

existujících modulů. Při návrhu bude kladen důraz na maximální obecnost a znovuvyužitelnost implementovaných modulů.

Předpokládaný postup práce

Fáze I. SW projekt

V rámci SW projektu bude navržena architektura frameworku, vstupy/výstupy, rozhraní jednotlivých plug-inů atd. Dále bude implementováno jádro frameworku, uživatelsky příjemná vizualizace dat, podpora uživatelských vstupů a jednoduché metody odvozování gramatiky a generování dat pro alespoň 3 různé jednodušší typy formátů.

Fáze II. Diplomové práce

V rámci případných navazujících diplomových prací (jejich realizace může probíhat při nebo po realizaci SW projektu) bude možné např. navrhnout a implementovat sofistikovanější metody pro analýzy netriviálních typů formátů (stromové, semi-strukturované, grafové, ...), efektivnější metody generování dat, podporu neobvyklých formátů jako je např. NASA formát FITS apod.

Literatura

1. *BigDataBench, A Big Data Benchmark Suite*. ICT, Chinese Academy of Sciences. <http://prof.ict.ac.cn/BigDataBench>
2. Zijian Ming, Chunjie Luo, Wanling Gao, Rui Han, Qiang Yang, Lei Wang, Jianfeng Zhan: *BDGS: A Scalable Big Data Generator Suite in Big Data Benchmarking*. CoRR abs/1401.5465 (2014). <http://arxiv.org/abs/1401.5465>
3. Sherif Sakr Mohamed Gaber: *Large Scale and Big Data: Processing and Management*.
4. Tilmann Rabl: *Big Data Generation*. Middleware System Research Group, University of Toronto.
5. Transaction Processing Performance Council (TPC): <http://www.tpc.org/>
6. <http://fits.gsfc.nasa.gov/>