Combining NL, ML and AI to Deliver Meaningful Solutions

How to architect a meaningful EIM framework that accelerates delivery

midwest
architecture community
collaboration

SEMANTIC SUITE

midwest architecture community collaboration

The New Imperative

Companies must build an "evidence-based knowledge culture"

- This implies that business decisions are more frequently based on
  - smart humans spending more time exploring possible scenarios because
  - they are not spending all their time fighting with Incomplete, incorrect and disjoint data
- We architect the platforms / frameworks that accelerate the rise of the Citizen Data Scientist
  - These are SMEs in every area of the company — including IT
  - Crowdsourcing is needed for a rising tide instead of peaks and valleys of good decisions
- IT fundamentally changes its focus toward empowering SMEs:
  - Less commoditized IT— maintenance work that looks roughly the same at every company
  - More differentiated work — Enabling IT to become a direct provider of business value

Evidence-Based Management (HBR)
https://hbr.org/2006/01/evidence-based-management

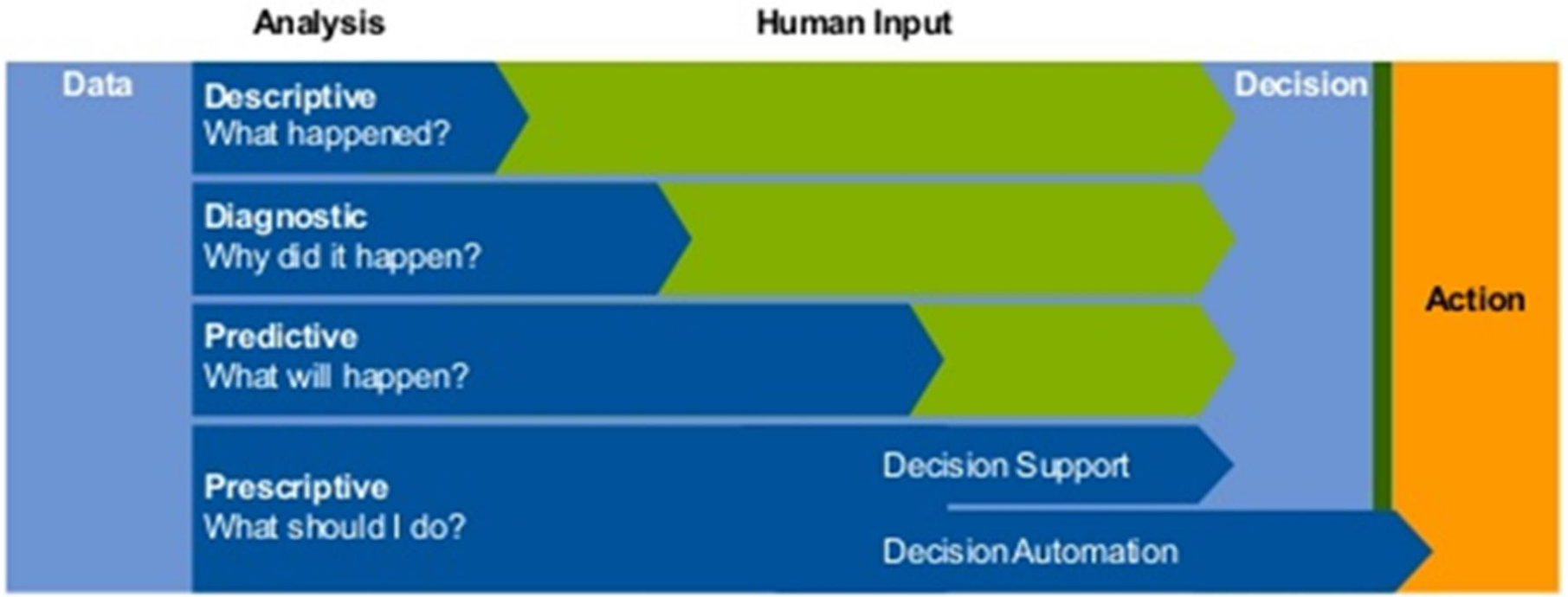
IKE SEMANTIC SUITE

Evidence-Based Management (CEBMA)
https://www.cebma.org/faq/evidence-based-management/

