Agenda:
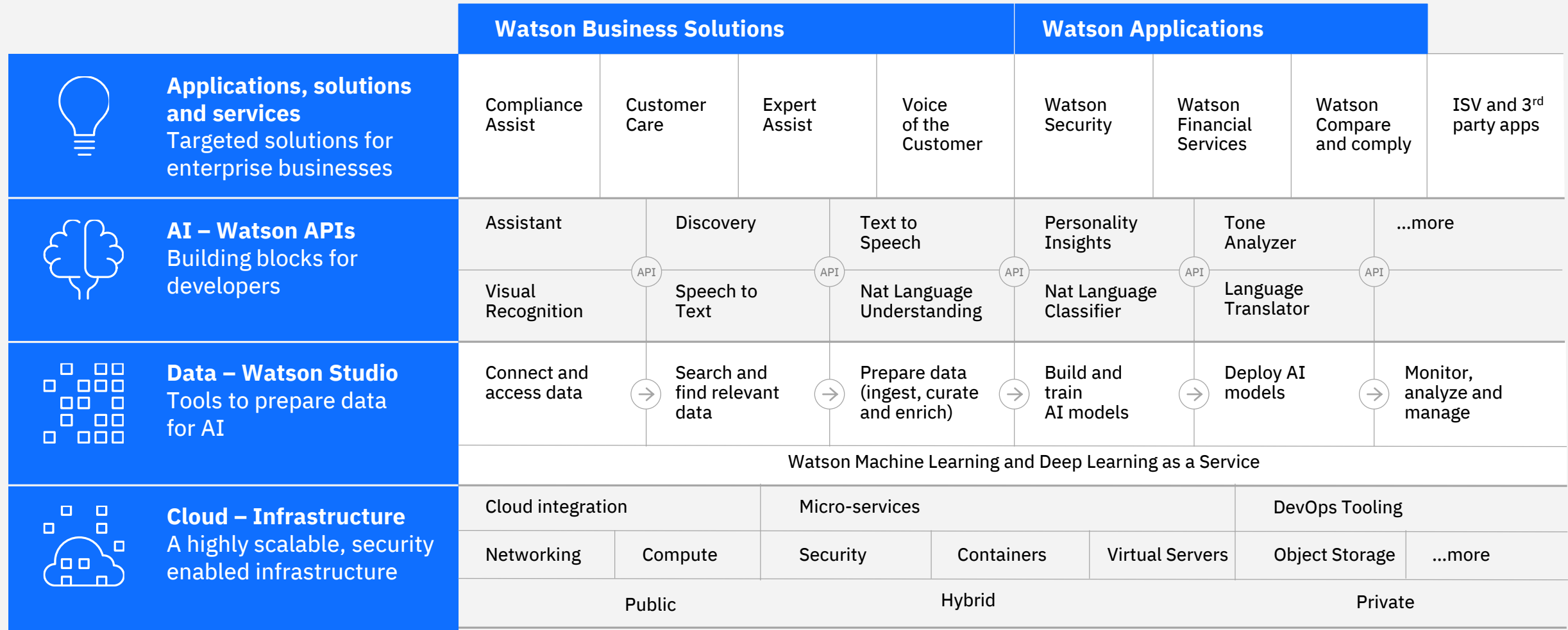
1. Overview of IBM Offerings in text analytics - Intro (5 min)
2. Service details: (60 min)
   a) Watson Discover
   b) Watson NLU
   c) Watson Tone Analyzer
   d) Watson NLC
   c) Watson Knowledge Studio
3. Demo - STB dataset (1 hr)
4. Hands on using STB dataset (1hr)
5. Q&A (15 min) & Wrap up

Watson is AI for smarter business

	Watson Business Solutions				Watson Applications			
Applications, solutions and services Targeted solutions for enterprise businesses	Compliance Assist	Customer Care	Expert Assist	Voice of the Customer	Watson Security	Watson Financial Services	Watson Compare and comply	ISV and 3rd party apps
AI – Watson APIs Building blocks for developers	Assistant Visual Recognition	Discovery Speech to Text	Text to Speech Nat Language Understanding		Personality Insights Nat Language Classifier	Tone Analyzer Language Translator	...more	
Data – Watson Studio Tools to prepare data for AI	Connect and access data	Search and find relevant data	Prepare data (ingest, curate and enrich)		Build and train AI models	Deploy AI models	Monitor, analyze and manage	
	Watson Machine Learning and Deep Learning as a Service							
Cloud – Infrastructure A highly scalable, security enabled infrastructure	Cloud integration		Micro-services			DevOps Tooling		
	Networking	Compute	Security	Containers	Virtual Servers	Object Storage	...more	
	Public		Hybrid			Private		

API API API API API

What are
Watson APIs?

– **A collection of cloud services for developing Enterprise AI solutions**
Protects your insights

– **Services can be trained using custom models**
Learn more with less data

– **Each service has a RESTful API**
Usable from virtually any language/platform

– **Provisioned at IBM Cloud**

– **http://ibm.com/watson**
Demos, documentation, API reference, sample apps, SDKs

– **https://developer.ibm.com/code/patterns/**
Comprehensive AI code patterns
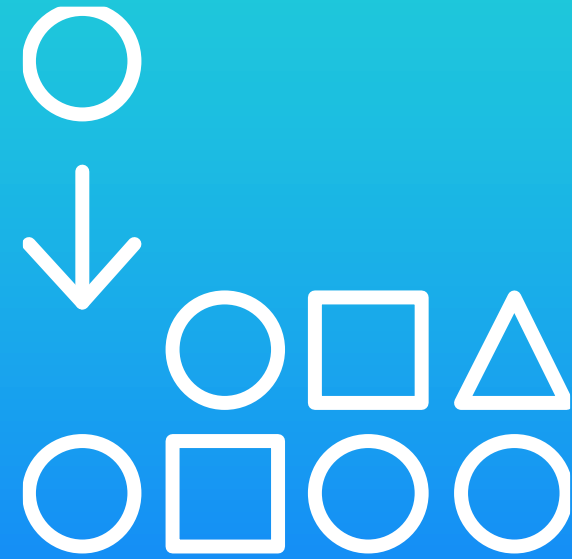
Four capabilities distinguish Watson from traditional programmed computing systems
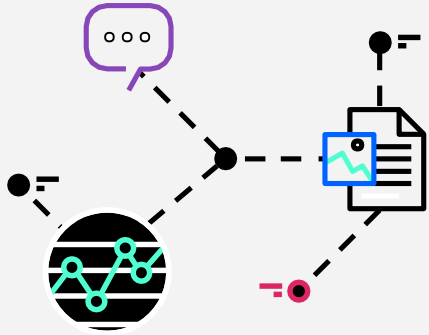
Understands
imagery, language and other unstructured data humans
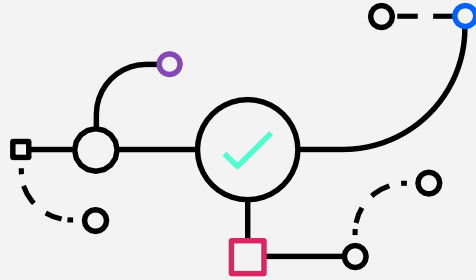
Reasons
forms hypotheses, infers and extracts ideas

