



The Hive Mind

New NLP techniques to cross-pollinate
your cloud and bloom new insights



Moderated by **Colin Gibson**

Senior Advisor, EDM Council

- Colin is a senior executive and consultant with 30 years' experience in financial services
- He specializes in architecture, data governance and data management
- Previously head of data architecture for the investment banking divisions of RBS and HSBC and head of enterprise architecture for Willis Towers Watson
- Colin combines an enthusiasm for all things data with experience in software development, technology operations and management of major change programmes



Today's speakers

Moderator



Colin Gibson

Senior Advisor
EDM Council



Ben Easley

Technical Evangelist
Datorama - Salesforce



Ian McCarty

Chief Product Officer
eContext.ai

The Big Bang of Unstructured Data

“Natural language understanding, the interaction between computers and humans, is the next frontier.”

– Satya Nadella, CEO Microsoft

- Massive data explosion
 - 80%–90% of data is unstructured, **and only 5% of unstructured data is estimated to be analyzed today**
 - Total amount of data stored in the cloud will grow to 100 ZB by 2025*
 - 463 EB/person/day of new data is created
- Overall desire to reduce time & cost of data analysis
 - 46% of enterprises do not have the right technology to take full advantage of their data

*Sources: Fortune, Cybercrime Magazine, Raconteur, Dun & Bradstreet

Technologies for an Unstructured World



Ingestion Technologies



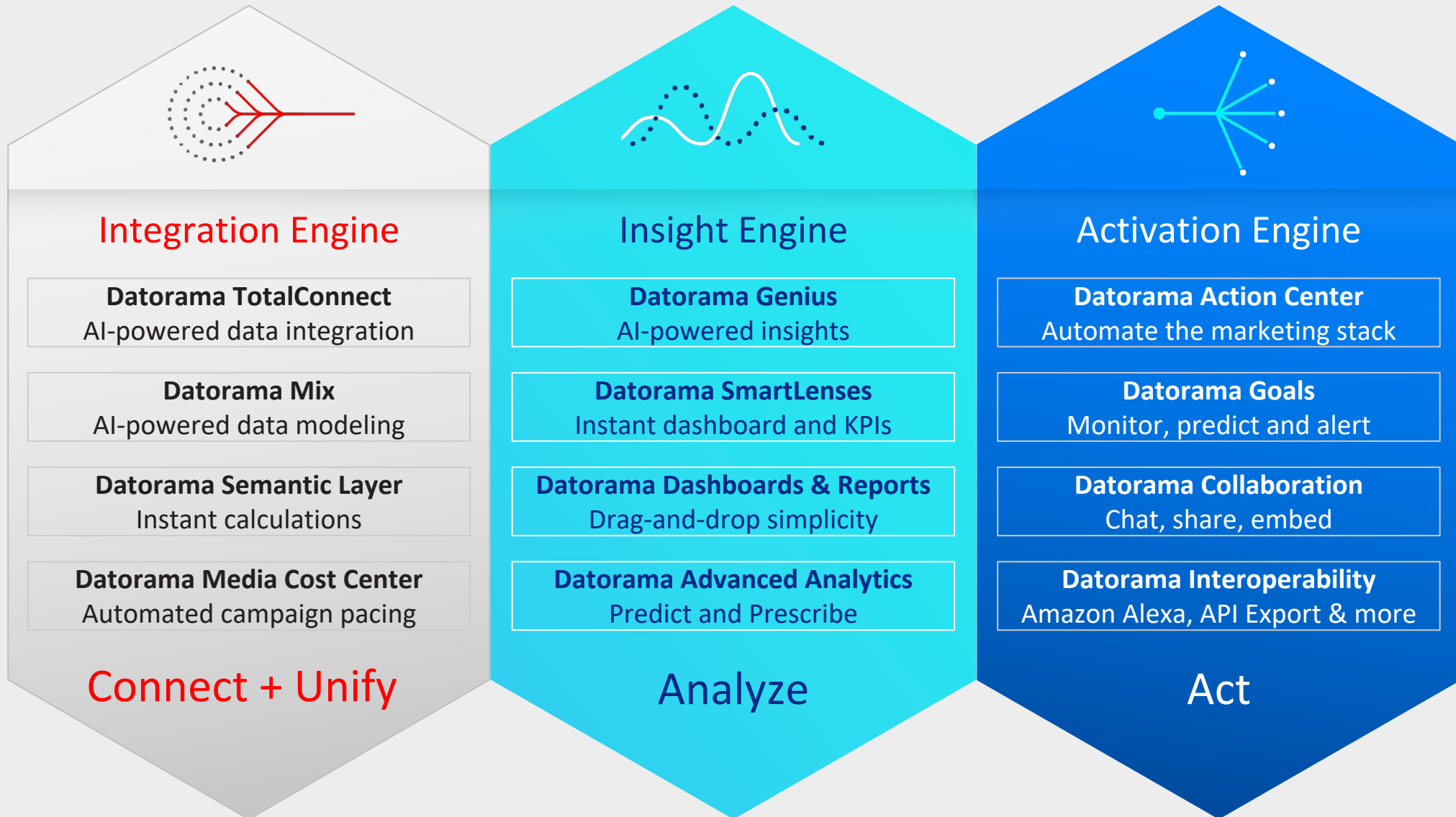
Storage Technologies



Analysis Technologies

- More and more, overlapping with storage technology
- This is a form of cross-pollination, & is usually limited to quantitative measures

Much more than a data warehouse!