midwest architecture community collaboration

AI Shifts Who Does What

Decision automation, including analytics, is supported by IT, run by SMEs



Gartner

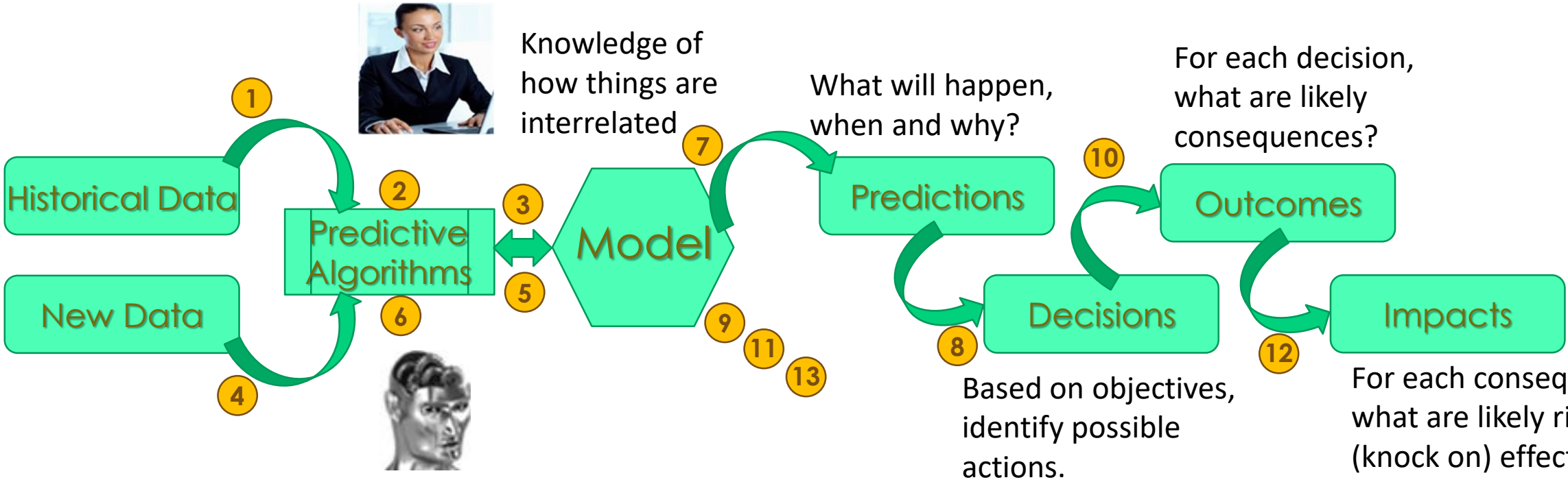
SMEs Curate Digital Meaning
http://understandingcontext.com/2016/12/curating-digital-meaning/

IKE SEMANTIC SUITE

midwest architecture community collaboration

Human or Machine Decision Flow

- Analytics and decision automation can be used to support all business processes



- An explanation utility is essential for validating complex inferences in a DSS or AI

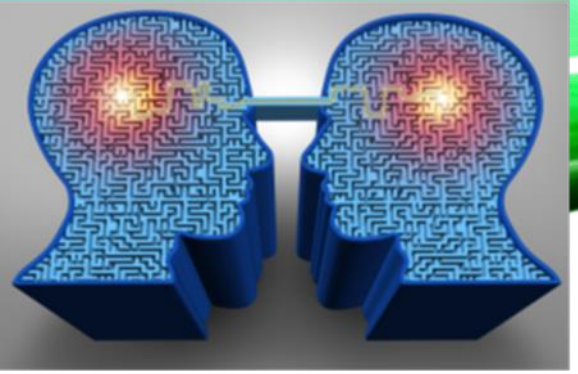
midwest architecture community collaboration

Information is Your Greatest Asset

- We have a clear sense of the general value of information
  - We spend millions to gather, manipulate, store and examine it, but …
- Most organizations do not know the value of most of the information they have
  - We know that we can often use information to gain competitive advantage, but …
  - We do not yet know how each database, document, video, image, audio file can help
  - We might not even know where to find them or how to deliver them to improve decisions
- Knowledge about your data, Metadata, can become Metaknowledge in four steps:
  1. Establish and maintain a robust enterprise information model
  2. Classify all information assets in a single metadata repository
  3. Establish a controlled vocabulary with KPIs, formulas and definitions
  4. Use these to architect a "Knowledge Framework" for AI and ML to improve search, BI, Analytics and *DECISIONS*

Proposed Information Architecture
http://understandingcontext.com/2016/06/enterprise-meta-model-governance/

IKE SEMANTIC SUITE



midwest architecture community collaboration

Commonality in All Information

Whether structured or unstructured, and no matter the subject, all information has…

Meaning

All conversations have utterances, each of which originates from a person with…

Intent

1. Meaning is inherent in the words and can be automatically interpreted with NLP
2. Intent models represent straightforward meaning, and attempt to infer intent from words
3. Intent is based in the mind of the originator, but the words' inherent meaning may not represent the intent, due to metaphor, sarcasm, irony, deception and other subtext
4. Today's AI is still struggling to accurately interpret straightforward intent

Subtext
http://understandingcontext.com/2012/07/understanding-subtext/

IKE SEMANTIC SUITE

Exformation
http://understandingcontext.com/2012/08/exformation-vs-subtext/

midwest architecture community collaboration

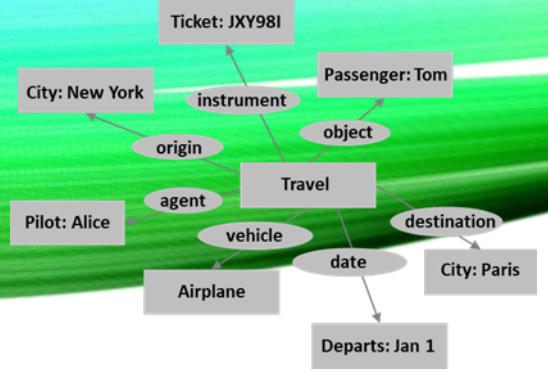
Where the Meaning Hides

Meta-knowledge is a *model* that describes context

- Knowledge and language (framework of meaning) may be even more complex than meets the eye
- Notice that in this model of all information, noise is part of context – this is critical to good filtering
- Current tools and techniques work with Data and Information, and are straining to reach Knowledge
- To automatically derive meaningful and actionable information, use heuristics with meta-knowledge
- Unclean data drags the content toward noise
- Good meta-models and heuristics can elevate the content and outputs toward actionable knowledge

