Knight Research Network

Tool Demonstration Day

Demonstrations of digital tools developed by universities and companies

Wednesday, October 13
10:30-3:00pm
A virtual event. Registration Required:
https://cmu.zoom.us/meeting/register/tJcsde2oqjwiHd2lfLWxTCALWlamF8zrJc2H

Botometer• BotBuster• BotHunter• BotSlayer• CauseBox• CoVaxxy• Co-Verifi COVID-19 news verification system• ExFacto• Exploring Relations between Online and Offline Data in a Pandemic• FBAdTRacker• Hashtag & URL Coordination• Hate Speech Detection• Hoaxy• Image Emotion Classifier• The Lumen Database• ORA• Netmapper• PIEGraph• Quiz Creator• smaberta• StoryGraph• Troll Hunter• Twitter V2 Conversation and Timeline Collector and V2 to V1 Tweet converter• urlExpander• youtube-data-ai

ABOUT

The Knight Research Network (KRN) was created in 2019 when the John S. and James L. Knight Foundation invested in centers and projects with the goal of "identifying how society can respond to the ways in which digital technology has revolutionized the production, dissemination and consumption of information" (https://knightfoundation.org/democracy-in-the-digital-age/)

Digital tools developed by members of KRN and our friends at other institutions further this goal by making it easier for researchers to gather and sort data, bring data-driven information to the public in innovative, visual formats, fact check news and other information, and detect bots and trolls online.

The KRN Tool Demonstration Day is a free, virtual event, open to the public.
Register here: https://cmu.zoom.us/meeting/register/tJcsde2oqjwiHd2lfLWxTCALWlamF8zrJc2H

The Knight Research Network Tool Demonstration Day is planned and organized by KRN member centers CMU IDeaS; NYU CsMaP; IU OSoMe; and UNC CITAP.

SCHEDULE

Track 1 10:30-2:00 https://cmu.zoom.us/j/93124047879?pwd=ck1QR1hFS0hvelYyQVNQVzBwN3NBZz09 Meeting ID: 931 2404 7879 Passcode: 100580							**Track 2** 11:00-2:45 https://cmu.zoom.us/j/96181988768?pwd=VzVIVUltWTFQUkdCdWN6a3BhNExLQT09 Meeting ID: 961 8198 8768 Passcode: 778537					
Theme		**Start Time**	**End Time**	**Tool**	**Demonstrator**		**Theme**		**Start Time**	**End Time**	**Tool**	**Demonstrator**
Opening Remarks		10:30	11:00	Welcome	Kathleen M. Carley							
Detection	1	11:00 AM	11:15 AM	Botometer	Kai-Cheng Yang		Data Collection/ transformation/ aggregation	1	11:00 AM	11:15 AM	youtube-data-api	Megan Brown
	2	11:15 AM	11:30 AM	BotBuster and BotHunter	Lynnette Ng			2	11:15 AM	11:30 AM	smaberta	Megan Brown
	3	11:30 AM	11:45 AM	BotSlayer	Pik-Mai Hui			3	11:30 AM	11:45 AM	urlExpander	Megan Brown
	4	11:45 AM	12:00 PM	Hate Speech Detection; TrollHunter	Joshua Uyheng			4	11:45 AM	12:00 PM	Quiz Creator	Jessica Collier
		12:00 PM	12:15 PM	BREAK				5	12:00 PM	12:15 PM	EXPO2O	Bohan Jiang
		12:15 PM	12:30 PM	BREAK				6	12:15 PM	12:22 PM	Twitter V2 Conversation and Timeline Collector and v2 to v1 Tweet Converter	Isabel Murdock
Fact checking/ verifiers	5	12:30 PM	12:45 PM	StoryGraph	Alexander Nwala			7	12:22 PM	12:30 PM	Image emotion classifier	Lynnette Ng
	6	12:45 PM	1:00 PM	Hoaxy	Christopher Torres			8	12:30 PM	12:38 PM	Hashtag & URL Coordination	Tom Magelinski
	7	1:00 PM	1:15 PM	CoVerifi COVID-19 news verification system	Nikhil Kolluri				12:38 PM	12:45 PM	BREAK	
	8	1:15 PM	1:30 PM	ExFacto	Anubrata Das				12:45 PM	1:00 PM	BREAK	
	9	1:30 PM	1:45 PM	FBAdTracker	Ujun Jeong		Data Visualizers	9	1:00 PM	1:15 PM	ORA	Kathleen M. Carley
	10	1:45 PM	2:00 PM	The Lumen Database	Adam Holland			10	1:15 PM	1:30 PM	NetMapper	Kathleen M. Carley
								11	1:30 PM	1:45 PM	PIEGraph	Deen Freelon
								12	1:45 PM	2:00 PM	CoVaxxy	Matthew DeVerna
								13	2:00 PM	2:15 PM	Twitter Simulation in Construct	Stephen Dipple
								14	2:15 PM	2:30 PM	CauseBox	Paras Sheth
							Closing Remarks		2:30 PM	2:45 PM		

