Machines Search, Users Find

Organization Beyond The Department Level

Information access is instrumental to solving business problems

Is Enterprise Search Worth It?
Instantly access answers to prospective client questions and current client needs.

Onboard new hires quickly and get them up to speed on products and services.
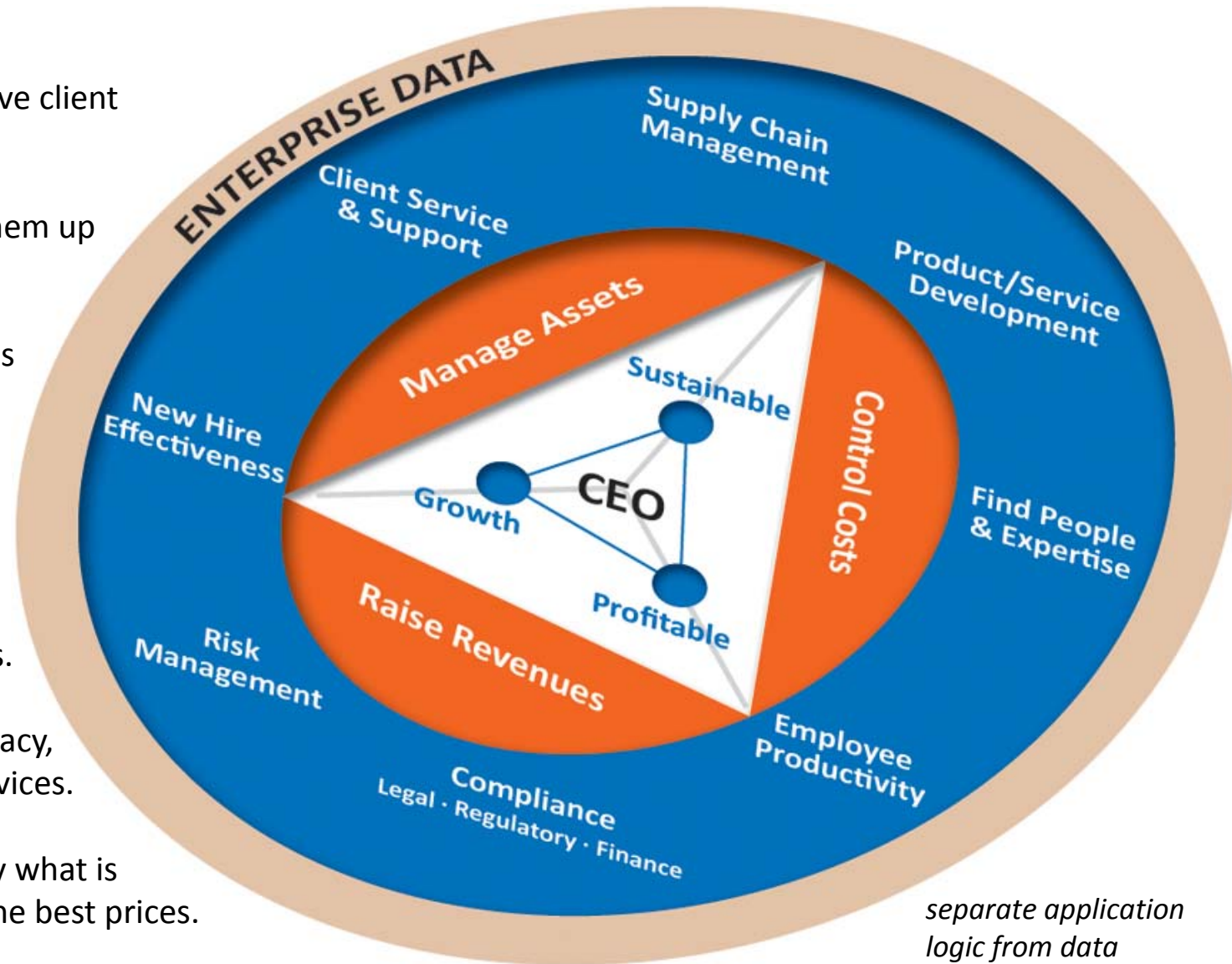
Comply with new civil procedure rules that mandate electronic discovery.

Find experts within the company by analyzing employee work product.