Language Complexity
http://understandingcontext.com/2014/04/form-and-substance-in-communication/

IKE SEMANTIC SUITE



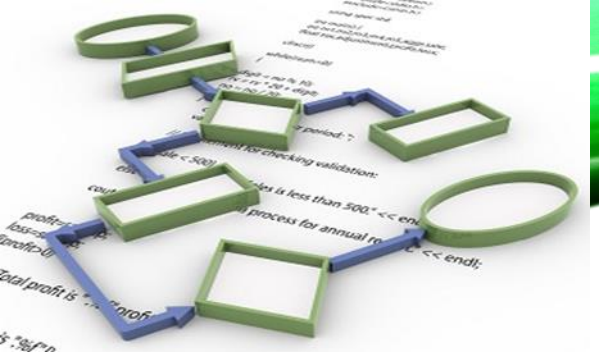
midwest architecture community collaboration

Components of the Model

- Knowledge Network (graph) Ontologies or Bayesian Networks describe concepts
- Metadata Repository End disjointedness with a single big table for all assets
- Controlled Vocabulary Build consistency with definitions, formulas and lineage
- Discovery Bots Building and maintaining it is too much for humans
- Curating/Validating Tools Bots' automated inferences may need to be tweaked
- Machine Learning Anomalies and ambiguities need not repeatedly offend
- Supervision Tools ML inferences may need to corrected, rejected or reinforced

How this Approach helps Enterprises
http://understandingcontext.com/2016/08/sustainable-software-for-the-enterprise/

IKE SEMANTIC SUITE



midwest architecture community collaboration

How to Process Meaning

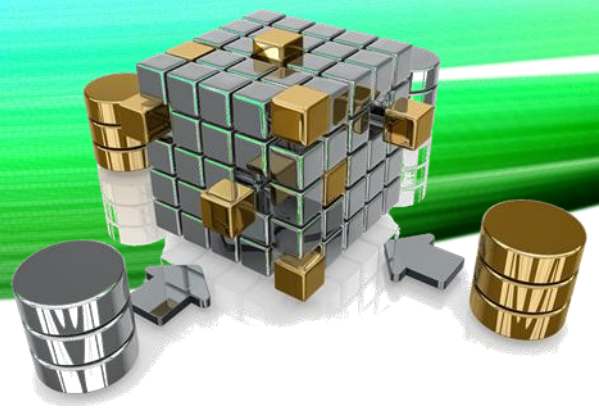
Tool	Use Case
• Heuristics	• any practical method to problem solving, learning, or discovery not optimal or but sufficient to reach a goal
• Inference Engine	
○ Forward chaining	○ When rules are complex but branching is known a-priori
○ Backward chaining	○ When path is not known but rules are known or discoverable
• Neural Network	• Low-dimensional problems with dirty or inconsistent data
• Bayesian Network	• For graph-type information spaces with many nodes
• Hidden Markov Model	• For problems where evolving facts could change outcome
• Genetic Algorithm	• For complex problems where ambiguity impacts process

Rules and Processes
http://understandingcontext.com/2014/02/workflow-and-business-rules-rings-of-power/

IKE SEMANTIC SUITE

Bayesian Inference
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4158865/

Bayesian Networks
https://towardsdatascience.com/introduction-to-bayesian-networks-81031eeed94e



midwest architecture community collaboration

What is Insight

1. the power or act of seeing into a situation : penetration
2. the act or result of apprehending the inner nature of things or of seeing intuitively

Merriam-Webster

Not to be confused with visualization:

1. the representation of an object, situation, or set of information as a chart or other image
2. the formation of a mental image of something

Given a set of inputs…

- Visualization is a graphical representation of some fact or set of facts
  - ➢ It must be analyzed by a human with the gifts of sight and intelligence
- Insight is the "meaningful" result of application of intelligence (human or machine)

Visualization
http://understandingcontext.com/2014/12/visual-knowledge-dimensions/

IKE SEMANTIC SUITE

Insight
http://understandingcontext.com/2018/07/anatomy-of-insight-prediction-and-qualitative-bi/

midwest architecture community collaboration

Quantitative → Measurable

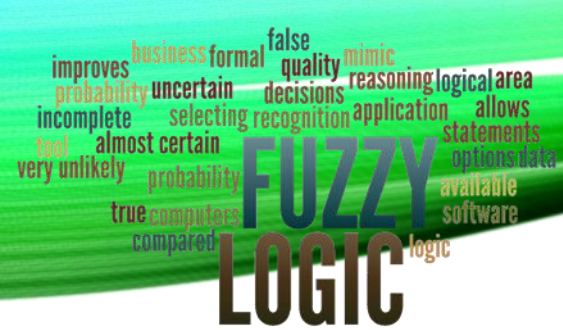
What happened (output/narrow outcome), when and where?