Learns
With every interaction to sharpen expertise
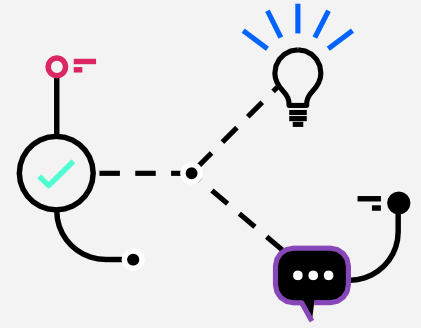
Interacts
With humans through seeing, talking and hearing

Language
Assistant

Language Translator

Personality Insights

Tone Analyzer

Natural Language Understanding

Speech
Speech to Text

Text to Speech

Vision
Visual Recognition

Data Insights
Discovery

Watson APIs

Watson Assistant

Language Translator

Speech to Text

Tone Analyzer

Natural Language Understanding

Natural Language Classifier
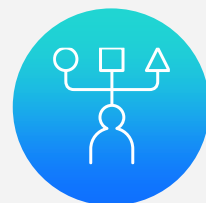
Personality Insights
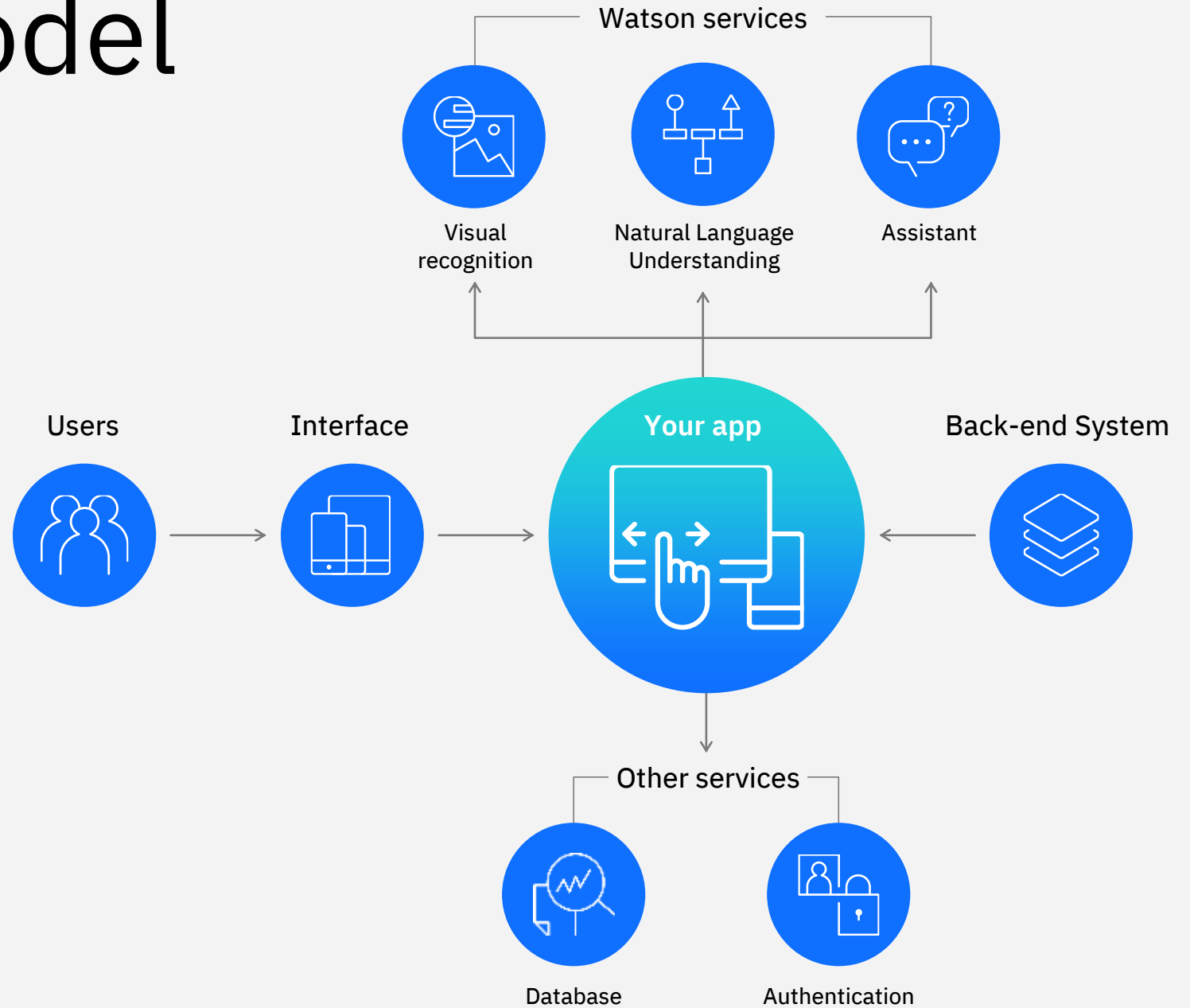
Visual Recognition