Marketers are facing foundational data challenges

salesforce

1

Lack of unified view of marketing investment, performance and results

2

Lack of real-time performance insight to optimize ROI and results

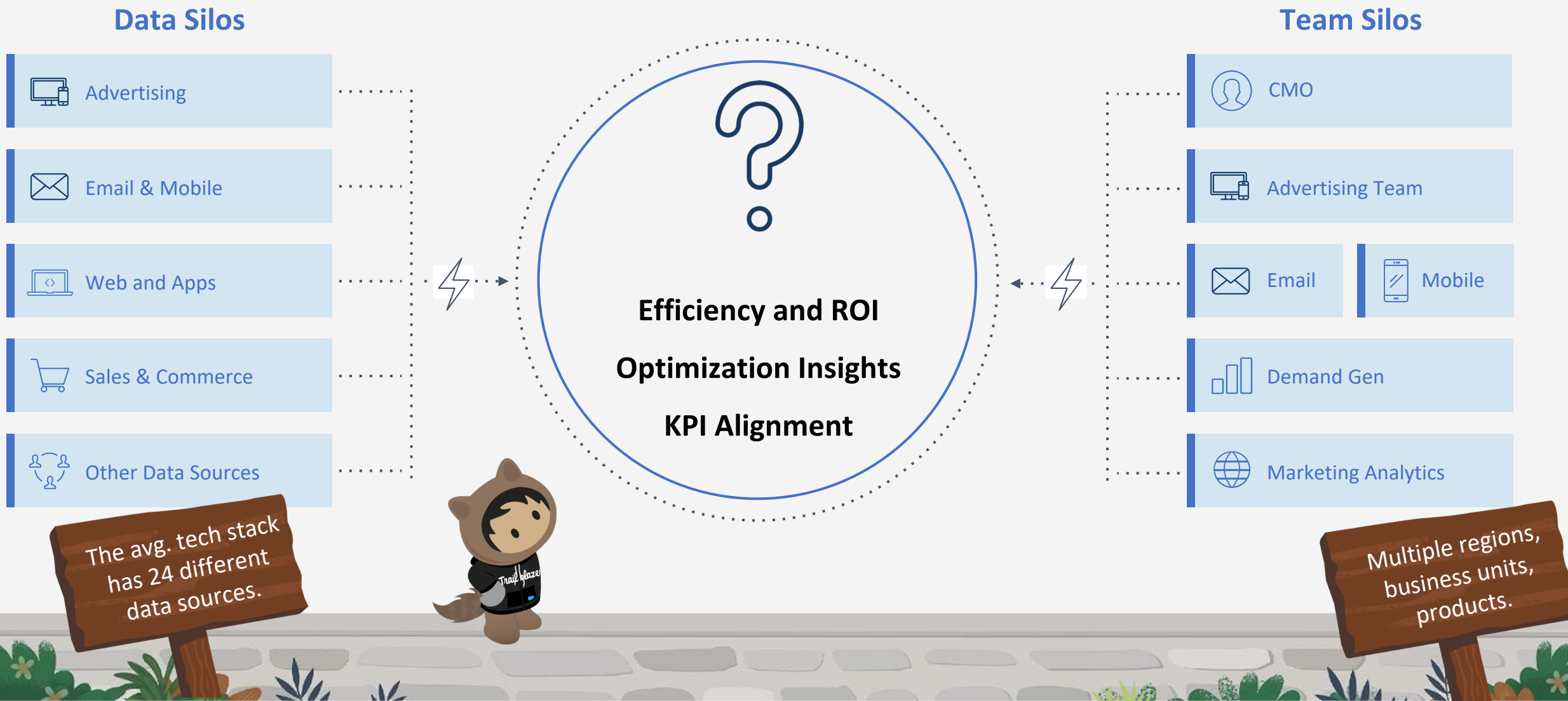
3

Lack of alignment and governance to drive collaboration across teams and regions



Marketers need a system of record