INFORMATION ON TOOLS

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
Botometer	https://botometer.osome.iu.edu/	Kai-Cheng Yang Pik-Mai Hui Christopher Torres Alexander Nwala Matthew DeVerna John Bryden Filippo Menczer	Indiana University Observatory on Social Media	Tweet data from Twitter's API.	Bot scores ranging from 0 to 1.	The tool has a website and API endpoint. The website is freely available to the public given that the user has a Twitter account. The API endpoint is free to use with a limited quota given that the user has a Twitter developer account. The API endpoint also has paid plans that allow users to make more requests.

Contact: yangkc@iu.edu
Description: Botometer checks the activity of a Twitter account and gives it a score. Higher scores mean more bot-like activity. https://doi.org/10.1038/s41467-018-06930-7

Tool	URL	Demonstrator	Company or Center	Input	Output	Free or purchase
BotBuster	forthcoming	Lynnette Ng	Carnegie Mellon University IDeaS/CASOS	Twitter V1, V2 JSON; Reddit JSON	CSV	Not currently available for public use

Contact: lynnetteng@cmu.edu
Description: BotBuster uses a mixture-of-experts approach to bot detection algorithm. This approach deals with incomplete data due to data collection limitations or account suspension. Each input (e.g. username, screen name, text) is trained individually with specific treatments to their quirks and separate predictions are performed corresponding to the available information. The predictions are then combined in a gating network to output a bot probability.

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
BotHunter v1	http://cerebro.isri.cmu.edu:8008	David Beskow	Carnegie Mellon University IDeaS/CASOS	Twitter JSON, v1	CSV	Free, limited access

Contact: info@netanomics.com

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
BotHunter v2	Forthcoming	Netanomics	Carnegie Mellon University IDeaS/CASOS	Twitter JSON, v1 & v2	CSV	purchase/educational discount

Contact: info@netanomics.com
Description: BotHunter - A tiered Approach to Detection and Characterizing Automated Activity on Twitter. http://www.casos.cs.cmu.edu/publications/papers/LB_5.pdf

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase

BotSlayer	https://osome.iu.edu/tools/botslayer	Pik-Mai Hui, Kai-Cheng Yang, Christopher Torres, Alexander Nwala, Matthew DeVerna, John Bryden, Filippo Menczer	Indiana University Observatory on Social Media	A user-generated query and a Twitter API key to collect tweets	Dashboard that ranks likely malicious entities and provides various statistics and visualization.	Free

Contact: huip@iu.edu

Description: BotSlayer is an application that helps track and detect potential manipulation of information spreading on Twitter. BotSlayer uses an anomaly detection algorithm to flag hashtags, links, accounts, and media that are trending and amplified in a coordinated fashion by likely bots. A Web dashboard lets users explore the tweets and accounts associated with suspicious campaigns via Twitter, visualize their spread via Hoaxy, and search related images and content on Google.
https://ojs.aaai.org//index.php/ICWSM/article/view/7370

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
CauseBox	https://github.com/paras2612/CauseBox	Paras Sheth, Ujun Jeong, Ruocheng Guo, Huan Liu, K. Selcuk Candan	Data Mining and Machine Learning Lab, Arizona State University	Treatment Effect Estimation model and Benchmark Data	Evaluation measures like PEHE, Policy Risk, error on ATE, etc.	The tool is free for use

Contact: psheth5@asu.edu

Description: CauseBox is a unified platform meant to serve as a benchmark for an ensemble of machine learning and deep learning based treatment effect estimation methods. It allows users to run and compare seven state of the art treatment effect estimation methods against benchmark datasets widely accepted in the causal inference literature. This tool is helpful for researchers who want to compare their own methods against benchmark methods. CauseBox supports GUI as well as command line interface.