Discovery

Text to Speech

Execution model

A main application

– Hosted on IBM Cloud

– Running on servers that you manage

Watson services

Visual recognition

Natural Language Understanding

Assistant

Users

Interface

Your app

Back-end System

Other services

Database

Authentication

Watson Solutions combine Watson Services, Watson APIs, and pre-built assets to quickly prove value and deliver foundational applications
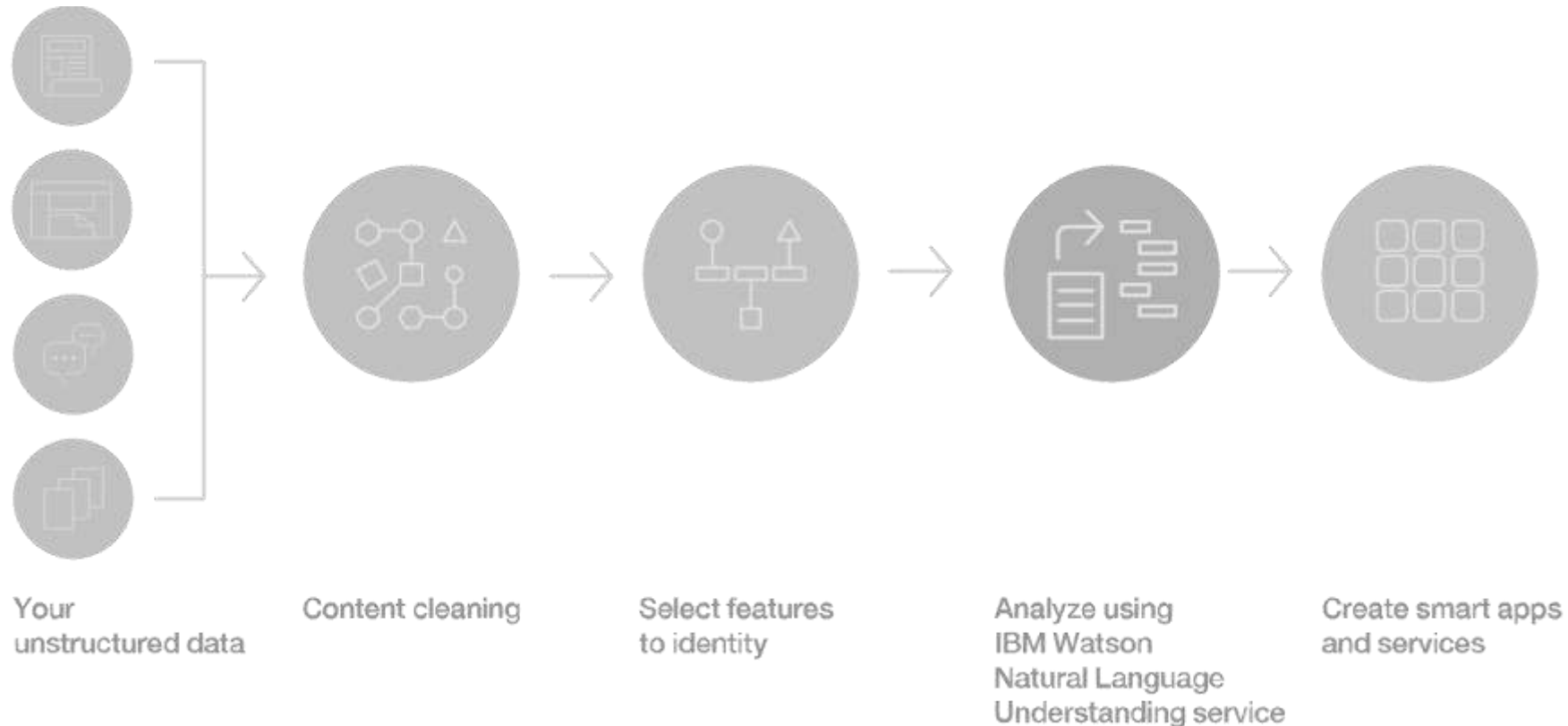
Watson
Services

Expert Services,
architecture, and
pre-built assets

Customer
data

Watson Solution
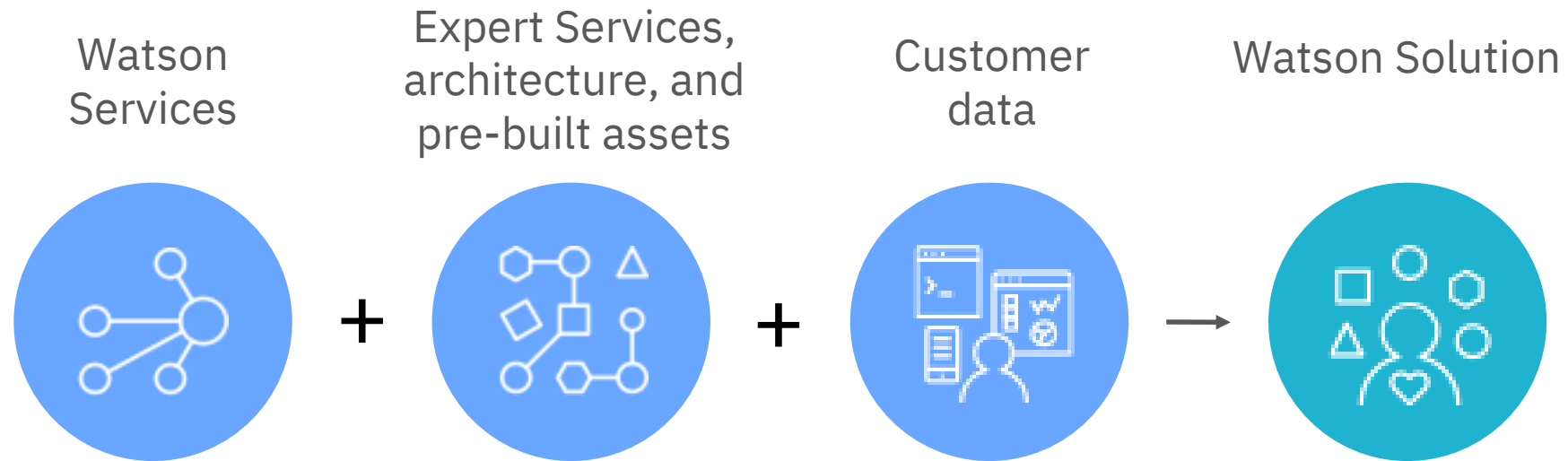
IBM Watson NLU

Natural language processing for advanced text analysis

A sophisticated suite of natural language processing capabilities to analyze text and extract meta-data from content such as concepts, entities, keywords, categories, sentiment, emotion, relations, semantic roles, with options for customization to specific industries and domains.

How does it work?

How it works

IBM Watson Natural Language Understanding makes value-driven decisions easy by giving you the full story behind all of your data.

Your
unstructured data

Content cleaning

Select features
to identity

Analyze using
IBM Watson
Natural Language
Understanding service

Create smart apps
and services

Highlights

Comprehensive NLP stack: A rich set of enrichments to extract metadata from unstructured text for deep analysis in various use cases

Broad Language Coverage: Depending on the feature, NLU can understand text in thirteen different languages

Customization: Industry and domain-specific annotation possible with custom models developed using Watson Knowledge Studio resulting in higher accuracy broader uses

Improved API: API Signature is consistent with all Watson APIs enabling quicker development

Scalability: Cloud-based service with the infrastructure to support millions of NLU enrichments per month

What are the main features?

Features

Keywords: Determine the most important keywords in your content.

Concepts: Identify general concepts in your content.

Categories: Categorize your content into a hierarchical 5-level taxonomy.

Entities: Detect important people, places, geopolitical entities and other types of entities.

Sentiment: Determine whether your content conveys positive or negative sentiment.

Features (cont.)

Emotion: Detect emotions such as anger, disgust, fear, joy or sadness that are conveyed by your content.

Relations: Identify relationships between entities in your content.

Semantic roles: Identify the subjects of actions, and the objects that they act upon.

Metadata: Get author information, publication date and the title of your text or HTML content.

Custom models: Use IBM Watson Knowledge Studio to customize text extraction.

What are the use cases?

Applications

Content Recommendation: How can I recommend content that readers might like?

Advertising Optimization: Where and how should I place my ads to the right audience?

Audience Segmentation: How can I segment customers to tailor strategies for different audiences?

Voice of Customer Analysis: How can I address customer feedback and improve customer experience?

Data Mining: How can I search a corpus and discover insights in an efficient way?