Examples of What, When and Where

- Weather:
  - Regional daily temperature was 5% higher than the 50 year average
- Sales:
  - US Sales increased by 3.6% compared to 4.2% increase in the prior year
- Election:
  - 12% more women voted for this party in this race than in the last election
- Process:
  - Plant 3 efficiency improved as shown by 31% increased throughput

IKE SEMANTIC SUITE

Quantitative Analysis (Chemistry and Business)
https://www.dictionary.com/browse/quantitative-analysis



midwest architecture community collaboration

Qualitative → Fuzzy

Why, what broader outcomes and what next?

Examples of Why

- Weather:
  - Nighttime cloud cover, ozone depletion, soil heat retention and warmer convection raised temperatures
- Sales:
  - Competing products with new features and market uncertainty impaired US sales growth in the quarter
- Election:
  - More women candidates on the ballot, and high profile appearances from female advocates raised turnout
- Process:
  - Lean manufacturing improvements reduced WIP at three critical points in the Plant 3 assembly line

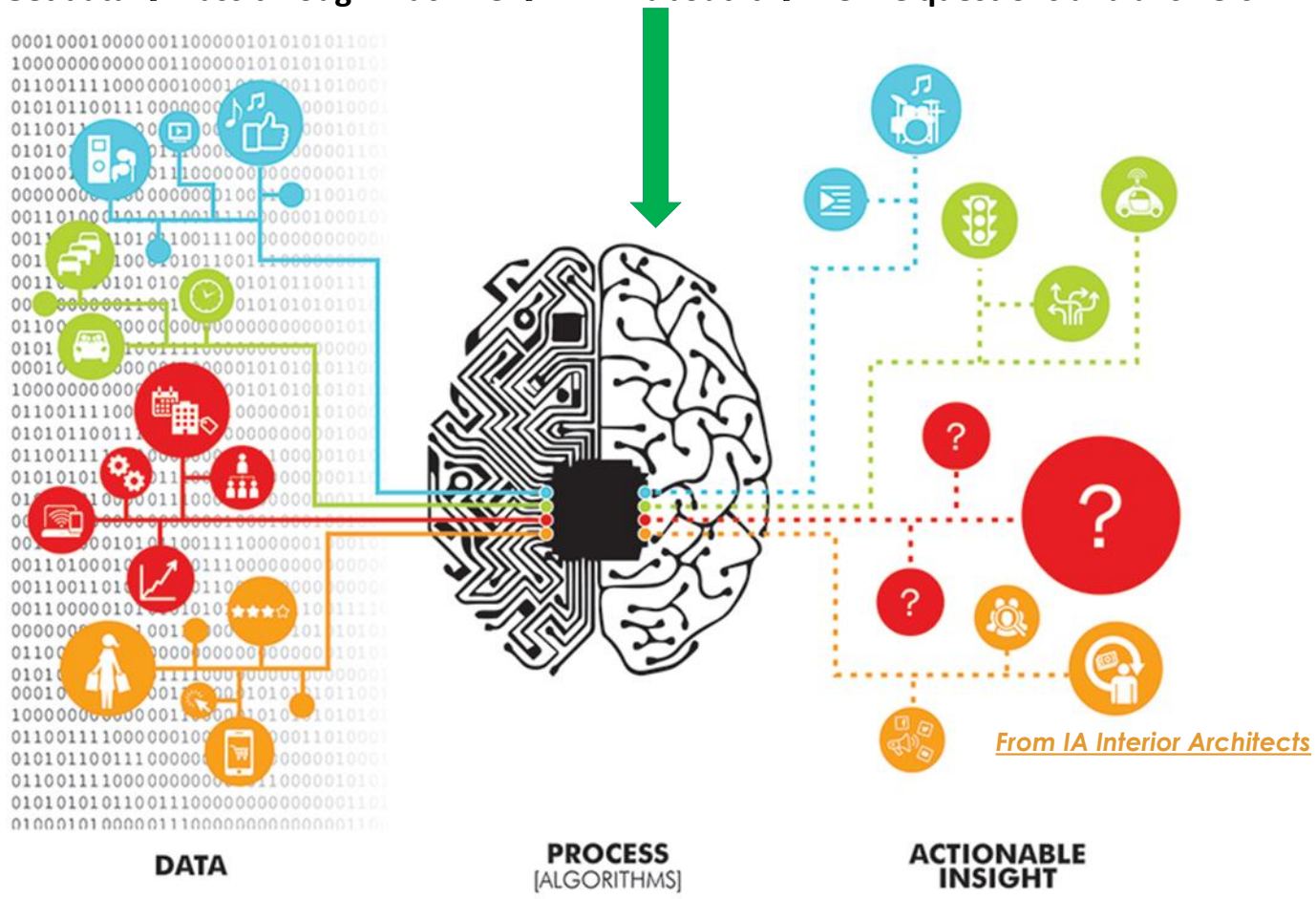
IKE SEMANTIC SUITE

Qualitative Analysis (Chemistry only)
https://www.dictionary.com/browse/qualitative-analysis

midwest architecture community collaboration

State of the Art in Insight

Get data → Pass through machine → Think about it → Derive questions and answers



From IA Interior Architects

Today's Challenges

- Incorrect assumptions (model)
- Inconsistent data quality
- Inconsistent skills
- Inconsistent results