To optimize marketing performance



Analyze & Act → Magic in the Middle



Datorama is not a “dashboard” – It is Marketing Intelligence



Connect & Unify
All Marketing Data



Dashboards and
Reporting



Insights and
Analytics



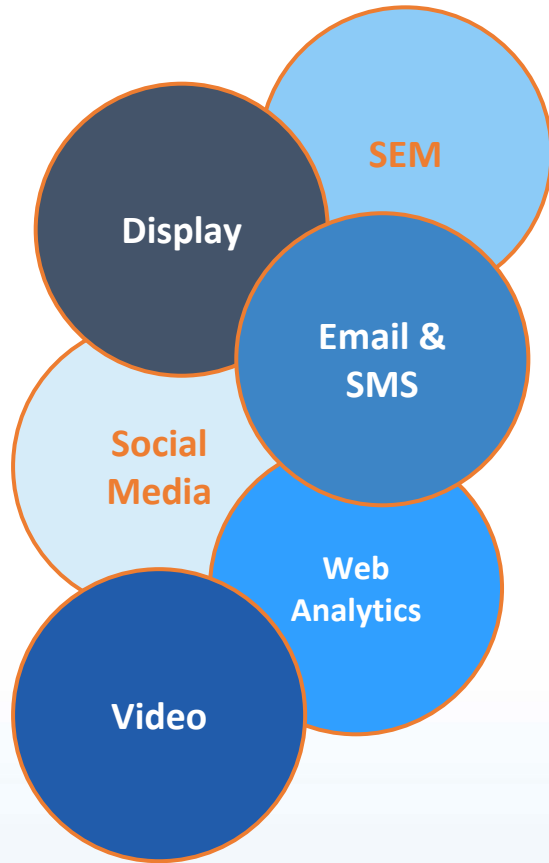
Goals, Alerts and
Activations



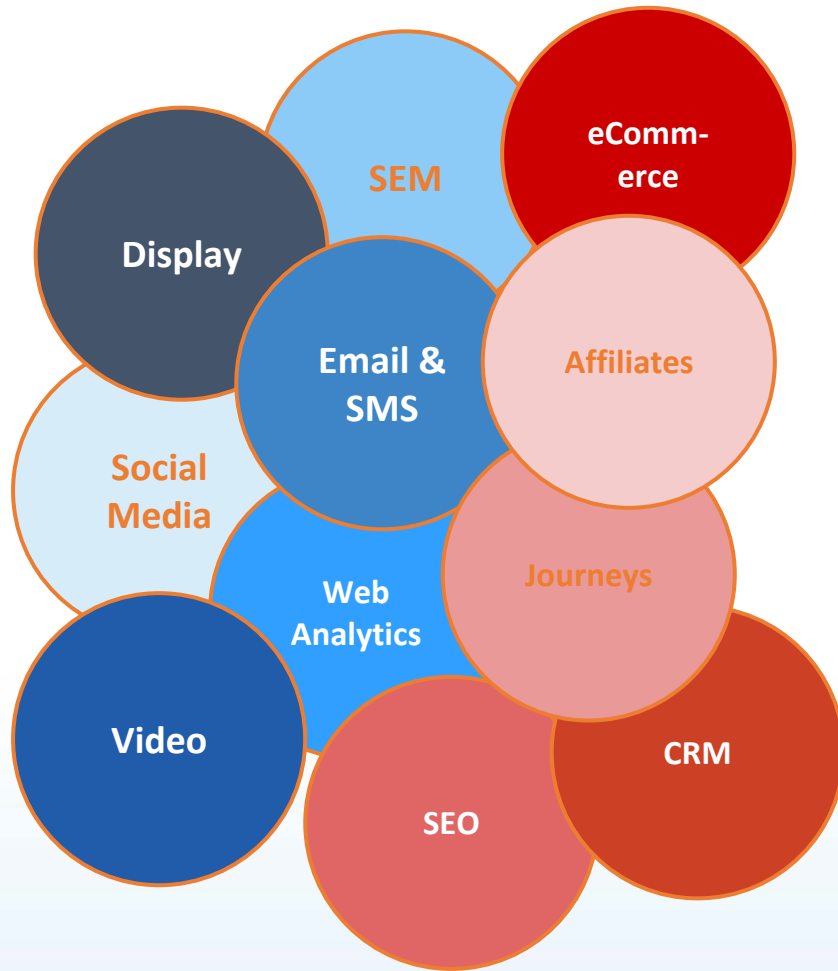
Integrating Complex, Disparate Data Sources