Use Case

Content Recommendation

Recommend similar content to what a user is interested in viewing or experiencing
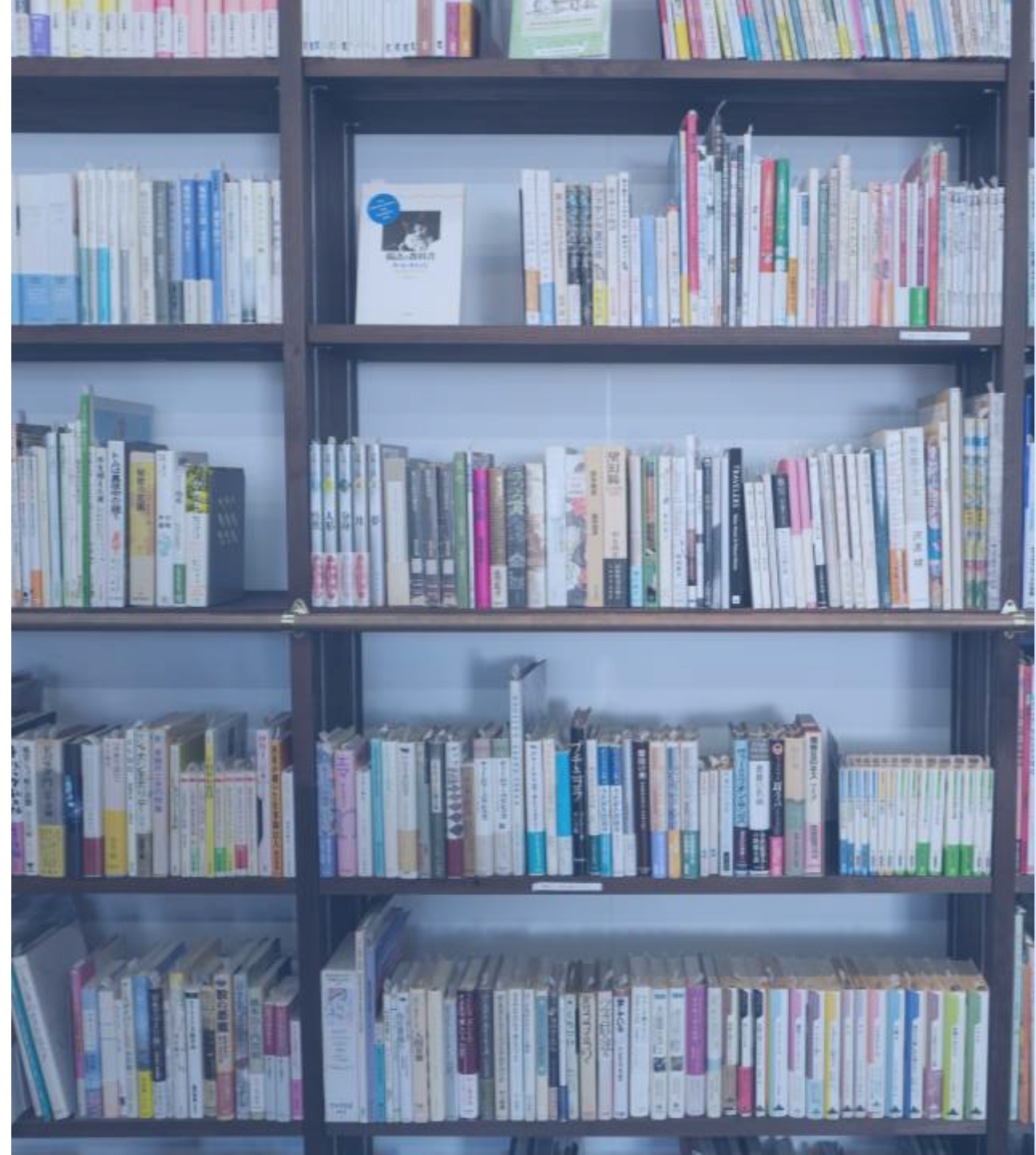
Use Case

Advertizing Optimization

Ensure proper placement of advertisements based on page content and viewer patterns

Use Case

Audience Segmentation

Identify key customer groups for market research purposes and to personalize campaigns towards different segments
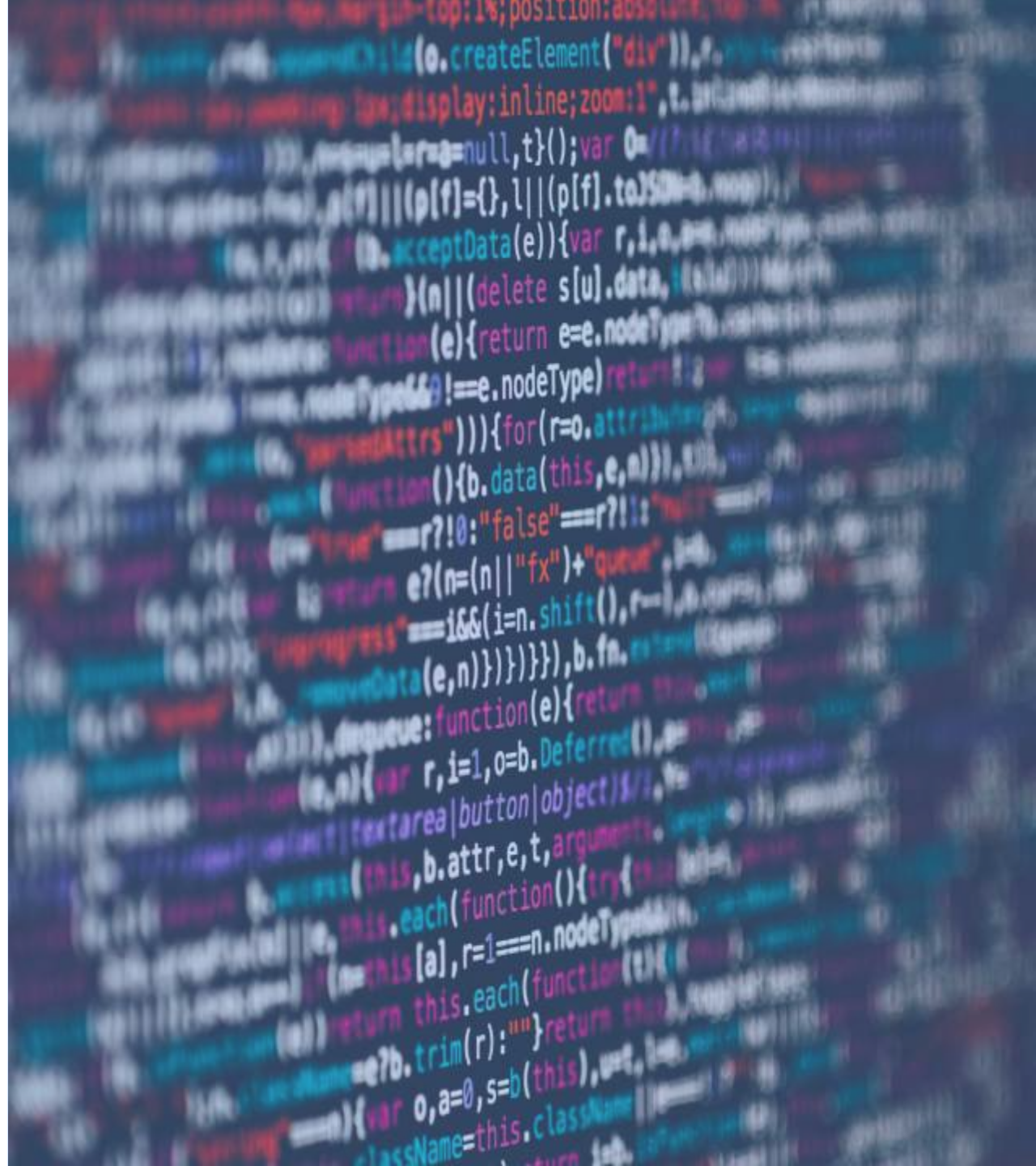
Use Case