To Improve Consistency

- Improve assumptions (model)
- Reduce disjointedness
- Automate more
- Use Brain-like methods

IKE SEMANTIC SUITE

Limitations of Current Approaches
http://www.dataversity.net/limitations-predictive-analytics-lessons-data-scientists/

midwest architecture community collaboration

Semantic Model Components

TIME and SPACE

- Where and/or when are critical components of most knowledge domains
- The data often tells you when and where things occurred, but qualitative insights are possible even in explicit data:
  - What prior changes or outcomes may have impacted current outcomes?
  - What changes in reporting may have caused greater variance from prior outcomes?

Taxonomy, Meronomy, Mereology, Metonymy,

Identity, Semiotics, Causality

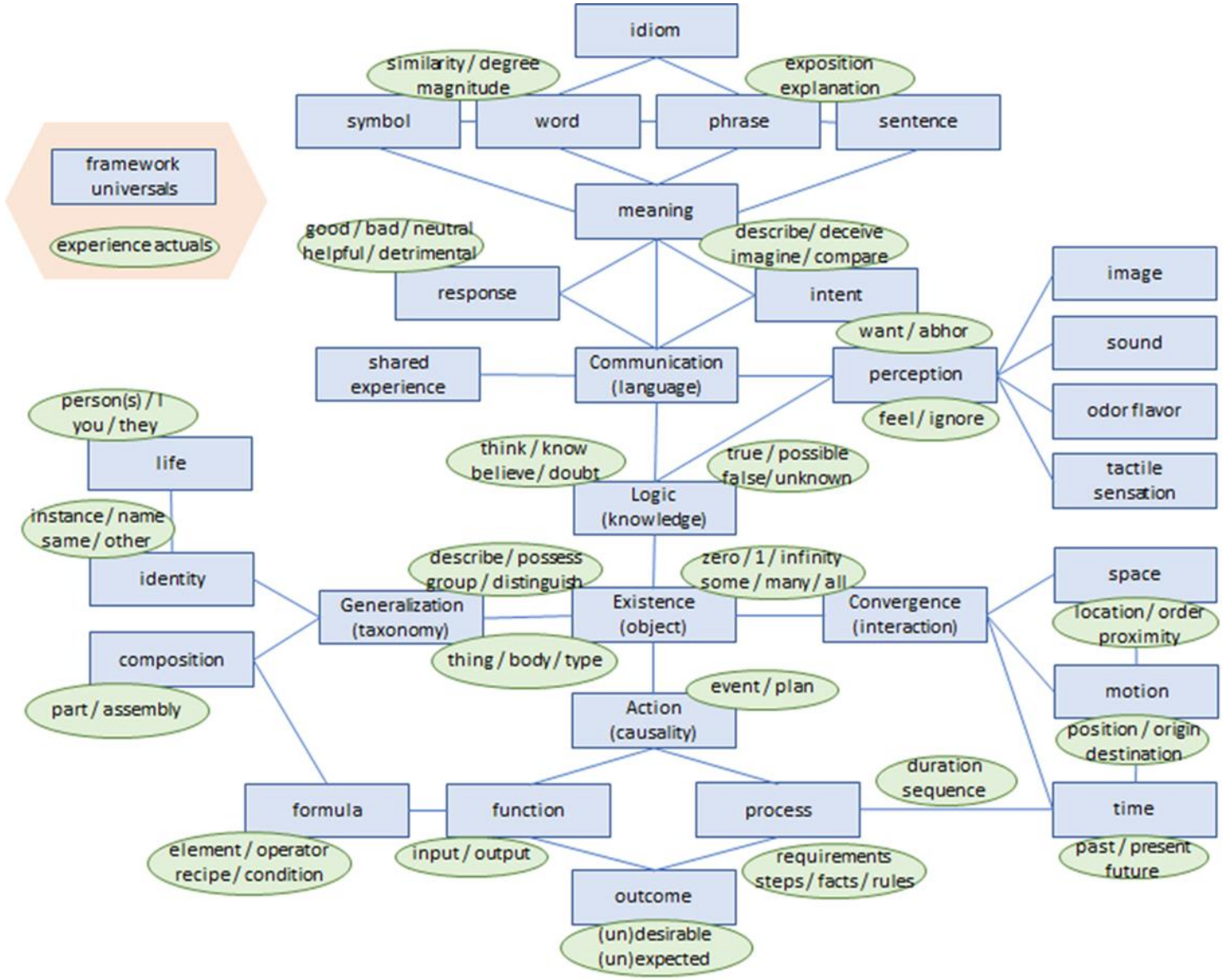
- The more accurate and complete the model, the better the outcomes

midwest architecture community collaboration

Meaning in the Model

What data do you need to answer why and use that as a basis for prediction and prescription?