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
CoVaxxy	https://osome.iu.edu/tools/covaxxy	Matthew DeVerna, Kai-Cheng Yang, Pik-Mai Hui, Christopher Torres, Alexander Nwala, John Bryden, Filippo Menczer	Indiana University Observatory on Social Media	Tweets related to COVID-19 vaccines collected in real-time (using the Twitter API's filtered streaming endpoint), since January 4th, 2020.	An interactive web-based data visualization dashboard. Pictures (.png) of visualizations. Tweet IDs for academic research via rehydration	Free

Contact: mdeverna@iu.edu

Description: CoVaxxy is a web-based data visualization dashboard that allows users to concurrently explore the relationship between COVID-19 vaccine conversations, vaccine uptake, and epidemic trends in the United States. The dashboard tracks and quantifies credible information and misinformation narratives over time, as well as their sources and related popular keywords. Furthermore, vaccine uptake and conversation statistics are visualized geographically at the U.S. state-level. The dashboard is updated daily and the data that the dashboard utilizes is made publicly available for others to rehydrate via the Twitter API. https://ojs.aaai.org/index.php/ICWSM/article/view/18122

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
CoVerifi COVID-19 news verification system	https://github.com/nlkolluri/CoVerifi	Nikhil L Kolluri	The University of Texas at Austin	Text data, social media API-feeds, News API feeds, and other API-derived data	It provides multiple outputs (including user ratings, machine learning outputs, and Botometer results) which indicate the likelihood of news being fake or fact. Users can also add their own assessments of this and these data are publicly outputted as votes are collected.	Free

Contact: nlkolluri@utexas.edu

Description: Manual fact checking is unable to cope with the large volumes of COVID-19-related fake news that now exist. To help address the need to classify this fake news proliferation in the COVID-19 'infodemic', we developed CoVerifi, an automated open-source tool to verify COVID-19 news and information. CoVerifi integrates crowdsourcing, newsfeeds, social media, and machine learning. Users of the web-based tool also have the ability to "vote" on news content, making the CoVerifi platform an effective method to collect labelled data. To develop our fake news detection tool, we built a crowdsourced dataset of ~7000 entries, which we tested and validated our CoVerifier model with. Ultimately, CoVerifi empowers users to make their unused consumption decisions by providing various points of data, rather than labeling news content as fake or fact. https://www.sciencedirect.com/science/article/abs/pii/S2468696421000070

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
ExFacto	https://exfacto.herokuapp.com/	Anubrata Das	University of Texas at Austin	Text data in a search box.	a) a set of evidence related to the claim b) Stance of each piece of evidence c) source reputation of the presented evidence d) overall veracity of the claim	We utilize this tool is for studying human-AI interaction in the context of fact-checking, not meant for production.

Contact: anubrata@utexas.edu

Description: Although we have seen a plethora of research in automated fake news detection, in practice, a large part of fake news detection efforts relies on human labor. Lack of adoption of fake news detection algorithms can be attributed to the complex nature of the problem and the high cost of error. We propose a tool that aims to close the gap by assisting human fact-checkers in their decision-making process. This tool adopts the methodology of evidence-based explainable fact-checking. Users can type a claim into a search box. With the help of search engines such as Bing, or Google, the tool retrieves a set of evidence relevant to the claim. Further, the tool calculates the stance of each piece of evidence and the reputation of each source. It aggregates the evidence to provide a veracity outcome. Users can also override model components (stance and reputation of the evidence) if the model makes a mistake. The model takes user input into account to update the claim veracity outcome. https://doi.org/10.1145/3242587.3242666; https://arxiv.org/abs/1907.03718

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
EXPO2O	TBD	Bohan Jiang, Mansooreh Karami, Anique Tahir, Huan Liu	Arizona State University DMML Lab	COVID-19 related online and offline geospatial data from different sources.	A concise, intuitive, and interactive data virtualization dashboard.	Free

Contact: bjiang14@asu.edu

Description: EXPO2O is a web-based dashboard that provides concise, intuitive, and interactive COVID-19 data visualization for users. Our dashboard allows users to visualize the potential relationship relations between various online-online, offline-offline, and online-offline data. EXPO2O also aims to improve interdisciplinary research on exploring relations between various types of data in a pandemic. In the demo, we will show preliminary findings and insights from the data we have collected so far:

1. Online data: Google trends data, social media data, news media data;
2. Offline data: COVID-19 related statistics, US census data, local events/protests/policies.

