IBM Watson

Introduction to **Alchemy API**

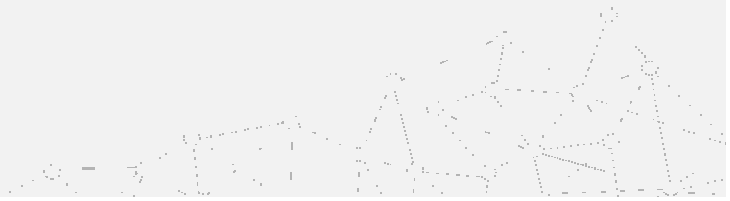
September 22, 2015

IBM

Who is **AlchemyAPI**?

Pioneer of web services for real-time text and image analysis

- Founded in 2005
- 60,000+ developers
- Used in 36+ countries
- > 1.5 Billion calls a month
- Acquired by IBM Watson
- Democratizing artificial intelligence



What can **AlchemyAPI** do for you?

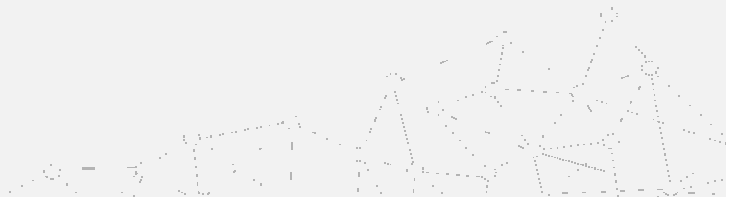
- Quickly extract meaningful information from unstructured data
- Categorize and label text and images
- Analyze sentiment
- Connect your data to additional knowledge sources

To get **started** you can…

- Get an API Key
- Download an SDK
- Check out the Getting Started Guides
- Ask me questions!

The **most popular** APIs

- *AlchemyLanguage*
  - Named Entity Extraction
  - Sentiment Analysis
  - Keyword Extraction
  - Relations Extraction
  - Taxonomy
- *AlchemyVision*
  - Image Tagging
  - Face Recognition
- *AlchemyData News*
  - News



Making **API Calls**

RESTful APIs for understanding data from text, URLs, images, etc.

- Endpoints available in API documentation

  http://access.alchemyapi.com/calls/text/TextGetRankedNamedEntities
- Required parameters

  ?apikey=<YOUR API KEY>&text=IBM Watson&outputMode=json
- Full URL

  http://access.alchemyapi.com/calls/text/TextGetRankedNamedEntities?apikey=<YOUR API KEY>&text=IBM Watson&outputMode=json

Named **Entity Extraction**

Extract people, places, organizations, etc. from blogs, news articles and other text.

See all
entity types

- **Examples:**
  - ***President Obama*** *works at the* ***White House****.* View Results
  - ***Zach Walchuk*** *was born in* ***Mankato, MN*** *and lives in the white house down the street.* View Results

View Results

Sentiment Analysis

Extract positive and negative opinions from any text.

- Document-level, entity-targeted, and keyword-targeted sentiment mining.
- Understands negations, amplifiers, diminishers, typos, hashtags and slang.

- **Examples:**
  - Overall Text Sentiment
    - *So happy the Patriots won the Super Bowl. Let's party. #PatriotsParade* View Results
  - Targeted Sentiment
    - *So happy the Patriots won the Super Bowl Seahawks suck. #BradyBunch* View Results Sea View Results Pats

- **Run on entities/keywords, if you do not have a specific target in mind.**
  - Overall Text Sentiment
    - *So happy the Patriots won the Super Bowl Seahawks suck.* View Results

Keyword Extraction

Extract topic keywords, with higher level abstractions via Knowledge Graph.

- **Examples:**
  - *The New England Patriots raised the Lombardi Trophy over the city of Boston on Wednesday.* View Results

Keyword extraction is like making a "word cloud" of important words in your document.

View Results

Relations Extraction

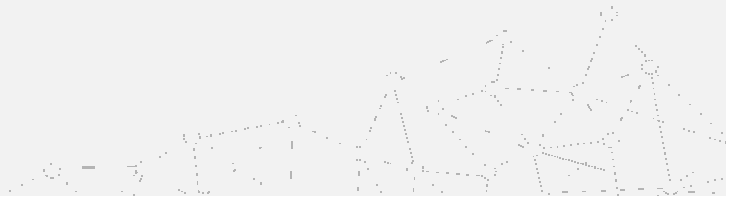
Extract the facts you need from raw text.

- Identifies Subject > Verb > Object
- **Examples:**
  - Article Title: *I want a sandwich* View Results

    “I” ➔ “want” ➔ “sandwich”

  - Article Title: *Charter buys Bright House* View Results

    “Charter” ➔ “buys” ➔ “Bright House”

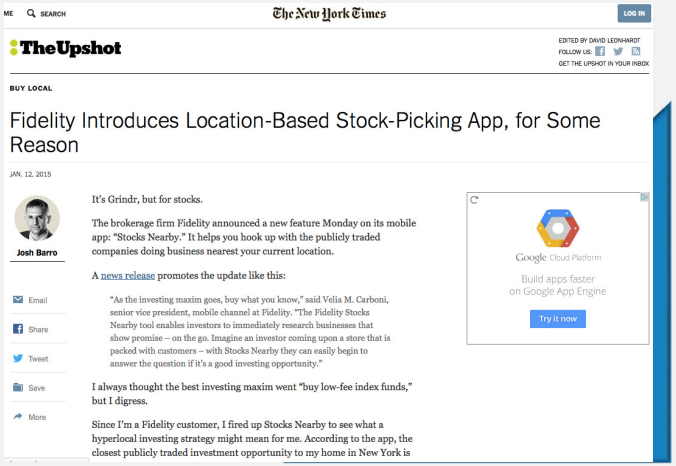


Taxonomy

Classify text documents based on 1000+ categories and subcategories.

