

1	Planning and Direction	2.3.3	Web crawler and/or web scraper	3.3.4	Statistical methods (e.g. aggregation via mean value, first-difference estimator)	4.3.4	Machine Learning (esp. Random Forest and Naive Bayes algorithms)
1.1	Cyber Security	2.3.4	Other open source tools (e.g. Impala shell tool)	3.3.5	Deep learning (text vectorization, e.g. using the word2vec algorithm)	4.4	Journalism, competitive analysis, general approach
1.1.1	Manual (e.g. manual definition of objectives and requirements)	2.4	Journalism, competitive analysis, general approach	3.4	Journalism, competitive analysis, general approach	4.4.1	Manual (expert analysis)
1.2	Health	2.4.1	Manually (e.g. via databases)	3.4.1	Manual (manual cleanup, labeling)	4.4.2	Deep learning (especially word embedding, for semantic analysis using different versions of the BERT algorithm)
1.2.1	Manual (e.g. manual requirements definition)	2.4.2	API interface	3.4.2	Standardized methods/algorithms/tools (e.g. parser and other word segmentation tools, for lemmatization, word/character filtering)	4.4.3	Machine learning (in particular Support Vector Machines or Random Forest and XGBoost algorithms)
1.3	Security	2.4.3	Web crawler and/or web scraper	3.4.3	Natural language processing filters and labeling methods/tools (e.g. using similarity algorithms and classification algorithms)	4.3.4	Machine Learning (esp. Random Forest and Naive Bayes algorithms)
1.3.1	Manual (e.g. manual hypothesis formulation)	3	Processing and Exploitation	4	Analysis and Production	5	Dissemination and Integration
2	Collection	3.1	Cyber Security	4.1	Cyber Security	5.1	Cyber Security
2.1	Cyber Security	3.1.1	Manual (e.g. labeling/review/verification/grouping via experts, search parameters in databases/search engines, manual cleansing)	4.1.1	Manual (content analysis)	5.1.1	Files/reports (e.g. CSV file, PDF dossier)
2.1.1	Manually (e.g. via databases and search engines)	3.1.2	Keyword/dictionary/hashtag filter (usually in combination with web crawler and web scraper)	4.1.2	Standardized methods/algorithms and tools (image hashing, open source tools, lexical approaches)	5.1.2	Dashboard/visualization map
2.1.2	API interface	3.1.3	Standardized methods/algorithms/tools (e.g. parser and other word segmentation tools, for normalization, lemmatization, word/character filtering, lowercase conversion and duplicate removal)	4.1.3	Tool stack	5.1.3	Web interface/web application/online platform
2.1.3	Web crawler and/or web scraper	3.1.4	Statistical methods (e.g. oversampling methods, N-gram algorithms, pattern matching)	4.1.4	Deep learning (especially word embedding, for semantic analysis using different versions of the BERT algorithm)	5.1.4	Automated alerts
2.1.4	Other open source tools (e.g. Python tools, data pipelines, stream listeners and lookup features)	3.1.5	Natural Language Processing filters and labeling methods/tools (e.g. heuristic methods, topic classification, part-of-speech tagging, entity and relation annotation, named entity recognition, e.g. using the CoreNLP toolkit or Python NLTK toolkit)	4.1.5	Machine learning (in particular Support Vector Machines or Random Forest, XGBoost, lightGBM, Naive Bayes and logistic regression algorithm)	5.1.5	Graph creation (e.g. using Matplotlib, Networkx, Pygraphistry or the Neo4j browser)
2.1.5	Web application (e.g. for the parallelization of searches in IoT search engines)	3.1.6	Deep learning (text vectorization, e.g. using the word2vec algorithm)	4.1.6	Artificial intelligence (e.g. neural network using BiGRU layers)	5.2	Health
2.2	Health	3.2	Health	4.2	Health	5.2.3	Web interface/web application/online platform
2.2.1	Manual (e.g. via databases, search engines and surveys)	3.2.1	Manual (e.g. selection by experts, expert interviews, search parameters in databases/search engines, manual cleansing)	4.2.1	Manual (content analysis)	5.3	Security
2.2.2	API interface	3.2.2	Keyword/dictionary/hashtag filter (usually in combination with web crawler and web scraper)	4.2.2	Statistical methods (meta-analysis)	5.3.1	Dashboard/visualization map
2.2.3	Web crawler and web scraper	3.2.3	Standardized methods/algorithms/tools (e.g. parsers and other word segmentation tools, for word and character filtering and file conversion tools)	4.2.3	Machine learning (in particular Support Vector Machines or XGBoost and Naive Bayes algorithms as well as Open Source tools)	5.4	Competitive analysis, general approach
2.2.4	Other open source tools (e.g. social media analysis tool and command line tool)	3.3	Security	4.3	Security	5.4.1	Dashboard/visualization map (e.g. using Power BI)
2.3	Security	3.3.1	Manual (e.g. verification via experts, search parameters in databases/search engines)	4.3.1	Standardized methods/algorithms and tools (e.g. dictionary comparison)		
2.3.1	Manual (e.g. via databases, search engines and surveys)	3.3.2	Keyword/dictionary/hashtag filter (usually in combination with web crawler and web scraper)	4.3.2	Statistical methods (correlation analysis, time series analysis, panel regression models)		
2.3.2	API interface	3.3.3	Standardized methods/algorithms/tools (e.g. parser)	4.3.3	Deep Learning (in particular vectorization algorithms)		