Voice-of-Customer Analysis

Spot trends in customer feedback to
identify business opportunities,
address concerns, reduce churn, and
drive revenue

Use Case

Data Mining

Search through large content repositories and get answers that are relevant and quick

Demo

Watson Discovery Components

Customer Apps

Query API

Watson Discovery

UI Tooling

News Collection

Natural Language Search

Passage Retrieval

Relevancy Training

Anomaly Detection

Knowledge Centric Reasoning

Query, Training and Analytics

Concepts

Emotion

Semantic Roles

Metadata

Categories

Keywords

Sentiment

Entities

Relations

NLP Enrichment

Custom Relationships
Custom Entities

Data Crawler

Document Conversion/Normalization

Ingestion API

Data Ingestion

World News

File Systems

CMIS

Databases

Customer Data

MS Sharepoint (and Sharepoint Online)

box

Websites

•••

Watson Knowledge Studio

Watson Discovery :
From Data to Insights - Beyond Enterprise Search

WORD   PDF

{JSON}   <HTML>
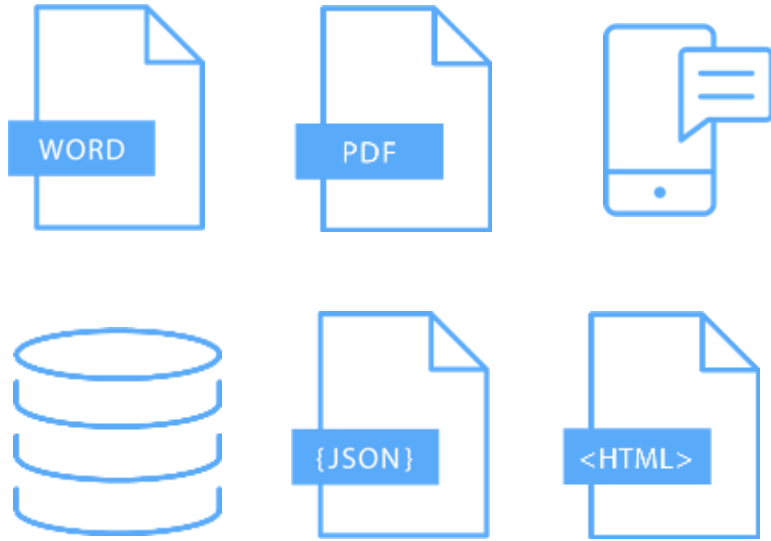
Documents and Data ⟶

"What phones are available that run Android and have good battery life under $400?"

"What plans include international roaming?"

"What are the most frequently mentioned problems with this phone model?"

"What category of questions are customers the most upset about?"

Smart Document Understanding

Teach Discovery how to interpret document structure and components

- Discover deeper insights by using extracted content (such as table header or abstract) in queries, to segment documents, ignore irrelevant parts and more

- Even identifies text in scanned business documents (OCR)

- Document formats supported: pdf, word, excel, powerpoint, html, png, jpeg, tiff and json

- Train Discovery to understand complex and domain specific tables

Rapidly Train

- Train AI model with point and click

- Label content in a small set of sample documents, including semantics like title, abstract, author or your own custom labels.

- Train on table semantics, header, footer, columns, more.

- Tests show 99% accuracy in as little as 200-300 labeled pages. Homogenous data sets in even less trained pages.

With labeled document sections you can :

- *Not index certain content*

- *Merge content*

- *Split content into separate documents*

- *Identify Q&A Pairs*

- *Perform targeted queries on specific content*
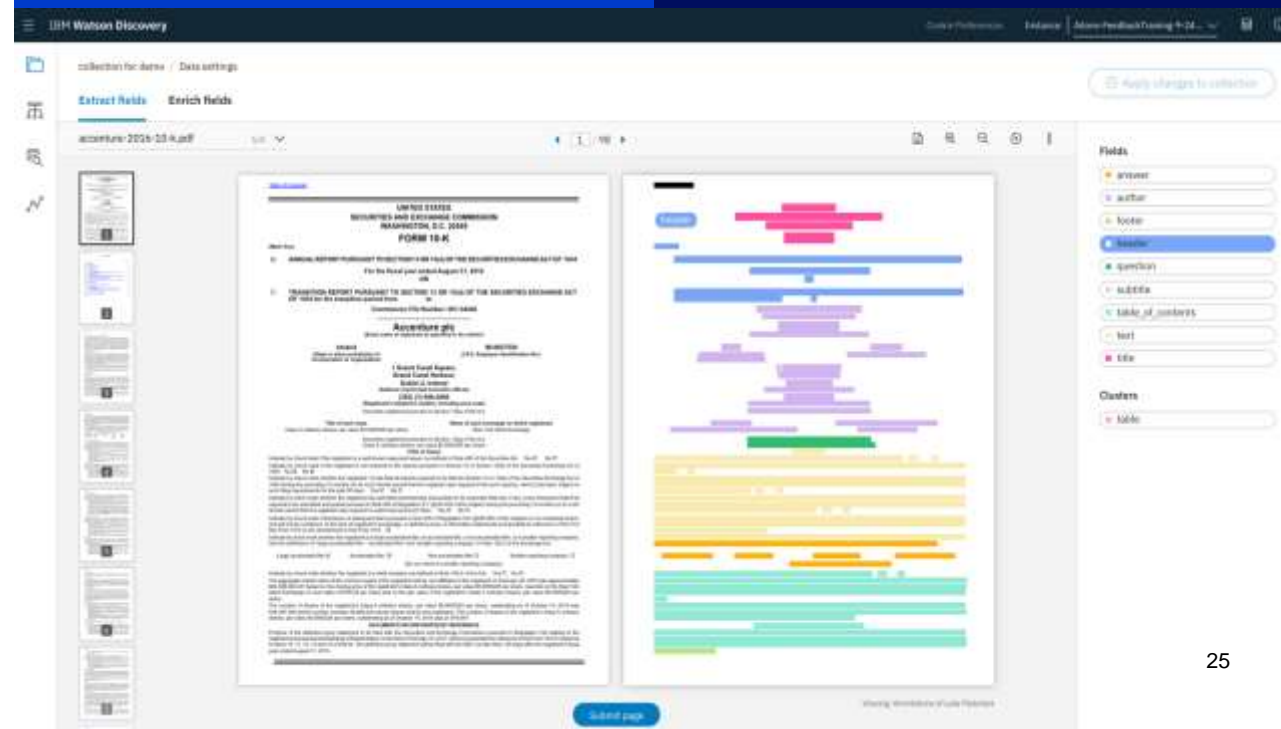
Relevancy Training

