DataScava

Unstructured Data Miner

315 Madison Avenue
Suite 901
New York, NY 10017
(646) 701-0055
www.datascava.com
@datascava

WHAT IS DATASCAVA?

SOFTWARE THAT INTERPRETS UNSTRUCTURED DATA USING PURELY DIGITAL (NON-SEMANTIC) LOGIC, YOUR BUSINESS INTELLIGENCE AND MACHINE TRAINING

DataScava
Unstructured Data Miner

PATENTS

U.S. PATENTS 7587395, 7702621
"PROFILE MATCHING OF UNSTRUCTURED DATA"

FIND THE DATA YOU NEED
EXTRACT ITS VALUE

DataScava
Unstructured Data Miner

FOUNDERS

Janet Dwyer, CEO

John Harney, CTO

DataScava
Unstructured Data Miner

80% of the world's data is
UNSTRUCTURED
90% has been created in the last two years

- IBM, May 2016

DataScava
Unstructured Data Miner

UNSTRUCTURED DATA GROWTH

International Data Group

- Unstructured data is growing at the rate of 62% per year. By 2022, 93% of all data in the digital universe will be unstructured.

Gartner

- Data volume is set to grow 800% over the next five years and 80% of it will reside as unstructured data.

DataScava
Unstructured Data Miner

DATA IS USELESS UNLESS YOU CAN

**FIND IT
USE IT
ANALYZE IT
MONETIZE IT**

DataScava
Unstructured Data Miner

2 TYPES OF SEARCH

Research Search

- In research search, the user tries to locate a number of documents which together provide the desired information.

Navigational Search

- In navigational search, the user utilizes the search engine as a tool to navigate to the best overall document.

DataScava
Unstructured Data Miner

3 WAYS TO SEARCH

1 BOOLEAN SEARCH

2 SEMANTIC SEARCH

3 DATASCAVA SEARCH

DataScava
Unstructured Data Miner

BOOLEAN SEARCH

- Uses sets of words with AND, OR, NOT

- Results are too literal and missed matches

- Lacks context, produces many false positives

- Requires skill, effort and SME to create query

- Inability to set required/desired score thresholds

- No analytics or ranking capabilities

- Inability to segment or ratchet up/down search results

- Cannot traverse markup language

DataScava
Unstructured Data Miner

SEMANTIC SEARCH

- Semantics is science of meaning in language

- A search for "Bank of America" finds American banks, banking in America, American banking

- Finds all word forms and no "not" capability

- Invisible, hard-coded and imprecise

- Ignores "noise words" (and, of, if, the)

- No tagging, scoring, matching, ranking, analytics

- Inability to set minimum score thresholds in search topics

- Produces a large number of false positives

- "Semantic is suitable for research NOT navigational search"
    Ramanathan V. Guha, PHD
    Creator of Google Custom Search
    https://en.wikipedia.org/wiki/Semantic_search

DataScava
Unstructured Data Miner

DATASCAVA SEARCH

- Converts unstructured data to structured data

- Non-semantic parse, index, score and match

- Uses your business nomenclature and jargon

- Weights time-sensitive synonym occurrences

- Segmented search and match

- User-defined minimum score thresholds

- Quantified text analytics & percentile scores

- Single click multidimensional rank and sort

- Editable taxonomies built out for I.T. & Finance

- Customizable to any domain or business

- Excels in jargon-intensive industries

- Brings accurate results quickly to the top

DataScava
Unstructured Data Miner

HOW WE DO IT

- Define what you need

- Re-define it as necessary

- Locate precisely where it is

- Transform it as required

- Store and index it

- Quantify its depth

- Categorize it by type

- Prioritize it on-the-fly

DataScava
Unstructured Data Miner

TALENTBROWSER Powered by DataScava

Skills Analytics, Patented Search and Job Matching

A — Indexes millions of data points

B — Using your business nomenclature

C — Matches people across jobs 24/7

D — Built out for I.T., Finance and more

E — Customizable to any industry

TalentBrowser

THE BENEFITS

1 Identify ripe opportunities for data monetization and mining to maximize your data investments

2 Make business decisions that correspond directly to what your data is telling you

3 Gain insights and visibility to improve decision making and support the demands of your business

4 Analyze text-heavy data efficiently & create a reliable, personalized indexer & matching engine

DataScava
Unstructured Data Miner

DataScava

Unstructured Data Miner

315 Madison Avenue
Suite 901
New York, NY   10017
(646) 701-0055
www.datascava.com
@datascava

Thank You!