Manage risks and fraud by detecting inappropriate documents and actions.

Develop new services by studying legacy, current and competitor products/services.

Maintain equipment and acquire only what is critical for continuing operations at the best prices.

ENTERPRISE DATA
Client Service & Support
Supply Chain Management
Product/Service Development
New Hire Effectiveness
Manage Assets
Sustainable
Control Costs
Growth
CEO
Find People & Expertise
Risk Management
Raise Revenues
Profitable
Employee Productivity
Compliance
Legal · Regulatory · Finance

separate application logic from data

How Does Search Work?

Simple example: I am searching for all "green" hay in a haystack

Step One: Recall
Give me all the green hay in the haystack

Challenge
How much green hay did I leave behind in the haystack?

Are my results missing relevant documents?
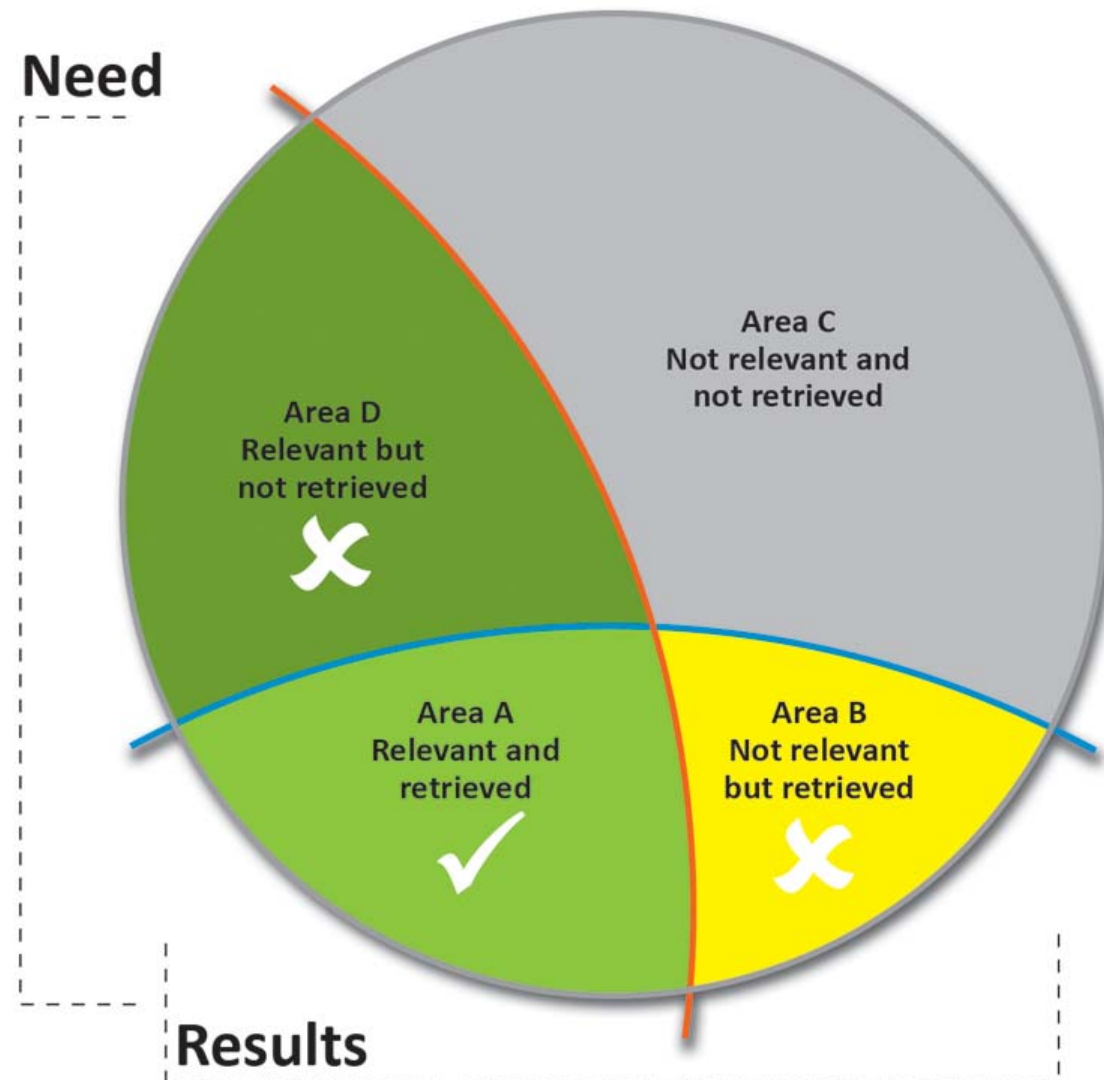
Recall = A/(A+D)

Step Two: Precision
I wanted green hay only, not yellow and green hay

Challenge
When asking for green hay, did I get results that did not match my query?

