

TabGenie : A Toolkit for Table-to-Text Generation

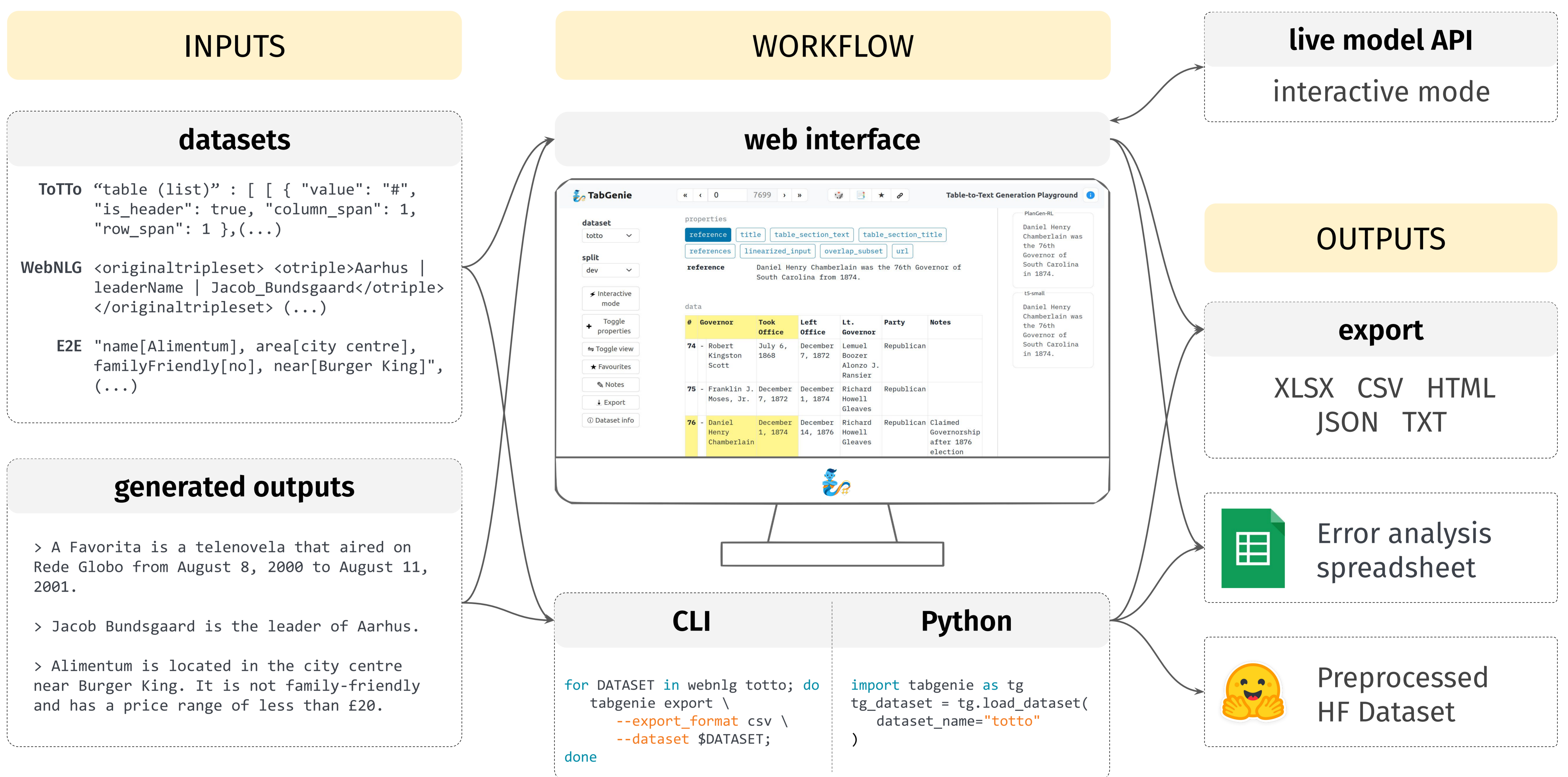
Zdeněk Kasner
kasner@ufal.mff.cuni.cz

Ekaterina Garanina
e.garanina@student.rug.nl

Ondřej Plátek
oplatek@ufal.mff.cuni.cz

Ondřej Dušek
odusek@ufal.mff.cuni.cz

Unified representation makes the work with data-to-text generation datasets easier



TabGenie is a software toolkit which provides a unified **web, command line, and Python interface** for visualizing and processing **16 data-to-text generation datasets**. ↓

Implementation

Backend: Flask + Jinja, HuggingFace datasets.

Frontend: HTML5, CSS, Bootstrap, JavaScript, jQuery.

Try it!

> `pip install tabgenie`

<https://github.com/kasnerz/tabgenie>

Code + live demo



Dataset	Source	Data type
CACAPO	van der Lee et al. (2020)	Key-value
DART	Nan et al. (2021)	Graph
E2E	Dušek et al. (2019)	Key-value
EventNarrative	Colas et al. (2021)	Graph
HiTab	Cheng et al. (2021)	Table
Chart-to-text	Kanharaj et al. (2022)	Chart
Logic2Text	Chen et al. (2020b)	Table + Logic
LogicNLG	Chen et al. (2020a)	Table
NumericNLG	Suadua et al. (2021)	Table
SciGen	Moosavi et al. (2021)	Table
SportSett:Basketball	Thomson et al. (2020)	Table
ToTto	Parikh et al. (2020)	Table
WebNLG	Ferreira et al. (2020)	Graph
WikiBio	Lebret et al. (2016)	Key-value
WikiSQL	Zhong et al. (2017)	Table + SQL
WikiTableText	Bao et al. (2018)	Key-value

Presented at ACL 2023 – System Demonstrations, Toronto.

Supported by the ERC (No. 101039303), the Charles University projects GAUK 40222 and SVV 260698, and the Apple NLU Research Grant for Heriot-Watt University and Charles University. Garanina's work is supported by the Erasmus Mundus Master's Programme "Language and Communication Technologies". Resources provided by the LINDAT/CLARIAH-CZ Research Infrastructure and Peregrine HPC cluster (University of Groningen).



CHARLES UNIVERSITY



university of groningen