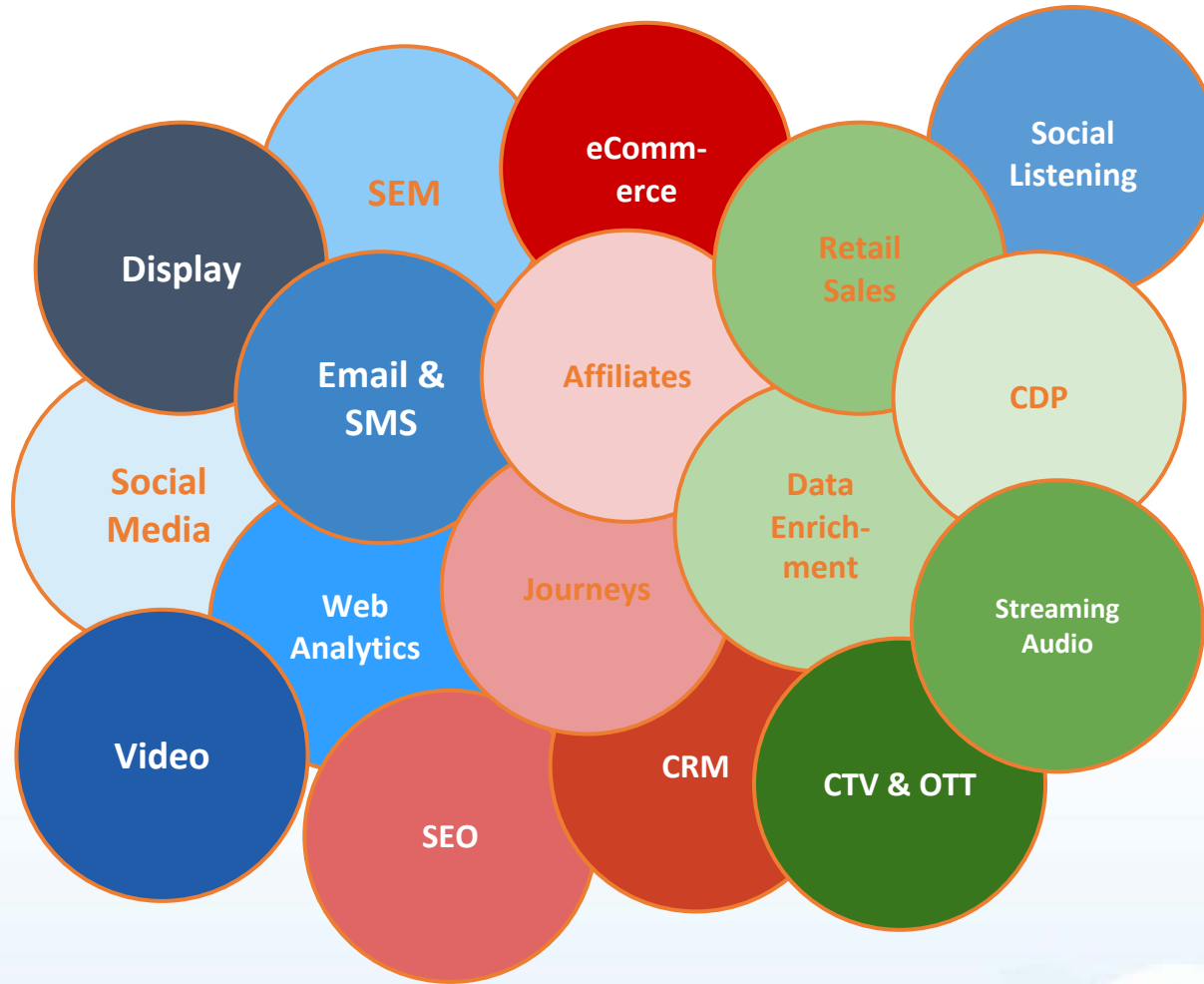
Pushing the Digital Analytics Envelope



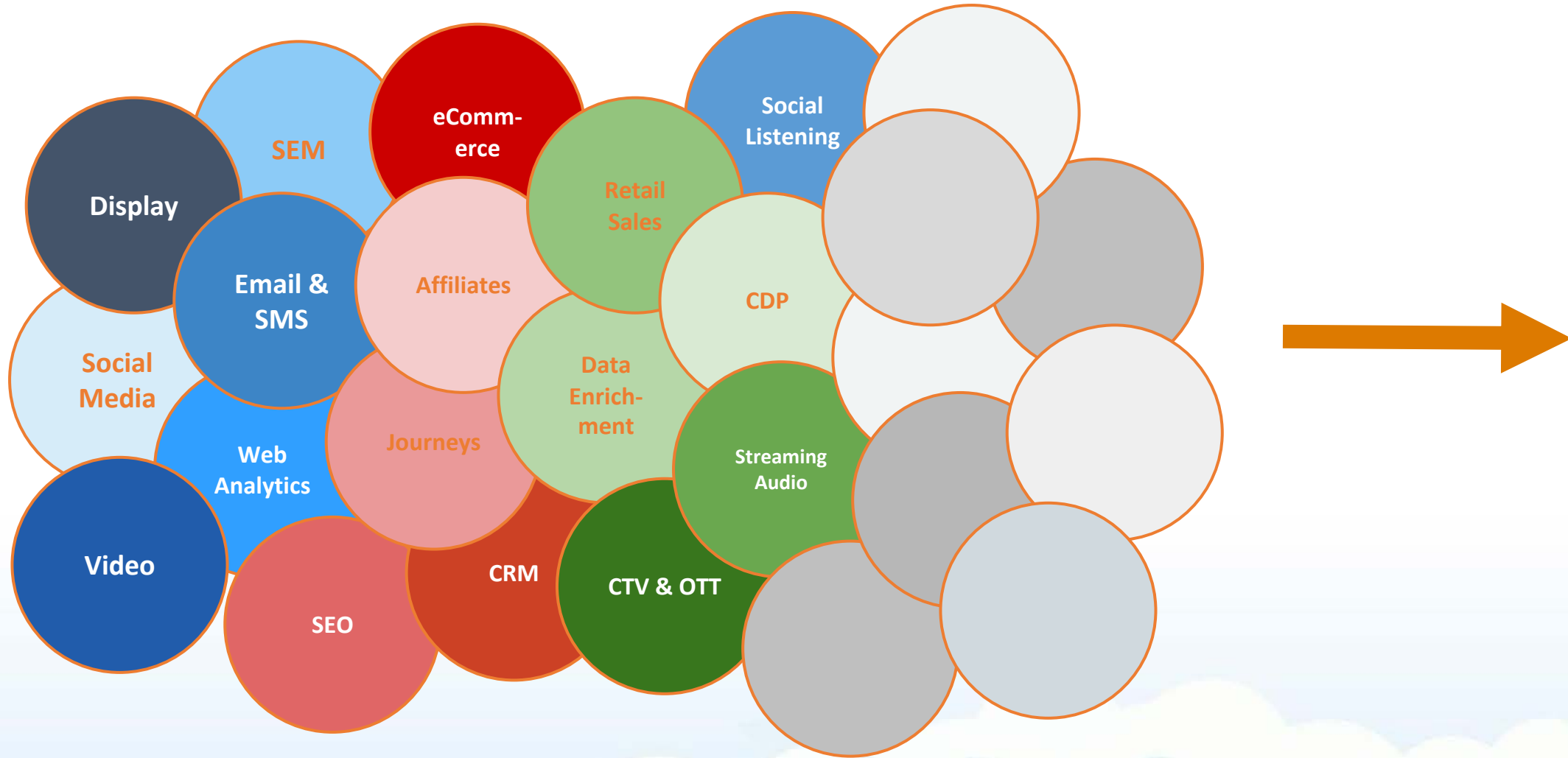
Pushing the Digital Analytics Envelope



Pushing the Digital Analytics Envelope



Pushing the Digital Analytics Envelope





It's Not Just About Integrating and Harmonizing Data...



The Hard Problem of Soft Data

- Many current analytical solutions depend on “hard”, numerical data
- There is usually accompanying “soft”, textual or content data that is under-utilized (or outright discarded)
- These problems compound as the data comes from multiple business units

How Text Data is Typically Processed

Level 1: Basic

- Brand & Keyword Mentions
- Word Clouds
- 3rd Party Shallow Labels

Level 2: Intermediate

- Sentiment Analysis
- Named Entity Recognition

Level 3: Advanced

- Word Embeddings (word2vec, sense2vec)
- Unsupervised Topic Modeling (Latent Dirichlet Allocation, Latent Semantic Indexing)

Customer Generated Taxonomies and Naming Conventions



Stitching it All Together



- Excels at data integration + harmonization

Opportunities

- Enrich existing customer data
- Omnichannel data enrichment unlocks omnichannel analytics in Datorama

Demand for Data Classification & Taxonomy



Strong demand for tools that deal with:

- Data classification
- Data enrichment
- Data governance



Naming Conventions & Taxonomies



- Customers are accustomed to creating/maintaining their own taxonomies
- Tedious and time-consuming endeavor
- Relies heavily on coarse grain, meta data
- Opportunity → create a cross-channel intelligent taxonomy at a more granular level
- Create an omni-channel taxonomy -- video, SEM, SEO, content, web analytics