Am I retrieving documents that are not relevant?

Precision = A/(A+B)

Need

Area C
Not relevant and not retrieved

Area D
Relevant but not retrieved
✗

Area A
Relevant and retrieved
✓

Area B
Not relevant but retrieved
✗

Results

Finding the Most Important or Authoritative Results

Challenge: Limit recall to highly relevant collections and provide iterative navigation

Link Analysis?
The importance of a page is judged by the number of pages linking to it as well as their importance

Challenge
Internal corporate documents do not rely on hyperlinks or external links to internal documents

Successful "consumer" search applications (Page Rank) do not translate well for corporate needs
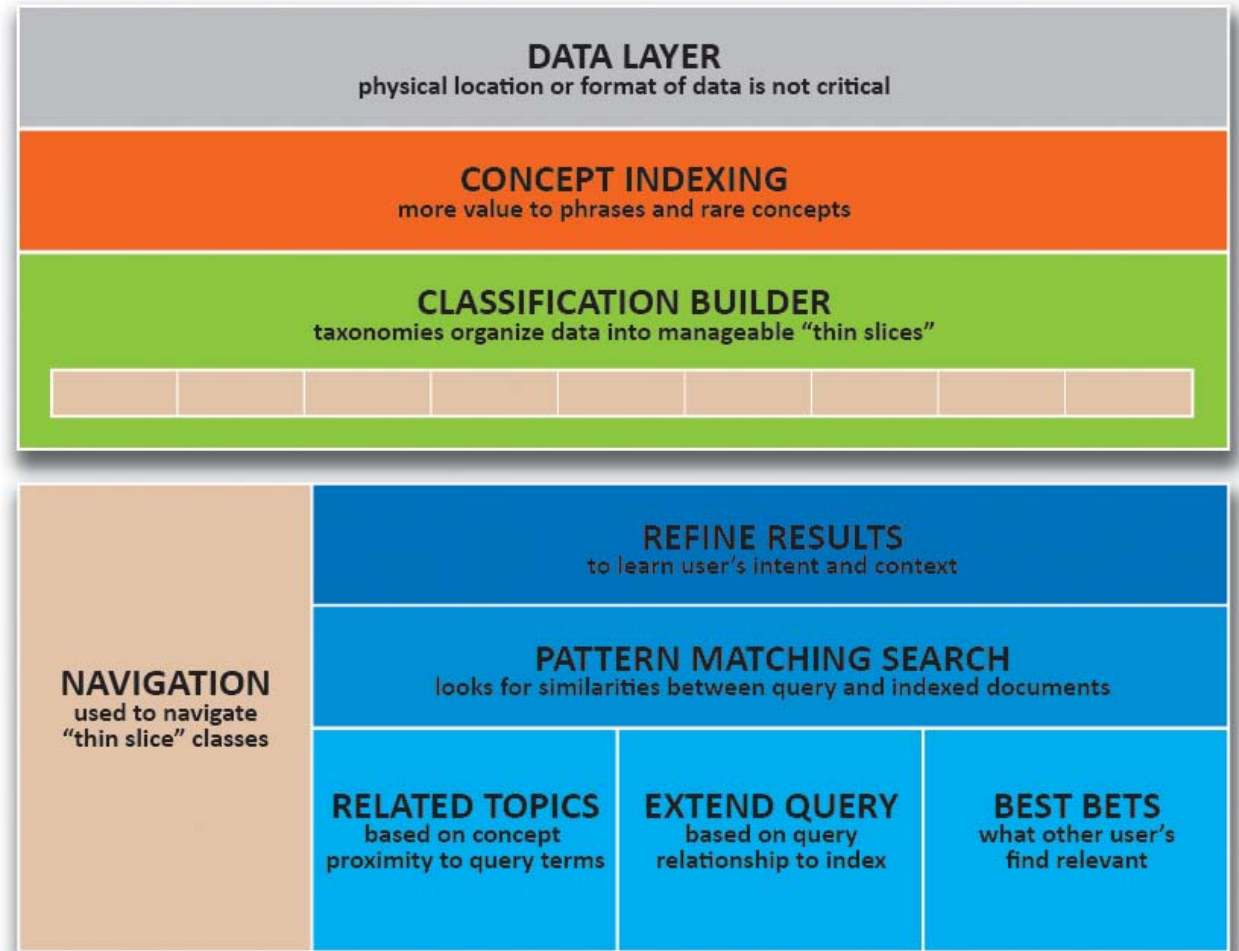
Content Analysis?
Roughly 95% of the text in all corporate documents is composed of the same 10,000 words