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
FBAdTracker	http://tweettracker.engineering.asu.edu:5001	Ujun Jeong, Kaize Ding, and Huan Liu	DMML in Arizona State University	Keywords and options for searching Facebook Advertisements	Collected Facebook Advertisements and analysis on advertisements/advertisers	This tool is provided by webpage and users can simply use the system by the interface

Contact: ujeong1@asu.edu

Description: The purpose of this application is to provide an integrated data collection and analysis system for current research on fact-checking related to Facebook advertisements. Our system is capable of monitoring up-to-date Facebook ads and analyzing ads retrieved from Facebook Ads Library. https://arxiv.org/abs/2106.00142

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
Hashtag & URL Coordination	Forthcoming	Isabel Murdock, Lynnette Ng, Tom Magelinski	Carnegie Mellon University IDeaS/CASOS	• Twitter Dataset- File (json or gziped json) or Directory (of jsons or gziped jsons) • Type of coordination (hashtag, URL, or both) • Time window (in minutes) • Output filename (csv)	Edgelist (user-user-type-weight) for the coordination network that can be imported into ORA (.csv import)	Free

Contact: iem@andrew.cmu.edu

This tool constructs networks of Twitter users who are tweeting the same hashtags or same URLs within a short time window, in which tight clusters correspond to coordinated users.

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
Hate Speech Detection	Found on CASOS Servers; a public version will be available soon.	Joshua Uyheng/Netanomics	CASOS Center, Institute for Software Research, Carnegie Mellon University	Netmapper Cues files.	CSV file with multiple levels of hate speech probabilities	Purchase with educational discount

Contact: juyheng@cs.cmu.edu

Description: The CASOS Hate Speech Detection model uses psycholinguistic features to predict the likelihood that a given tweet is hate speech. Due to multiple - and at times conflicting - definitions of hate speech, our model is trained on multiple datasets and produces likelihoods optimized for these different definitions. This affords users the ability to select the most appropriate predictions depending on their definition of choice, or run experiments with multiple definitions of hate speech for robustness. Across definitions, the model is trained using theoretically anchored features that allow for meaningful interpretations of results in relation to social identities and conflicts.

Hate Speech Detection: https://doi.org/10.1007/978-3-030-80387-2_12

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
Hoaxy	https://hoaxy.osome.iu.edu/	Christopher Torres, Kai-Cheng Yang, Pik-Mai Hui, Alexander Nwala, Matthew DeVerna, John Bryden, Filippo Menczer	Indiana University Observatory on Social Media	The application uses Twitter data.	A network visualization of interactions, such as retweets, quote retweets, replies, between different accounts. The tool also allows the user to see how the network evolved over time, displays bot scores, and provides links to the tweets. The output can be downloaded to allow future reproduction or external analysis.	Free

Contact: torresch@indiana.edu

Description: Hoaxy provides an easy-to-use way to visualize the spread of information on Twitter. The user has the ability to visualize the spread of articles which they query from a list of URLs that are associated with fact-checking sources and low-credibility domains. Additionally, we leverage the Twitter API to allow users to visualize any search query that works on the Twitter search bar. Anatomy of an online misinformation network (http://doi.org/10.1371/journal.pone.0196087)

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
Image Emotion Classifier	Forthcoming	Isabel Murdock, Lynnette Ng, Tom Magelinski	Carnegie Mellon University IDeaS/CASOS	Image	CSV	Free

Contact: iem@andrew.cmu.edu

Description: Image Emotion Classifier is makes use of machine learning to provide the probability of images identifying with each of Plutchik eight emotional categories: anger, fear, sadness, disgust, surprise, anticipation, trust and joy. It is trained using 8000 images tagged with the respective categories on Flickr. It has been applied in a case study to identify emotions in images in an emotional event - the Kashmir Black Day event. https://link.springer.com/chapter/10.1007/978-3-030-80387-2_18

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase

The Lumen Database	www.lumendatabase.org	Adam Holland Shreya Tewari Chris Bavitz Peter Hankiewicz	Berkman Klein Center for Internet & Society, Harvard University	We accept copies of requests to remove content from the web, usually in the form of fielded data sent through our API. Requests to view the data can come through the API or a browser interface. notices can be human-readable or JSON.		Free