A semantic model with a robust set of primitives that can be used to classify data well enough to derive "intent" as the human brain does.



Semantic Primitives as core of a Semantic Model
http://understandingcontext.com/2018/05/what-it-means-to-be-and-why-ai-needs-to-know/

IKE SEMANTIC SUITE

midwest architecture community collaboration

Inferring Cause vs. Correlation

- **The following roles can be played by a word that implies a causal association:**

  Agent, Instrument, Object, Action ***(Standard roles in traditional semantics)***

action	event	origin	proxy
agent	function	participant	reaction
behavior	gauge	path	recipient
byproduct	instrument	phenomenon	response
catalyst	material	precursor	result
cause	medium	prerequisite	source
conduit	method	process	substitute
consequence	object	producer	terminator
consumer	obstacle	product	transition
destructor	opponent	proponent	tributary
			vehicle

Note: The colored words may be more temporal, spatial, constructive, descriptive or other type of association and used for analyses other than inferring causality

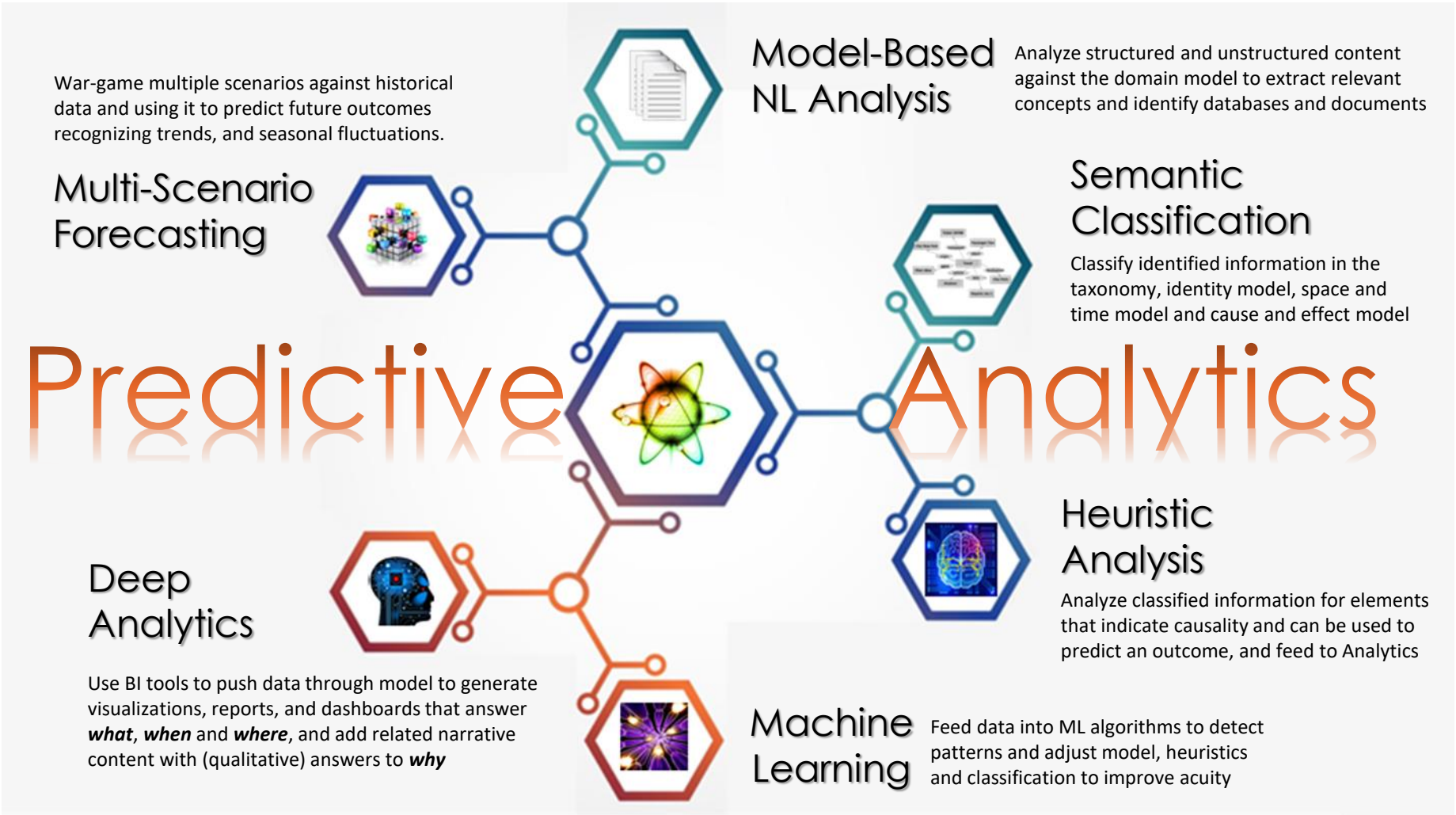
Causality
http://understandingcontext.com/2014/04/causal-chains/