Challenge
A huge number of documents will contain the words in a user's search phrase

Density of search terms can assist recall and precision, but do little for ranking search results

Needed: A Content Topology

DATA LAYER
physical location or format of data is not critical

CONCEPT INDEXING
more value to phrases and rare concepts

CLASSIFICATION BUILDER
taxonomies organize data into manageable "thin slices"

NAVIGATION
used to navigate "thin slice" classes

REFINE RESULTS
to learn user's intent and context

PATTERN MATCHING SEARCH
looks for similarities between query and indexed documents

RELATED TOPICS
based on concept proximity to query terms

EXTEND QUERY
based on query relationship to index

BEST BETS
what other user's find relevant

Does the Right Hand Know What the Left Hand is Doing?

Challenge: Quantifying the importance of sharing information across the enterprise

2007 Information Management Trends
Large company middle managers spend up to 2 hours/day searching for information and 50% of what's retrieved is useless

Other Findings:
60% cannot find information they know exists somewhere in the company

42% accidently use the wrong information once a week

40% say other parts of the company are not willing to share information

53% say less than half the information they receive is useless

accenture >
January 2007 Information Management Study

Users are willing to:
- upload documents to a <u>server</u> for sharing with other employees in the organization
- <u>validate</u> "tags" that describe a document at the time of uploading
- mark documents that are out of date, misfiled or contain an error
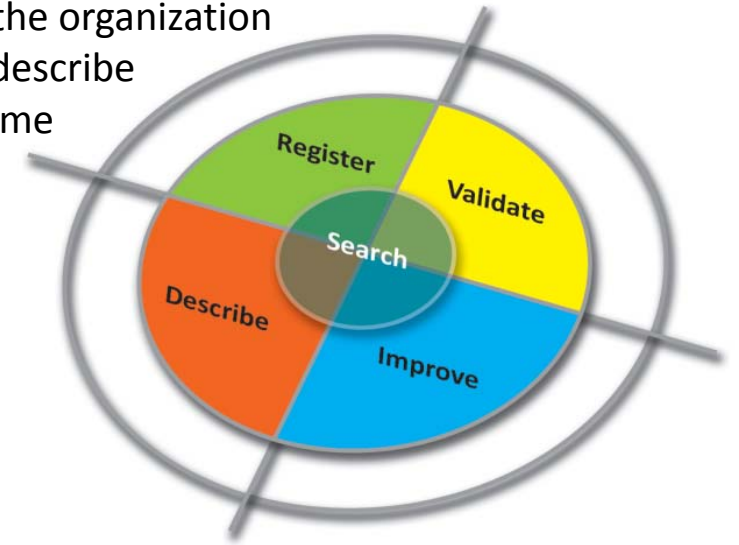
But users need:
- a Google-like search experience that does not require boolean skills or an advanced search screen
- way to improve search query with a single click from the search results page
- way to dog-ear specific documents for later recall and sharing with immediate team members

