

pro ubyCat formální definici anotačních schémat (jako ontologií lingvistických termínů) pro různé lingvistické fenomény v přibližně 70 jazycích. Na konci kapitoly autoři píší o propojení lemonUby s dalšími lexikálními zdroji jako příklad používají WordNet, kde speciálně u verze 3.0 docílí až 99.99% pokrytí.

V samotném závěru autoři zmiňují využití lemonUby, o kterém je zmínka na začátku tohoto textu a také nastiňují možný budoucí vývoj. Za cíle vytyčují další propojování jak mezi prvky lemonUby, ale také s okolními zdroji na internetu a prohlubování využití v aplikacích zpracování přirozeného jazyka postavených na RDF za účelem zkoumání přínosu takového zpracování dat.

[1] ECKLE-KOHLER, Judith, Christian CHIARCOS a John PHILIP MCCRAE. LemonUby: a large, interlinked, syntactically-rich lexical resource for ontologies. Semantic Web [online]. IOS Press, 2014, , 1-7 [cit. 2016-11-25]. Dostupné z: http://www.semantic-web-journal.net/system/files/swj515.pdf

[2] Lemon: The Lexicon Model for Ontologies [online]. 2016 [cit. 2016-11-25]. Dostupné z: http://lemon-model.net/

[3] GUREVYCH, Irina, Judith ECKLE-KOHLER, Silvana HARTMANN, Michael MATUSCHEK, Christian M. MEYER, Christian WIRTH: UBY – a large-scale unified lexical-semantic resource based on LMF, Proceedings of EACL, pp. 580–590, 2012, Avignon, France.