

Torch.AI™ Advances the Capabilities of Acquiring and Enriching Publicly Available Information (PAI)

Nexus™ human-centric machine intelligence transforms and enriches data in-flight to deliver speed-of-mission access to PAI

Torch.AI's pioneering methodology of enabling PAI and automatically creating knowledge graphs allows users to focus on the critical task of analyzing processed information. Our triple-entity extraction process builds-out the picture of an initial network in real time, and identifies new entities for users to explore. With Nexus in the data ecosystem, the user can obtain analytic outcomes in much less time.

Nexus connects the dots so you don't have to.

PAI Acquisition

PAI acquisition and analysis is driven by the Nexus platform, an open architecture of microservices, ML-enabled algorithms, and data intelligence capabilities developed by Torch.AI. Nexus enables massively scaled, ultra-high-performance text and document-based data processing.

It employs several techniques for acquiring PAI, including:

- ✓ Non-attributable and obfuscated web scraping and caching
- ✓ Crawlers and spider services
- ✓ Ingestion and parsing of social media feeds
- ✓ Native integration with Dark Web aggregators, including communication and market sites such as Reddit, China's Weibo, Russia's VK, and Telegram

Entity Extraction

Nexus transforms and enriches data in flight by automating the tagging, normalization, and deduplication processes. Nexus includes a set of Natural Language Processing (NLP) models and file parsers that can identify attributes in the data such as language, sentiment, personally identifiable information (PII), locations, actors, currencies, dates, company names, and human names. The NLP models are continually retrained on language- and domain-specific content for increasing accuracy.

Nexus from Unstructured Data

Torch.AI frees the user from manually mining unstructured data to make connections by automatically creating knowledge graphs from contextual language. Knowledge graphs reveal entities ("nodes" in graph theory), their relationships ("edges"), and properties of both. Through Nexus, users create connections across nodes, analyzing relationships, patterns, and thematic topics.

Nexus creates knowledge graphs by applying the following NLP-enabled processes to the ingested data:

- ✓ Extracting noun-verb-noun triplets from unstructured content. A triplet is a subject, relationship, and object. For example, rather than extracting “Vladimir” (human name entity) and “Moscow” (location entity), we extract the triple “Vladimir lives in Moscow.”
- ✓ Disambiguating triplets to determine whether the person or entity represented in the string matches a person or entity with the same name in the knowledge graph. For example, we can resolve that “Vladimir” who lives in “Estonia” is not the same entity as “Vladimir” who lives in “Moscow,” thus we create a new node in the graph.
- ✓ Performing coreference resolution to identify all expressions that refer to the same entity in a text. Following our example, if the text, “he has a brother named Sergei” is encountered nearby, “he” is resolved to “Vladimir.” By replacing all pronouns with nouns, the data is prepared for higher-level analysis tasks such as document summarization, question answering, and information extraction.
- ✓ By combining these three NLP-enabled processes, using models trained on domain-specific data, Nexus can instantly create large and powerful knowledge graphs.
- ✓ Nexus can access a variety of third-party PAI sources. Exclusive financial market data such as S&P Global, Churuchbase, Bloomberg and Pitchbook are accessible as well as the SEC’s EDGAR database. In addition, Nexus has a rich set of foreign language algorithms, which include speech recognition, language identification, and optical character recognition (OCR).

The Value of Our Solution

Torch.AI’s pioneering methodology of enabling PAI and automatically creating knowledge graphs allows users to focus on the critical task of analyzing processed information. Our triple-entity extraction process builds-out the picture of an initial network in real time, and we identify new entities for users to explore. With Nexus in the data ecosystem, the user can obtain analytic outcomes in much less time.