Wishlist for Cross-Channel Intelligence



Plug & Play

Simple to operate & maintain by most enterprise employees



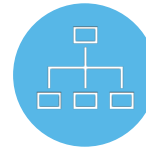
New Knowledge

Infuses new information into the system, rather than a new puzzle



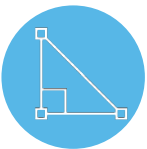
Humanistic

Mirrors the operators understanding



Taxonomic

Conceptually connected within a structure



Calibration-Free

Doesn't need recalibration for every new source it is applied against



Extensible

Integrates and supports existing naming conventions or taxonomies

How Should Text Data Be Processed

Level 4: Expert

- Supervised classification models & services
- Large-Scale Hierarchical Text Classification System, or
LSHTCS

How Should Text Data Be Processed

A New Challenger!

Scale-Optimized Channel-Agnostic Text Classification Hierarchy
or, **SO CATCHY**

- Supervised Model with a very large number of classes (10,000+), balanced across different disciplines and verticals
- Trained on the *smallest piece of data* likely to exist in practice (“quark”)
- Pre-processing techniques to turn longer documents into quarks
- Organizational hierarchy for the labels

Processing Text Data within a Hierarchy

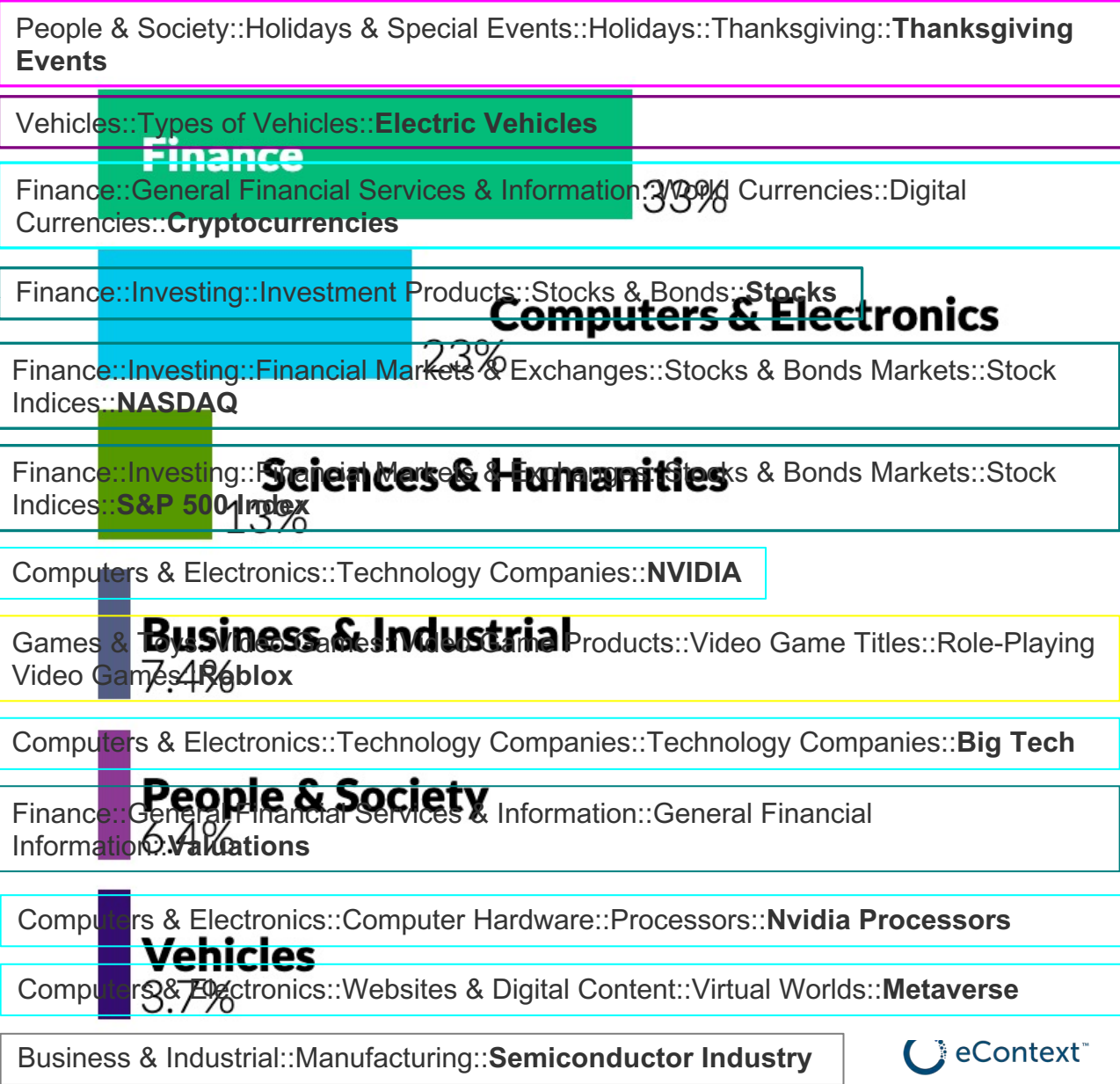
The Metaverse, Crypto and EVs Are Among 2021's Big Tech Winners

When Americans gathered around the Thanksgiving table last week, the blistering rally in technology, electric vehicles and crypto-related stocks was part of their conversations.