Information Management Framework

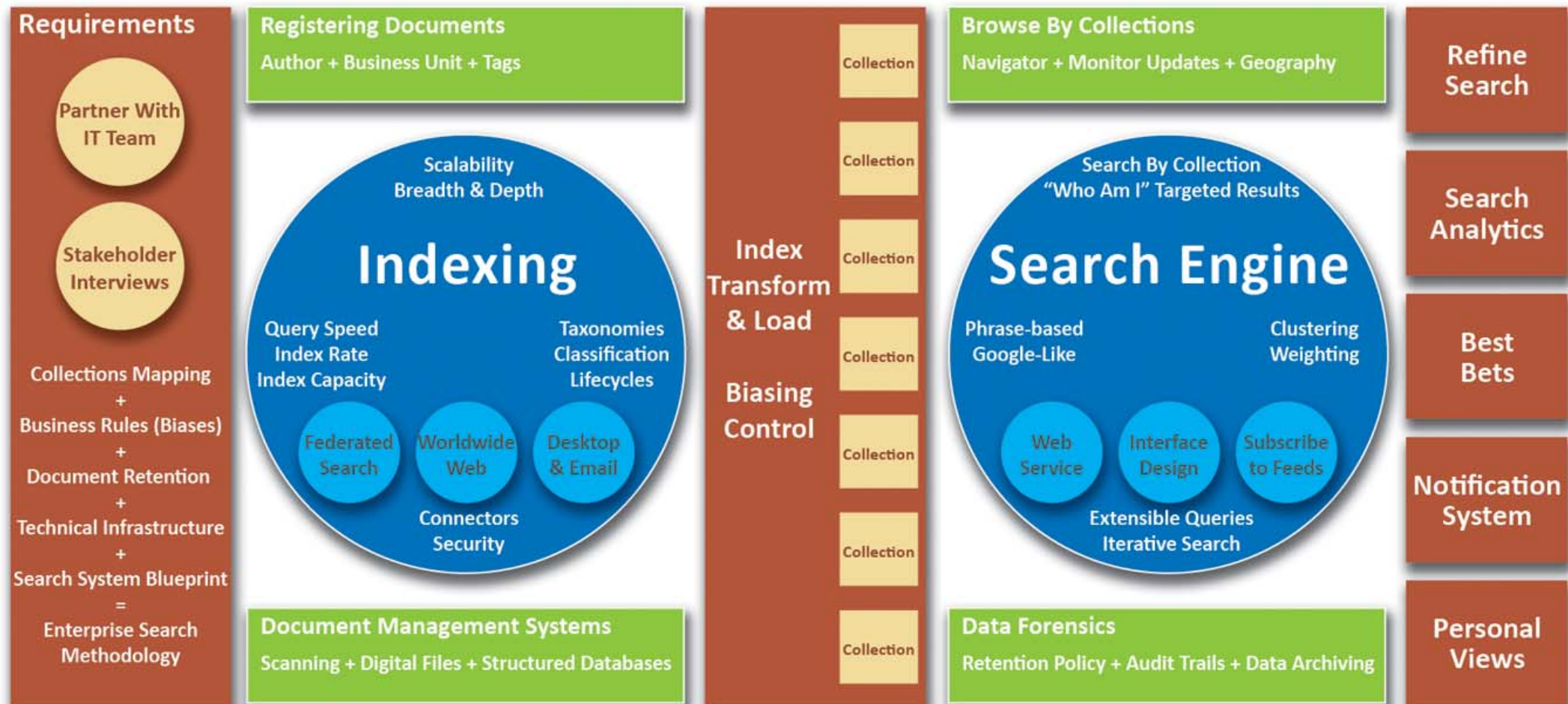Challenge: Build a flexible model that adapts to specific organizational needs
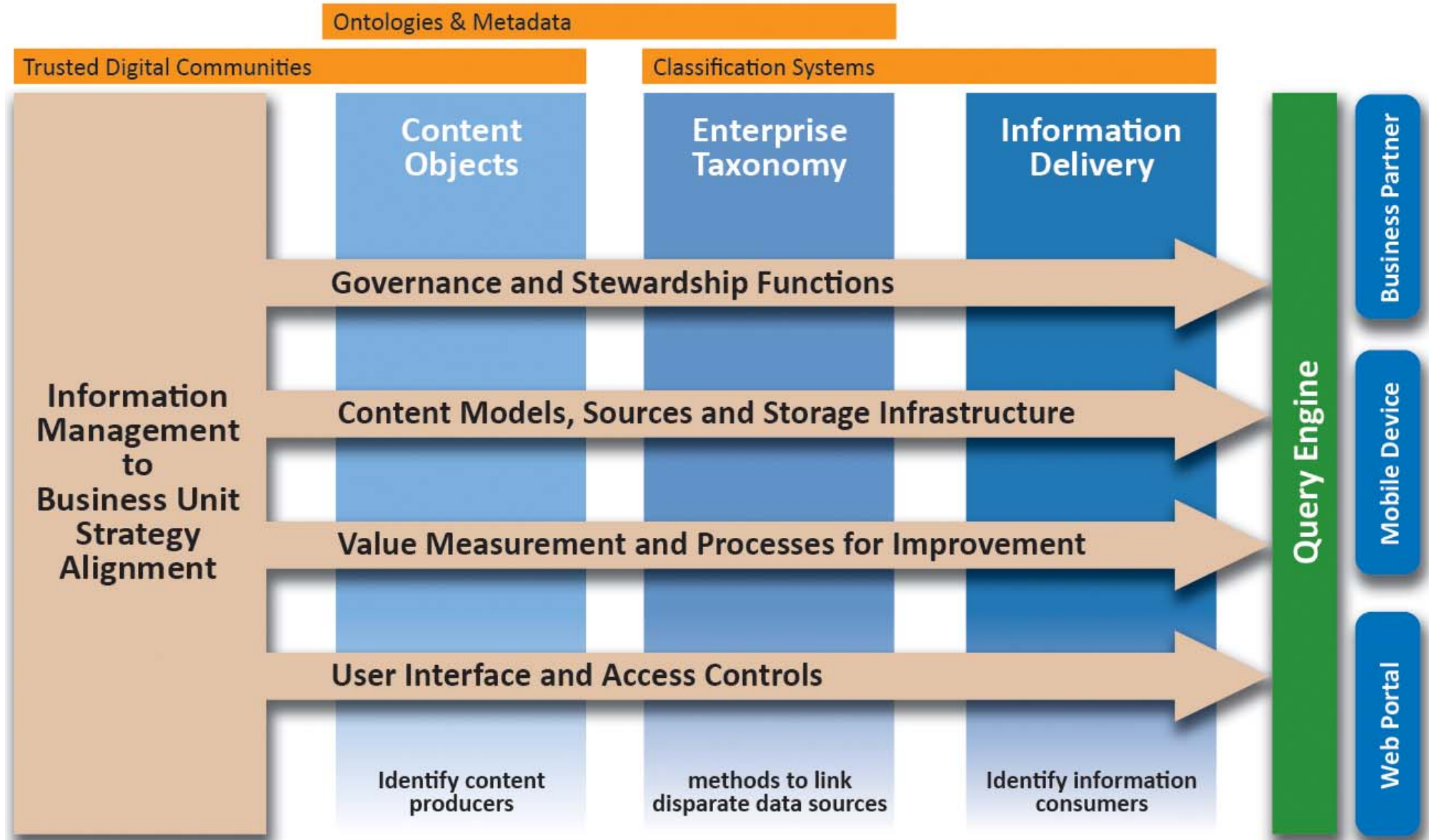
Information Management Architecture