Teach Discovery how to find relevant results with representative examples

- Increase overall relevancy of results across a large set of answer documents without training for specific content

- Scales to learn from general, corpus-wide signals that indicate relevance rather than query or document specific signals
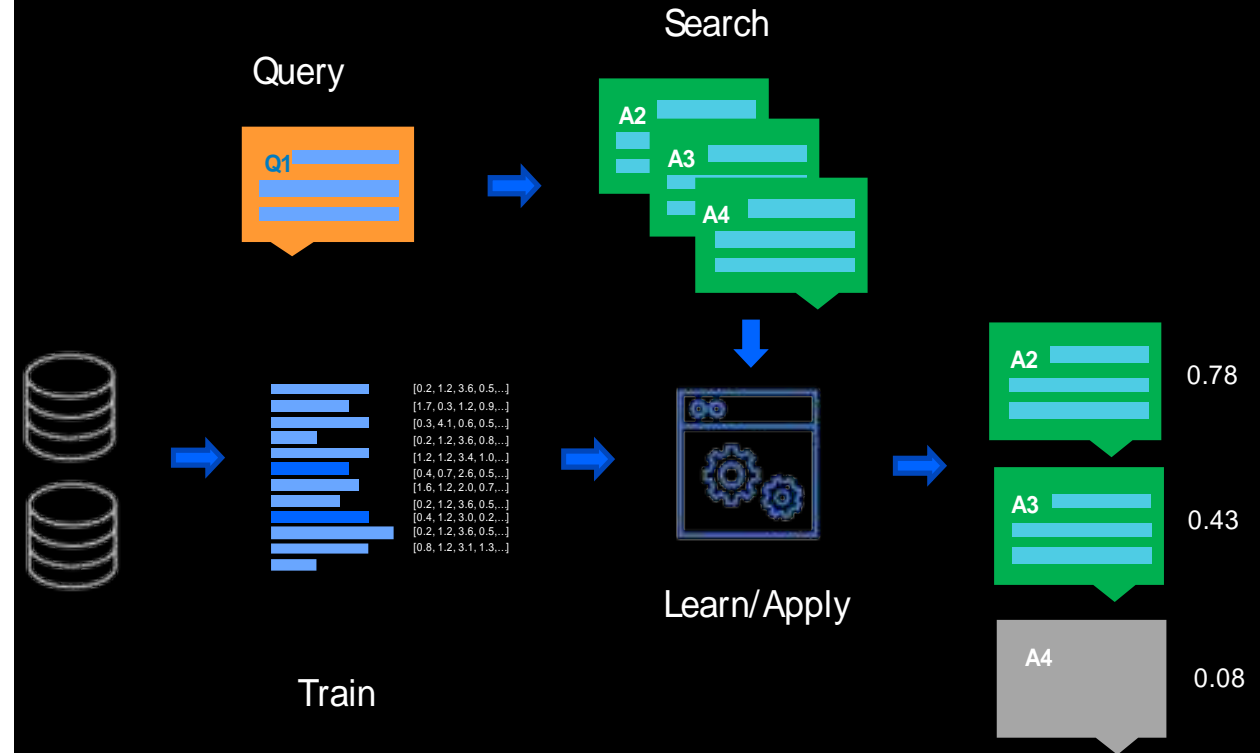
Connectivity

Tooling based connectivity framework for easy loading of data

Available In Q2

- Box*

- SharePoint (Online and On-Prem)*

- SalesForce

- Web Crawl

Available in a July fixpack

- Fileshare*

- Database (ODBC)

* Includes support for Doc Level Security

Query confidence prediction

Discovery uses pre-trained models to estimate the relevance of a result and return a normalized confidence score

Allows for **setting thresholds** to improve overall relevance

Can be used in combination with Watson Assistant to determine when to use long-tail results

Planned Improvements:

- **More accurate** scoring of results using semantic and statistical features
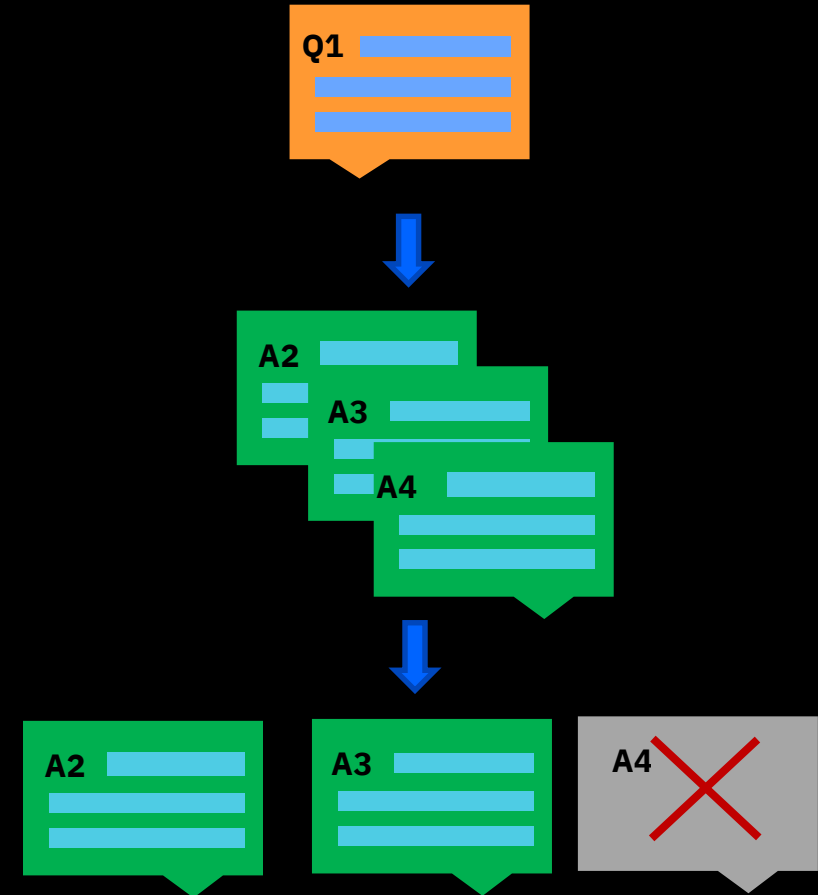
Custom Model Enrichments