IKE SEMANTIC SUITE

midwest architecture community collaboration

Predictive Analytics Cycle

- *Predictive analytics* does not prophesy the future, but seeks patterns in data sets to predict future outcomes and trends in terms of what, where, when and how much.
- Adding *qualitative insight* can describe why and what could change the likelihood



Interesting Forecasting Infographic
https://www.educba.com/predictive-analysis-vs-forecasting/

IKE SEMANTIC SUITE

midwest architecture community collaboration

Prescriptive Analytics

Current Standard Approach

Rearchitected Focus on Citizen

Prerequisite: Know the source data models, lineage, KPIs and formulas *(Data Governance)*

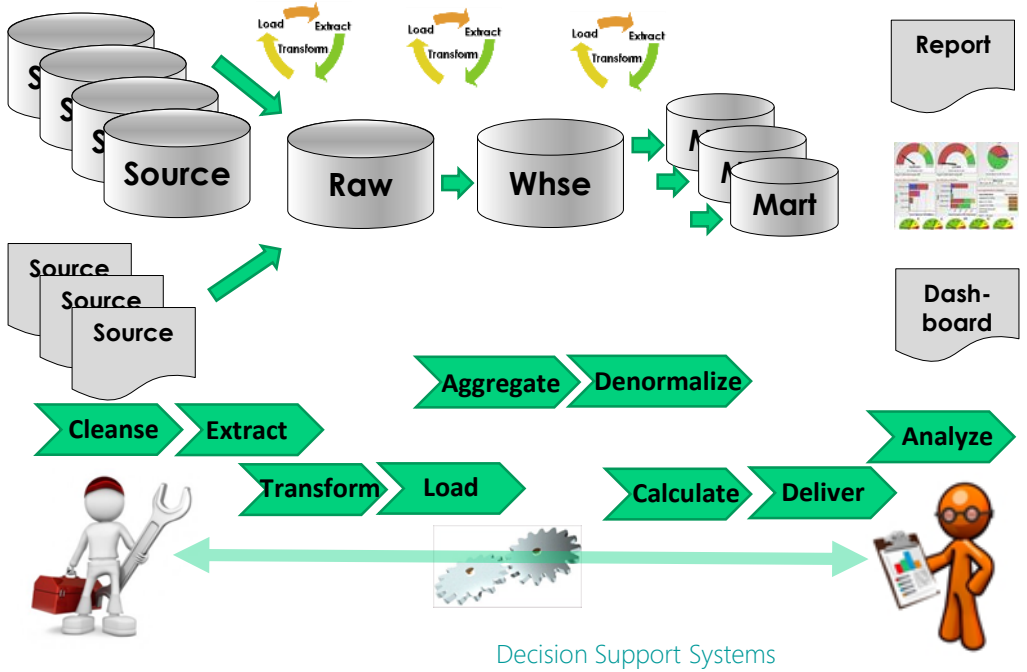
Prerequisite: Manage Data Lineage carefully

Identify, model and replicate data to non-transactional systems to improve both OLTP and OLAP performance

Denormalize and store data for optimal reporting, dashboards, analytics and visualization

Data Preparation

Data Presentation



Decision Support Systems
http://understandingcontext.com/2015/01/whats-in-a-decision/

Prerequisite: Classify data in an enterprise level semantic model of key domains and concepts

Wrangling, more than data prep, includes searching internal and external databases and documents

