

Datanami article – Taming Unstructured Data with Cognitive Computing



Contending with unstructured data is no longer a priority reserved for the most well-financed, IT-savvy organizations, like Google and Facebook. As the world's data continues to increase at nearly exponential rates, the reality is the majority of that data is unstructured and incongruent—in its native form—with time-honored tables and SQL-based modeling. The prominence of a critical confluence of technological forces including mobile, social, and the cloud has produced a situation in which the majority of unstructured data is created by consumers.

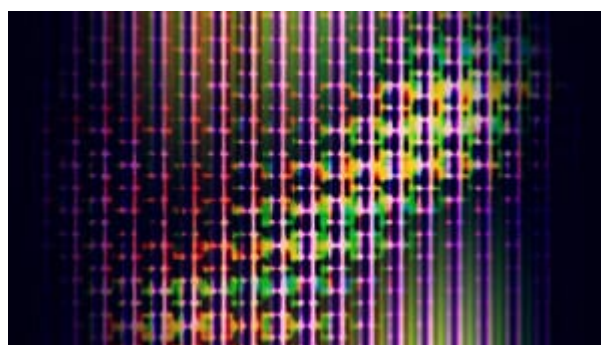
Consequently, traditional methods of managing, transforming, analyzing and, in some instances, even applying that data are no longer sufficient when incorporating such external, unstructured new data into the enterprise. Time-sensitive data simply cannot wait for conventional preparation processes and ETL for historic business intelligence analysis.

Instead, a new paradigm, one heavily reliant upon cognitive computing and its myriad capabilities, is necessary to effectively tame semi-structured and unstructured data, and to do so with the degree of accuracy and celerity necessary for business processes in contemporary times.

Pattern Discovery Creates Context and Structure

Pattern discovery is the basis of cognitive computing's ability to derive meaning from unstructured data. Machine learning algorithms (augmented in certain instances by deep learning and neural networks) can establish precedents from current attributes and uses of data to determine future ones.

Th
e
re
su
lt
is
a
de
gr
ee
of
co
nt
ex
t
th
at
is
va
lu
ab
le
fo
r
's
tr
uc
tu



ring,
unstructured data for nearly every facet of its use. The flexibility of th

es
e
al
go
ri
th
ms
ca
n
de
te
rm
in
e
wh
at
so
rt
s
of
da
ta
ar
e
in
te
gr
at
ed
wi
th
on
e
an
ot
he
r

an
d
ho
w,
as
we
ll
as
cr
ea
te
ac
ti
on
re
la
te
d
to
cl
ea
ns
in
g,
ET
L,
en
ri
ch
me
nt
an
d
im
pu
ta
ti

on

.

They can also expedite the data discovery process and provide analytical insight into vast amounts of data. Importantly, these dynamic algorithms can also greatly inform the results of analytics from recommending actions to providing explanations to quantifiable results—as well as selecting the most eminent dashboards and visualizations to demonstrate their relevance. Their capabilities for determining relationships and discerning context are enhanced by semantic graphs, terminology systems, ontologies and other aspects of semantic technology.

The New Data Modelers

Flexible algorithms are fast becoming the data modelers for unstructured data when deploying cognitive computing technologies, which are vital for generating insight and action for sets of big data.

The conventional modeling process of implementing business requirements, creating the various types of models (logical, conceptual, enterprise, etc.) and redoing them when the data or requirements change, is both arduous and time consuming. Organizations can forego that process by utilizing these versatile algorithms that change as the data is changing, according to use cases. They can also work in conjunction with ontological models for specific requirements such as regulatory concerns.

This automated process enables organizations to effectively include much more data than they previously could and use them for more targeted use cases than was previously possible. Additional boons include increased agility, expedience, and the ability to adapt to changing business requirements.

End User Results

Subsequently, the end user is able to achieve more with unstructured data via cognitive approaches than he or she otherwise could have, enhancing both individual and organizational performance. Whether deployed in master data management systems, data lakes, or big data governance procedures, cognitive computing enables the end user to issue ad hoc queries without code and lengthy IT processes.

These delays have typically restricted the use of data for the business. In the life sciences field, individuals can traverse virtually any number of disparate sources, including different medical conditions and terminology references, to analyze multiple variables in a single query in near real time. Such analysis frequently requires a combination of structured, semi-structured and unstructured data of extreme complexity.



In the retail and ecommerce fields, it is possible to reference factors based on products, customer type, location, weather, sentiment analysis and other relevant variables to determine aspects of marketing and product development. One of the more recent points of focus for the medical industry is chronic diseases. Cognitive computing technologies can readily analyze patient specific data generated by any number of wearable devices to augment internal data and tailor treatments and diagnoses for individuals.

AI Today

The focus of the renewed emphasis on artificial intelligence and the impact that cognitive computing is creating upon the contemporary data landscape is on enabling organizations to do more with external, non-proprietary data.

Cognitive computing technologies allow for the ready

integration of such data with typical structured data to provide a comprehensive overview of whatever particular business function for which those data are used.

They accelerate the processing of such data and deliver more profound insights in less time for the end user. In such a way, cognitive computing brings the enterprise closer to decision automation, while helping to elucidate the context of what has historically been termed dark data.

About the author: Jans Aasman is Ph.D. psychologist and expert in cognitive science, as well as CEO of semantic graph database and analytics company, Franz. As both a scientist and CEO, Jans continues to break ground in the area of artificial intelligence and semantic databases. Jans spent a large part of his professional life in telecommunications research, specializing in intelligent user interfaces and applied artificial intelligence projects. He gathered patents in the areas of speech technology, multimodal user interaction, recommendation engines while developing early versions of the iPad and Siri From 1995 to 2004. He was also a part-time professor in the Industrial Design department of the Technical University of Delft.

