Design Automation Renegades

Globetrotting Division

BibTeX Analytics: For Automating Reference Management and Recognizing Emerging Trends

Zhiyang Ong [1]

A document on *Python*-based BibTeX analytics
For Reference Management ...
and Emerging Trend Recognition

May 21, 2018

[1]Email correspondence to: ongz@acm.org

Abstract

This documents how the repository of the BibTeX *Analytics* project is organized, and its software architecture. It also describes the future goals of the project for using a data analytics approach to recognize emerging trends in research, especially emerging research trends in electrical and computer engineering, computer science, and other fields, such as medicine, agriculture, and environmental science.

Insert abstract here.

More stuff to be included.

Revision History

Revision History:

1. Version 0.1, May 21, 2018. Initial copy of the report.
2.

Contents

Chapter 1

Organization of the BibTeX Analytics Repository

The main deliverables of the BibTeX *Analytics* project are a *Python*-based software to perform reference management, and data analytics on BibTeX entries to recognize emerging research trends.

The organization of the BibTeX *Analytics* repository is described as follows:

1. `automated_regression_testing.py`:
   (a) *Python* script to automate regression testing.
2. big_input:
   (a) Data set for stress testing the software deliverable of the BibTeX *Analytics* project.
3. database
   (a) *bibtex_database* (`bibtex_database.py`) class to represent a BibTeX database of BibTeX entries.
   (b) `bibtex_database_test.py` is a *Python* script to test the functionality of the *bibtex_database* class.
   (c) *entry* (`entry.py`) class to represent each (instance of a) BibTeX entry.
   (d) `entry_test.py` is a *Python* script to test the functionality of the *entry* class.
   (e) `key_check.py` is a *Python* script to check if each BibTeX entry is valid.
   (f) `key_check_test.py` is a *Python* script to test the functionality of the *Python* script `key_check.py`.
   (g) `key_frequency_pairs.py` is a *Python* script to demonstrate how to sort a set of 2-tuples based on its first/former field and its second/last/latter field.
4. `duplicate_BibTeX_entries.py`:
   (a) A *Python* script to determine if duplicate BibTeX entries exist in a BibTeX file/database. If such entries exist, warn the user that duplicate BibTeX entries exist.
5. `editions.py`:
   (a) A *Python* script to display a set of editions from all the BibTeX entries in a BibTeX file/-database.
6. `incremental_test.py`:
   (a) A *Python* script to incrementally test features for performing data analytics operations with BibTeX files/databases.
7. input:
   (a) A set of BibTeX files to test my *Python*-based BibTeX *Analytics* software.

8. `institutions.py`:
   (a) A *Python* script to display institutions from each BIBTEX entry in a BIBTEX file/database.
9. `journal_titles.py`:
   (a) A *Python* script to display journal titles from each BIBTEX entry in a BIBTEX database.
10. `keywords_display.py`:
    (a) A *Python* script to display keywords from each BIBTEX entry in a BIBTEX database.
11. makefile:
    (a) For build automation of *Python* scripts, not placed in subdirectories, in the repository.
12. notes:
    (a) `gpl-license.text`, `LICENSE`, and `mit-license.text` are text files of the GNU General Public License (GNU GPL) (`gpl-license.text`) and The MIT License (`LICENSE` and `mit-license.text`)
    (b) guidelines:
        i. A document containing a set of guidelines on how to collaborate with me.
    (c) report:
        i. This document that describes the organization of the BIBTEX *Analytics* repository, the software architecture of the BIBTEX *Analytics* software, and future work of the BIBTEX *Analytics* project.
13. `organizations.py`:
    (a)
14. `publishers.py`:
    (a)
15. `readme.md`:
    (a)
16. `rm_bibtex_metadata.py`:
    (a)
17. sandbox:
    (a)
18. statistics:
    (a)
19. `tutti_series.py`:
    (a)
20. utilities:
    (a)
21. `validate_url.py`:
    (a)
22. `z_booktitles.py`:
    (a)
23.

Chapter 2

Software Architecture of the BibTeX Analytics Project

Software Architecture of the BibTeX *Analytics* Project

Chapter 3

Future Work

Future work of the BibTeX *Analytics* project is described as follows:

1.
2.
3.
4.
5.
6.
7.
8. Find emerging research trends to consider pivoting towards, or to get involved in side projects
   (a) E.g., benchmark adiabatic quantum computers with topological computers and universal quantum computers [1].

Bibliography

[1] Prateek Tandon, Stanley Lam, Ben Shih, Tanay Mehta, Alex Mitev, and Zhiyang Ong. Quantum Robotics: A Primer on Current Science and Future Perspectives, volume 10 of Synthesis Lectures on Quantum Computing. Morgan & Claypool Publishers, San Rafael, CA, January 2017.