AI/NLP Bots look for meaning in every asset and classify based on model

Instead of constantly fighting with data, citizens curate meaning

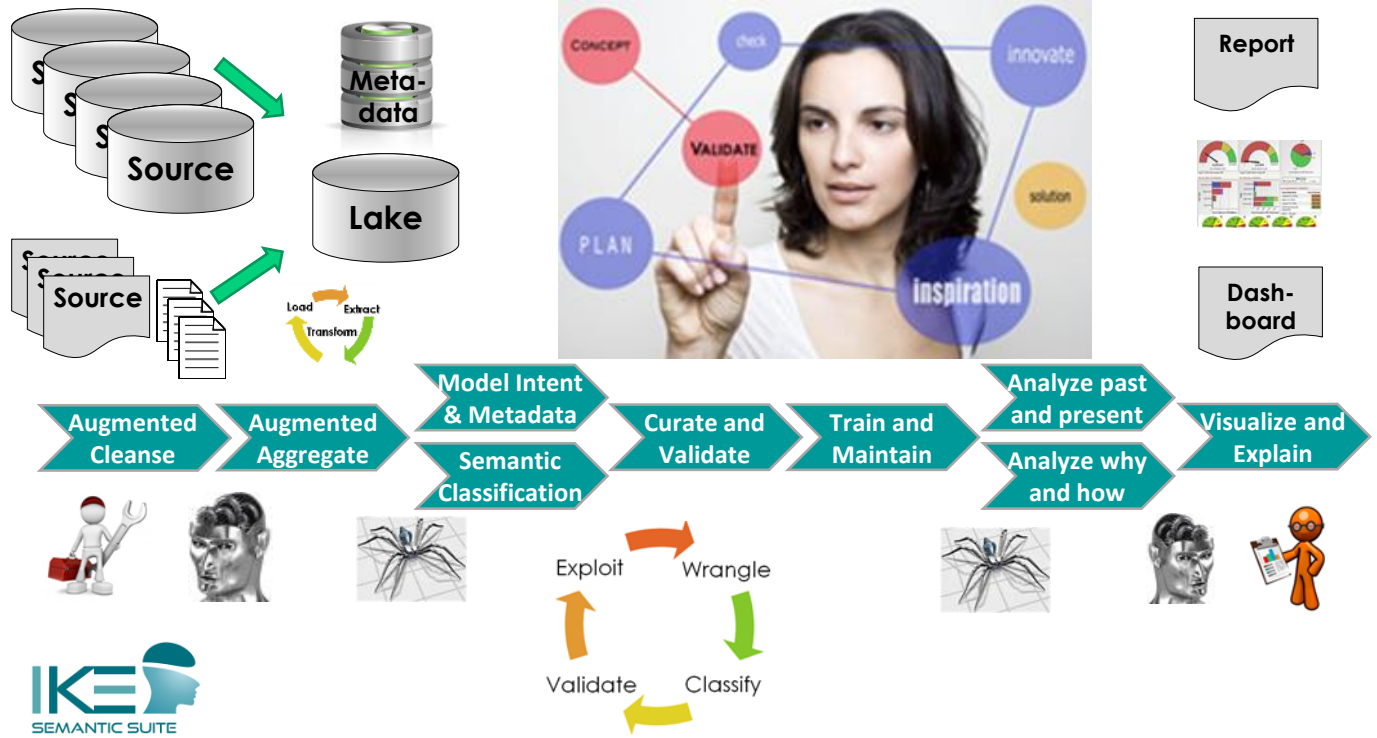
Deep understanding of the meaning lets users explore deeper and broader

Wrangle Data

Classify Data

Validate Info

Exploit Info



IKE SEMANTIC SUITE

midwest architecture community collaboration

Wrangling and AI

Exploit

Wrangle

Validate

Classify

With the help of AI decision support, can all these tasks be moved to the “Citizen”

Data wrangling, sometimes referred to as **data** munging, is the process of transforming and mapping **data** from one "raw" **data** form into another format with the intent of making it more appropriate and valuable for a variety of downstream purposes such as analytics. Using the EIM graph, IKE can auto-map data from new sources to target.

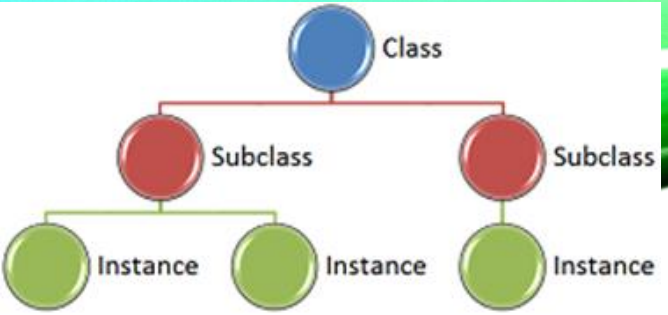
Wranglers are not ETL technicians: Wrangling tools are for the people who best know the data (Business Analysts, LOB SMEs) to explore and prepare that data. The data is more diverse, some poorly- or un-structured. And use cases are exploratory, explanatory, predictive or prescriptive.

Augmented Cleansing means more than SQL string functions, which require you to know the nature of data issues and know exactly what to do to fix them: AI tools identify, assess and characterize data inconsistencies, then propose solutions, often heuristics, to resolve them.

https://www.techopedia.com/definition/30968/citizen-developer
https://www.gartner.com/it-glossary/citizen-developer/

IKE SEMANTIC SUITE

Citizen Data Scientist
https://www.forbes.com/sites/stevebanker/2018/01/19/the-citizen-data-scientist/



midwest architecture community collaboration

Classification

What is the Process?

- Semantically or Numerically grouping similar items by characteristics, (see also "cluster analysis"). The results can be critical to model building and validation, and can find hidden or historically overlooked groups.

Who are the Agents?

- Discovery tools/Bots can do it → SMEs validate it, often with help from AI

What is the Object?

- Classify databases, table names and column names by concept into metadata
- Classify documents by title, tag, author and content by concept into metadata
- Associate each asset in metadata with a node in the model, and adapt model

Unless IT has deep conceptual understanding–this belongs to citizen

http://understandingcontext.com/2014/10/knowledge-in-non-neural-models/

http://understandingcontext.com/2014/10/neural-conceptual-dependency/

IKE SEMANTIC SUITE

http://understandingcontext.com/2014/09/gnostic-learning-model/

Conclusion:

The “citizens” own and manage the model

We architect the suite of tools

To deliver them quickly they need good tools

Thank you for attending

midwest
architecture community
collaboration

IKE
SEMANTIC SUITE