There's a reason it dominated the small talk: The tech-heavy Nasdaq 100 is now worth almost half as much as the benchmark S&P 500 -- the highest ever -- and the megacap tech stocks alone represent a third of the S&P 500. Nvidia Corp. and Roblox Corp.'s sprint stood out in a year when the rest of the big tech names jogged to new highs, defying several calls to sell the sector around last year's thanksgiving due to soaring valuations.

Hot Chips

Chipmaker Nvidia has soared 148% as booming chip demand and a foray into the metaverse made it the best performer on the Nasdaq 100. Applied Materials Inc. and Advanced Micro Devices Inc. were other winners, each rising about 80% and outperforming

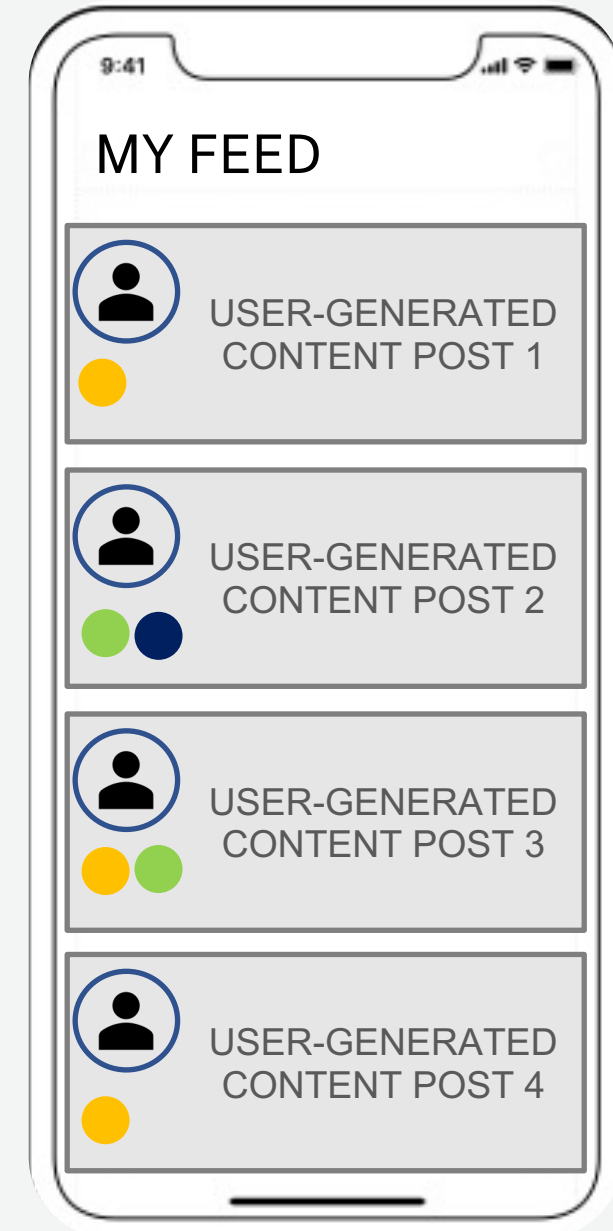
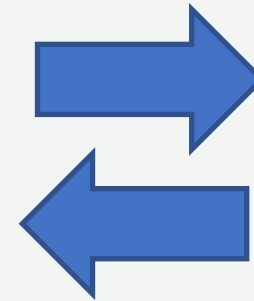
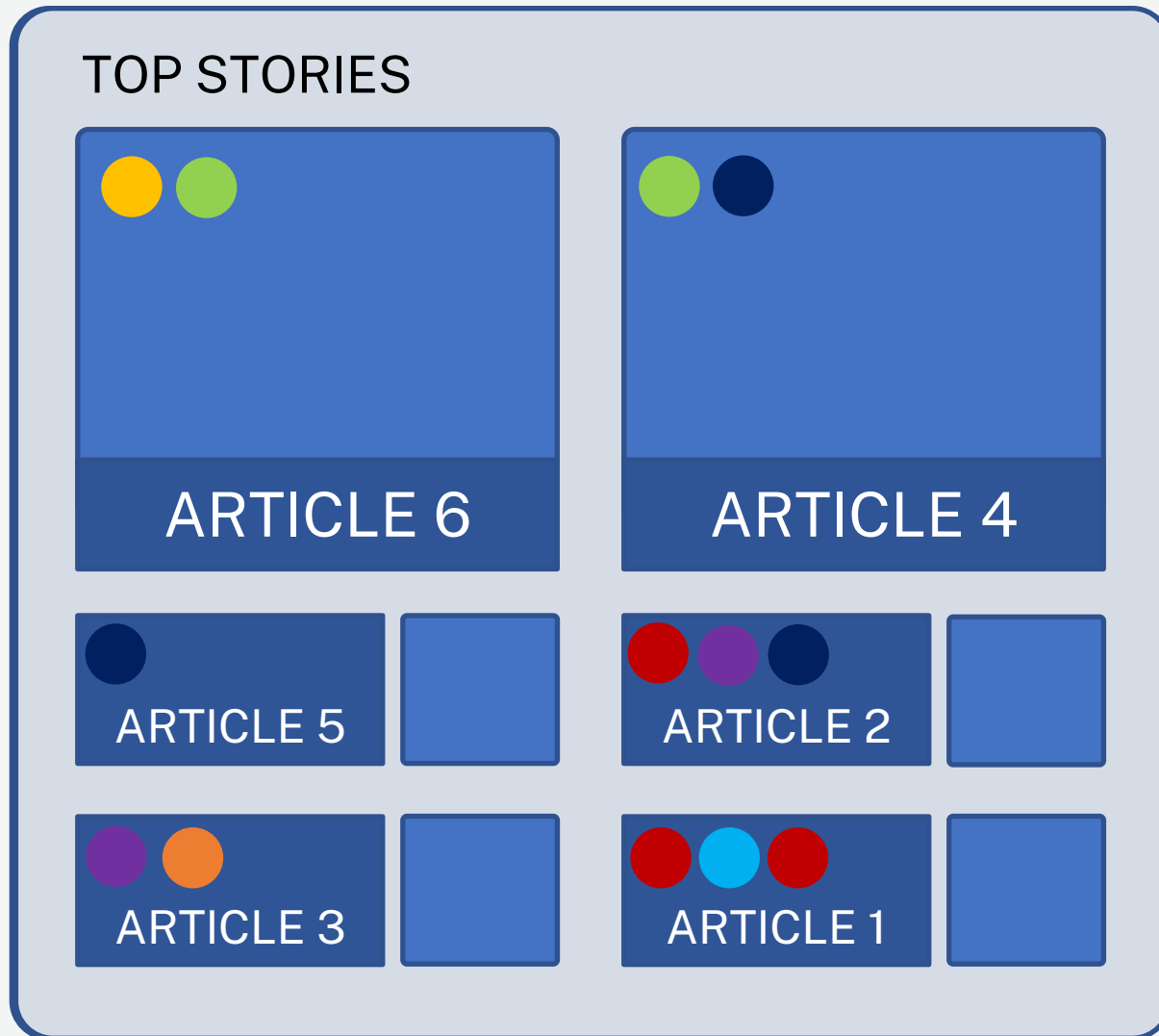


New Analytical Opportunities

- Query Once, Search & Retrieve Everywhere
- Unite Behaviors across Touchpoints
- Mirror Operational Success Across Channels
- Find New B2B Opportunities From Past Wins

New Analytical Opportunities

Personalize Customer Experiences



How to Pursue & Capitalize from Scratch

- Invest in a wide and deep hierarchy
- Collaboration – More cooks improve the broth
- Train at the smallest common denominator
- Rely on NLP, until you can't
- Know when to split & know when to fold

How to Extend and Improve

- Identify intersections between existing enterprise taxonomies
- Discover intersections with “tracer” documents
- Consider a hub-and-spoke approach

eContext – Turning Unstructured Data Into Deep Knowledge



World's Largest Text Classification Engine – **550,000 topics within hierarchy up to 20 tiers deep**



Understand any digital data, including short forms of text



Unites Enterprise Hierarchies



Identify customer interests, organize content, target ads, and more

New Media



Social Media



Podcasts



Video



News

Contextual Targeting



RTB/App Inventory



Brand Safety



Market Research



Search & Browsing



Surveys



Shopping

Customer Service



VAs & Chatbots



Email



Call Centers

Questions?



For more information and a Free Trial,
contact Ian McCarty:
imccarty@econtext.ai

Read our testimonials at
www.econtext.ai/econtext-testimonials/