Challenge: The importance of a requirements phase and understanding stakeholder needs

Perfect Search is a Moving Target

Challenge: Improving search capabilities over time to meet changing business needs

Configuration Improvements:

Narrow Search Results	**Stemming Algorithms**	**Stop Word Dictionary**	**Highlight Search Terms**	**Synonym Lists**	**Weighting Keywords**
Users narrow the index to a specific collection	Looks for multiple variations of the same word	Removes common words that can bias search results	Very useful when using the refine feature	A search for "car" will find sedan, auto, automobile	Frequently read content can be moved to top

Understanding User Experience:

Analyze past search behavior » When do user's initiate search? » Do they exit the site from the search results page?
Analyze search logs » What are the most common searches? » Evaluate root causes for "null" search results
"Did You Find What You Are Looking for?" feedback form on all search results pages » respond to all user feedback

Improving Content:

Ask users to validate/add descriptive tags when uploading » mark out-of-date content and documents misfiled in system
Pull rich metadata from "people" database for user, for business unit and based on physical location of uploaded document
Classify content into multiple areas of the enterprise system » set different thresholds for inclusion based on "thin slice"

A Consulting Firm That Listens and Delivers Sustainability

Challenge: Finding a firm that solves a problem, is realistic and <u>leaves</u> after the project.