University leverages the power of third-party data

The Hebrew University of Jerusalem is the first academic institution to use the LexisNexis® Bulk Content API to train an advanced machine-learning algorithm.

A team of political scientists led by Professor Tamir Sheafer, Dean of the Social Sciences Faculty at the Hebrew University of Jerusalem, is using the LexisNexis® Bulk Content API to deliver massive amounts of content to its news-monitoring system.

The system is designed to identify trends in media discourse. Using content delivered via the Bulk API, the research team are training an advanced machine-learning algorithm to discover deep patterns not easily discerned by the human reader.

Advanced machine learning. Powered by the Bulk Content API.

The research team leverages massive amounts of unstructured data from the unrivaled LexisNexis® content collection.

The Bulk Content API provides fast, one-time access to huge third-party datasets. Users are free to pull and host all the LexisNexis® publication-level licensed news content—whether it's five years of everything or 35 years from a single source.

The team from the Hebrew University of Jerusalem are building their Bulk delivery from the vast LexisNexis® content library—which incorporates niche and industry publications as well as globally recognized outlets.

Thanks to the Bulk Content API, the team has accessed more than enough data to train an advanced machine-learning algorithm to perform discourse analysis. The algorithm is able to spot trends and variations within media coverage—enabling the news-monitoring system to discover deep patterns not easily discerned by the human reader.

Designing better research

The Bulk API's vast reach has allowed the team to track changes in media discourse across location and time period.

The ever-expanding LexisNexis® library of unmatched, vetted news content allows the research team—specialists in applied machine learning—to comparatively analyze a huge corpus of relevant news. As well as news gleaned from the LexisNexis® content library, the data studied by the team at the Hebrew University includes the public records of political institutions (laws, debates, speeches, and committee protocols).

With the freedom to analyze content from multiple news agencies, countries, and time periods, researchers can study communication phenomena in many different contexts. Insights are unlocked that reveal trends in representation of events, values, and actors. Discourse attributes are uncovered.

Big data. Big thinking.

With access to over 35 years of archived publications, the team can discover trends and validate theoretical research models in different contexts.

The ease-of-use inherent in the Bulk API has allowed the team to replicate the search process whenever they need to. Theoretical models are validated and refined within the almost endless expanse of the Bulk content delivery, which gives room to test and retest without ever running out of runway.

Looking beyond his own research, Professor Tamir Sheafer is positive about the potential of the Bulk API to inform other disciplines. Business schools could benefit from access to vast third-party data for the analysis of commercial campaigns. Computer scientists could benefit from access to such large resources of textual data.

What could you achieve with this much data? To give your team access to our developers' portal, sample files, and sample schema, please contact:
Hatim El Faiz on 31 630172525, email Hatim.elfaiz@lexisnexis.nl or visit
internationalsales.lexisnexis.com/solutions/Data-as-a-Service

RELX Group™