Contact: team@lumendatabase.org

Description: Lumen is an independent research project studying cease and desist letters concerning online content. We collect and aggregate requests to remove material from the web. Initially focused on requests submitted under the United States' Digital Millennium Copyright Act, Lumen's database, which offers API access for notice submitters and researchers, now includes complaints of all varieties, including those concerning trademark, defamation, and privacy, both domestic and international. Currently, the Lumen database contains approximately 17 million of removal requests, referencing 4.5 billion URLs, and grows by more than 20,000 notices per week, from companies such as Google, Twitter, YouTube, Wikipedia, Reddit, Medium, Github, Vimeo, and Wordpress.

Complete API documentation is available at https://github.com/berkmancenter/lumendatabase/wiki/Lumen-API-Documentation

https://www.wsj.com/articles/google-dmca-copyright-claims-takedown-online-reputation-11589557001 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3687861

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
ORA	http://www.casos.cs.cmu.edu/projects/ora/	Kathleen M. Carley	Netanomics	csv or json network or attribute files	html,csv, json	There is a lite free version and a full professional version for purchase. Educational discount available

Contact: kathleen.carley@cs.cmu.edu

Description: A network analysis toolkit for graphical, statistical and visual analytics on both social networks and high dimensional networks that can vary by time and/or space. ORA is a full function network analytics package that supports the user in creating, importing, exporting, manipulating, editing, analysing, comparing, contrasting, and forecasting changes in one or more networks. ORA pro can handle networks with millions of nodes and includes BEND analytics and a stance detector.

ORA: A Toolkit for Dynamic Network Analysis and Visualization (pdf) http://www.casos.cs.cmu.edu/publications/papers/CMU-ISR-20-110.pdf

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
NetMapper	https://netanomics.com/netmapper-government-commercial-version/	Kathleen M. Carley Eric Malloy	Netanomics	json .txt .pdf .csv	csv output or xml for networks	For purchase with educational discount

Contact: kathleen.carley@cs.cmu.edu

Description: Computational linguistic tool for extracting semantic networks, meta-networks, sentiment and cues from texts and social media posts. Netmapper operated in over 40 languages. It also can extract phone numbers, emojis and emoticons. NetMapper User Guide v12 9/2021 (pdf). http://sbp-brims.org/2018/proceedings/papers/Demos/ORA%20&%20NetMapper.pdf

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
PIEGraph	https://pcad.ils.unc.edu/	Deen Freelon,	University of North	Web domains from the user's Twitter timeline	See description	Free

		Drew Crist, Meredith Pruden	Carolina at Chapel Hill			

Contact: freelon@email.unc.edu

Description: PIEGraph is an interactive chart that displays web domains that have recently appeared in the user's Twitter feed in a scatterplot. The x-axis represents the domains' left-right ideological orientation, while the y-axis represents content credibility. The values for both axes were generated by the Media Bias Fact Check organization (https://mediabiasfactcheck.com/). The size of each bubble represents the relative prevalence of each domain--domains appearing more often have larger bubbles.

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
Quiz Creator	https://mediaengagement.org/quiz-creator/	Jessica Collier	Center for Media Engagement	The tool allows participants to respond to multiple choice or slider questions	The tool output includes views, quiz starts, completion percentages, and correct/incorrect responses.	Free

Contact: jessica.collier@austin.utexas.edu

Description: The Quiz Creator is a simple online tool using a step-by-step process to create a quiz in as little as 3 minutes. The interface is customizable to assist in seamless integration on any site. The quiz is embeddable to any page on a website and allows for tests of audience knowledge and response rate. Users can also create A/B tests to see which quizzes are most effective. We have done a series of projects to test the benefit of this tool: - https://www.tandfonline.com/doi/full/10.1080/19331681.2016.1230920; https://www.tandfonline.com/doi/full/10.1080/19331681.2019.1680475; and a working paper here: https://jessicareneecollier.files.wordpress.com/2021/08/quizzes-working-paper-1.pdf

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
smaberta	https://pypi.org/project/smaberta/	Megan Brown, Rachel Connolly	Center for Social Media and Politics	Labelled text data	A trained transformer model	Free

Contact: meganbrown@nyu.edu

Description: smaberta is a Python client for creating transformer-based classifiers in Python. Adapted from simple transformers, smaberta allows researchers to train, evaluate, and predict using minimal code. https://csmapnyu.org/scholarly-articles/