- Watson Knowledge Studio now included as part of Bundle/Add-On

- Custom Model enrichment support now available on WD on ICP4D
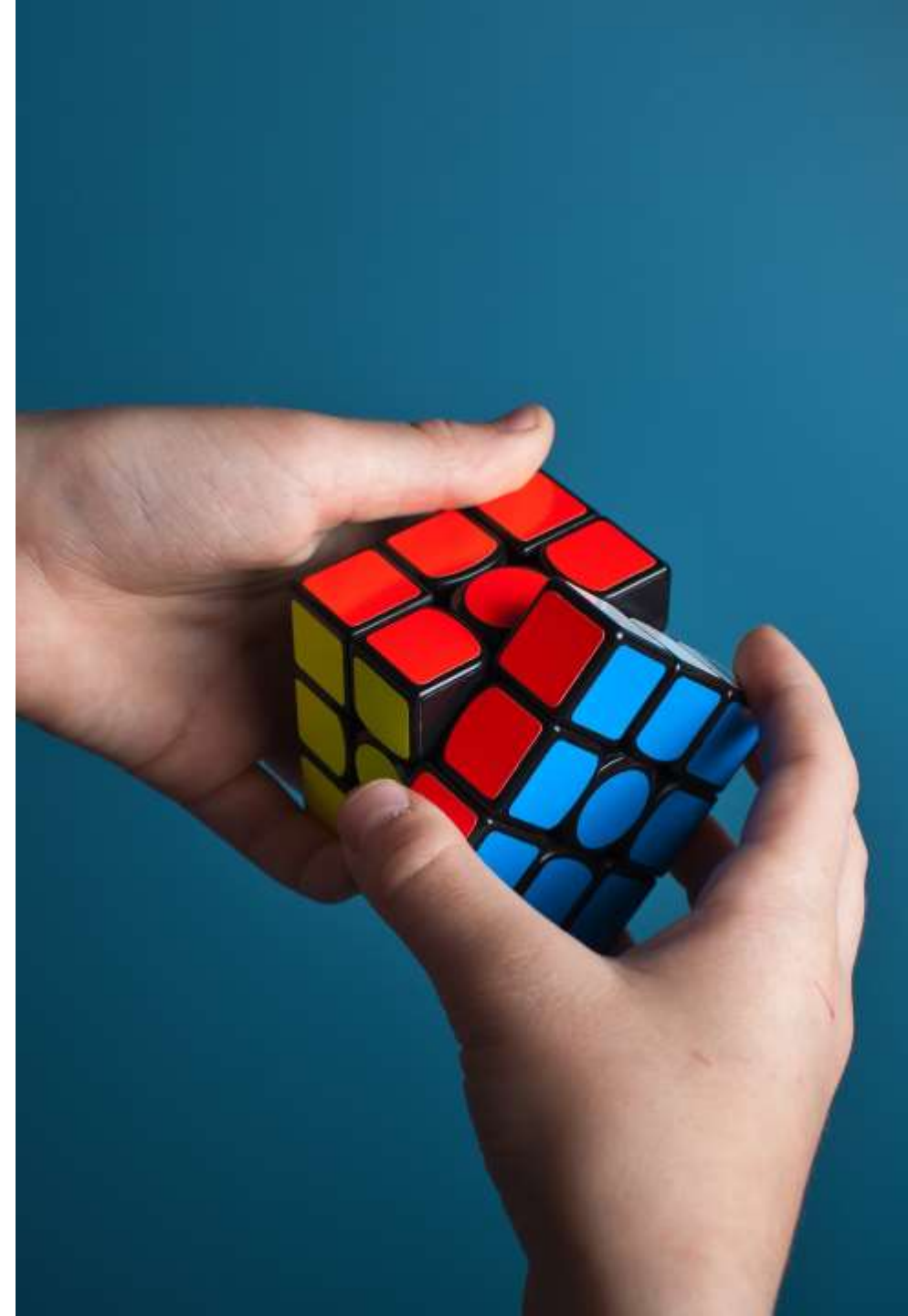
Watson Assistant Search Skill

- **Improve the coverage of your assistant** by calling out to a "search skill" from any dialog node

- **Search results are formatted automatically** by search skill

- **Ingestion and Sync to common knowledge base and CMS products**, like Salesforce Knowledge, Box, Sharepoint, and others...

Language Support (Same on Public and IBM Cloud Pak for Data)

Arabic

Japanese

Dutch

Chinese Simplified

English

Portuguese (BR)

Italian

German

Korean

Spanish

French

Some Languages Supported
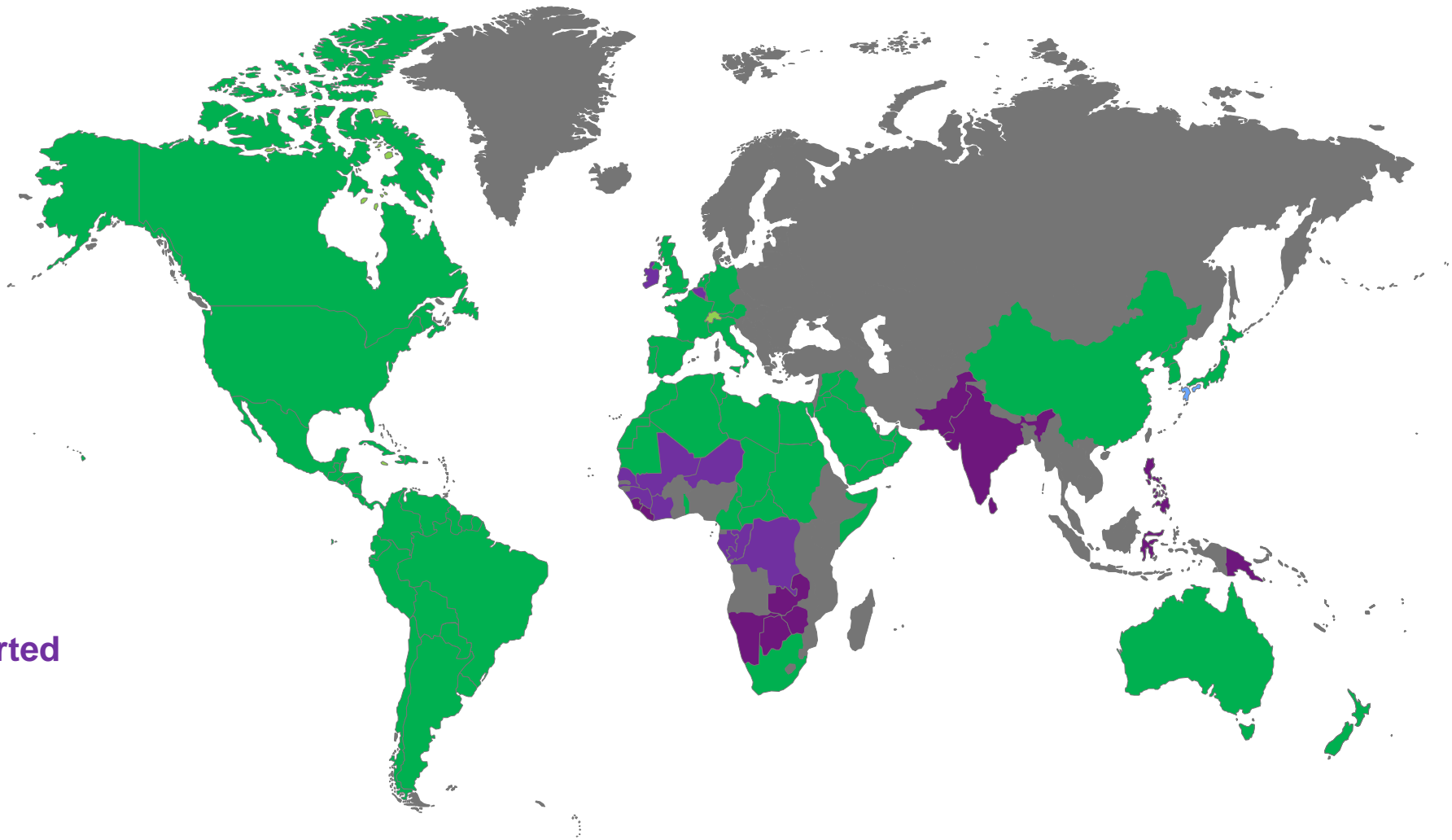
WA + Discovery for all types of Qs

"How do I turn on the headlights? "

"My exhaust is making a rattling sound, how do I troubleshoot the problem?"

"How do I restart my device?"

"My device won't power on. I am concerned if I do a reset that I will lose my data. What should I do?"

Frequency of Questions

Simple Questions

Complex Questions

100s

100,000+

Unique Intents
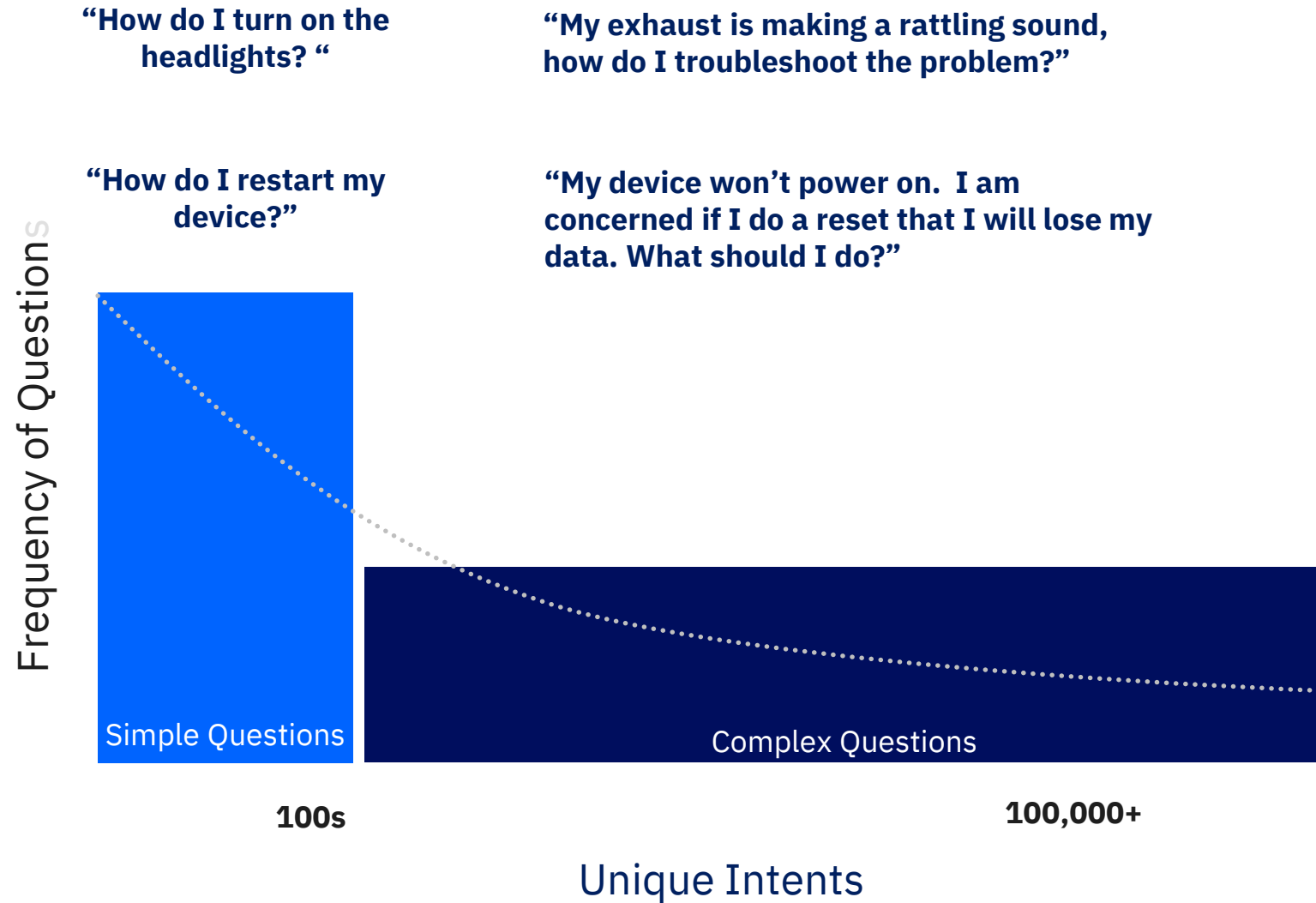
Watson Assistant

Watson uses reasoning strategies that focus on the language and context of the **question** to answer simple questions.
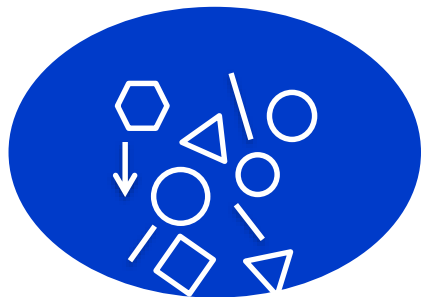
Watson Discovery

Watson uses reasoning strategies that focus on identifying the most appropriate **answer** for complex questions using the context of the entire corpus.

Watson Knowledge Studio

Teach Watson the language of your domain.

Complication

Ambiguity of natural language makes it complex for people to train machines to read it.

Isolation

Isolated development environments make it challenging for domain experts and developers to collaborate.

Application

Custom natural language models are difficult to actually apply to data, requiring further technical work and complex architectures

Teach by Example

Empower domain subject matter experts to teach Watson the language of your industry or organization without requiring deep technical skills or coding.

Engage your Experts

Bring your experts together into a collaborative online environment to teach Watson how to understand the linguistic nuances of your domain.

Use Everywhere

Apply what you've taught Watson in multiple applications by using models from Watson Knowledge Studio in Watson Discovery, Watson Natural Language Understanding and Watson Explorer.

Demo