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
StoryGraph	Website: https://storygraph.cs.odu.edu/ Twitter Account: https://twitter.com/storygraphbot	Alexander Nwalam, Kai-Cheng Yang, Pik-Mai Hui, Christopher Torres, Matthew DeVerna, John Bryden, Filippo Menczer	Indiana University Observatory on Social Media	The application reads the RSS feeds of 17 US news organizations. No query input is required by the user.	A graph visualization where the nodes represent news articles, and an edge between a pair of nodes represents a high degree of similarity between the nodes (similar news stories).	Free

Contact: anwala@iu.edu

Description: StoryGraph quantifies the level of attention given to new stories by 17 US left, center, and right news media organizations. The service generates a news similarity graph every 10-minutes, where each news story is assigned an attention score indicating the magnitude of attention it receives collectively from the news media organizations. 365 Dots in 2019: Quantifying Attention of News Sources (https://arxiv.org/abs/2003.09989)

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
TrollHunter	forthcoming	Joshua Uyheng	CASOS Center, Institute for Software Research, Carnegie Mellon University	NetMapper Cues files.	CSV file with troll probabilities	Will be available soon. Professional version available from Netanomics in 6 months.

Contact: juyheng@cs.cmu.edu

Description: TrollHunter uses psycholinguistic features to predict the likelihood that a given account is a troll. Due to conflicting definitions of trolling, we opt for an empirically grounded operationalization that emphasizes the use of abusive, targeted language. TrollHunter relies on properties of not only the messages of the account of interest, but also any messages to which the account of interest may be interacting. This allows for context-aware predictions that align with our understanding of trolling as an interactive - and disruptive - phenomenon.

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
Twitter Simulation in Construct	http://www.casos.cs.cmu.edu/projects/construct/	Stephen Dipple	CASOS Center, Institute for Software Research, Carnegie Mellon University			

Contact: kathleen.carley@cs.cmu.edu

Description: Construct, developed by CASOS, is a multi-agent model of network evolution. In Construct individuals and groups interact communicate, learn, and make decisions in a continuous cycle. The program takes into account how agents learn through interaction conducted over different media and change their information, beliefs, and activities based on what they learn. This can be used for forecasting how a network can evolve and seeing if two groups that appear identical on one dimension actually evolve in the same way. Training and Sample Data: http://www.casos.cs.cmu.edu/projects/construct/sample.php

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase

Twitter V2 Conversation and Timeline Collector and v2 to v1 Tweet Converter	https://github.com/CASOS-IDeaS-CMU/twitter_conversation_collection	Isabel Murdock, Lynnette Ng, Tom Magelinski	Carnegie Mellon University IDeaS/CASOS	No data needed for the collection scripts. For the v2 to v1 format tweet converter, input tweets should be in the format of the direct response JSON from the v2 API.	Twitter data in the format of the Twitter API v2.	Free

Contact: iem@andrew.cmu.edu

Description: Set of python scripts for collecting Twitter user timeline data, conversations, recent search tweets, full archive search tweets, sampled stream tweets, filtered stream tweets, and user profile information using the Twitter API v2. The scripts take care of the requests and writing out the collected data. Additionally, scripts are provided that convert the collected tweets from v2 format to v1 format so that they can be compatible for existing tools/software created for the v1 format of data. The v2 to v1 converter code can also be run in a standalone fashion for data collected through other methods.

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
urlExpander	https://pypi.org/project/urlexpander/	Megan Brown, Rachel Connolly	Center for Social Media and Politics	Link text (or JSON payloads from some social media sites)	JSONs of the expanded link	Free

Contact: meganbrown@nyu.edu

Description: This package makes working with link data from social media and webpages easier. It not only expands links, but catches errors, and makes parallel link expansion quick and efficient. https://csmapnyu.org/scholarly-articles/

Tool	URL	Demonstrator(s)	Company or Center	Input	Output	Free or purchase
youtube-data-api	https://pypi.org/project/youtube-data-api/	Megan Brown Rachel Connolly	Center for Social Media and Politics	Query inputs	JSON outputs	Free

Contact: meganbrown@nyu.edu

Description: youtube-data-api is a Python client to download public YouTube data about channels, videos, and searches. https://csmapnyu.org/scholarly-articles/