- **Examples:**
  - *Lebron James has been struggling lately in his recent return to the Cavs.*
    View Results

View Results

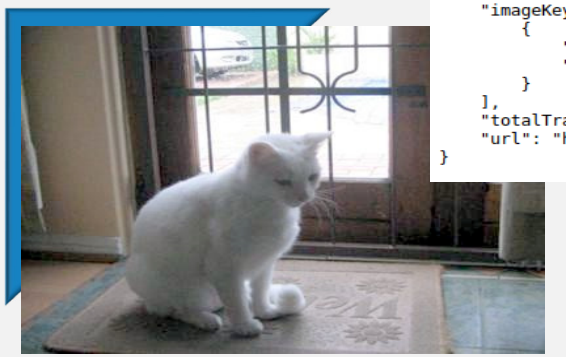


View Results

Image Tagging

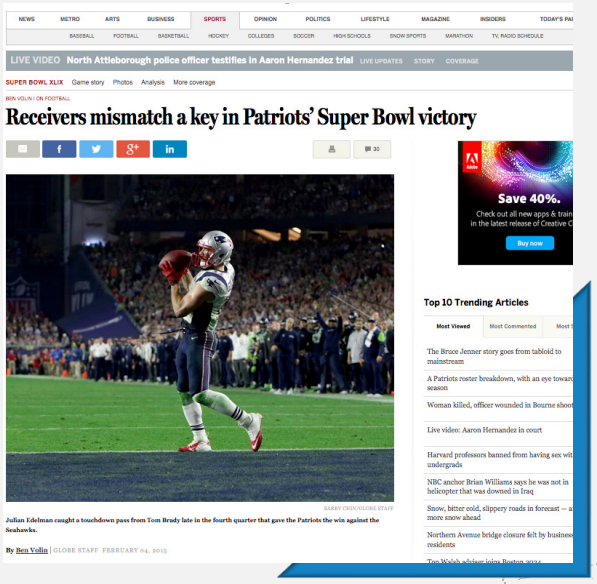
Extract keywords to categorize your image libraries.

- **Examples**:




```
{
    "status": "OK",
    "imageKeywords": [
        {
            "text": "cat",
            "score": "0.983697"
        }
    ],
    "totalTransactions": "4",
    "url": "http://www.pictures-of-cats.org/images/turbo-the-cat-sat-on-the-mat-21349914.jpg"
}
```


View Results

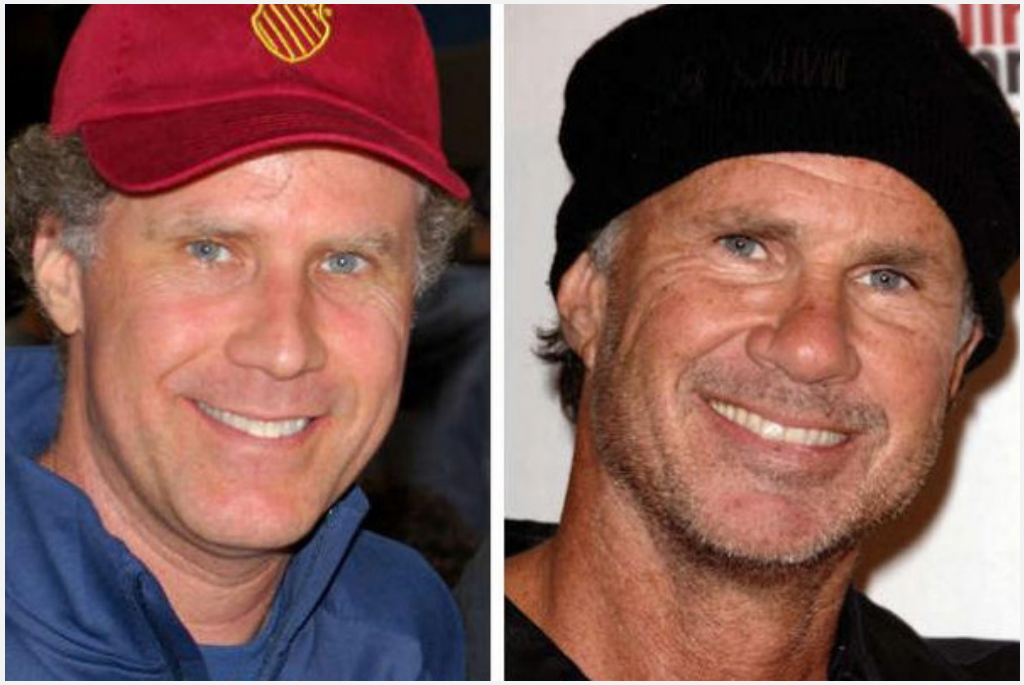


Face Detection

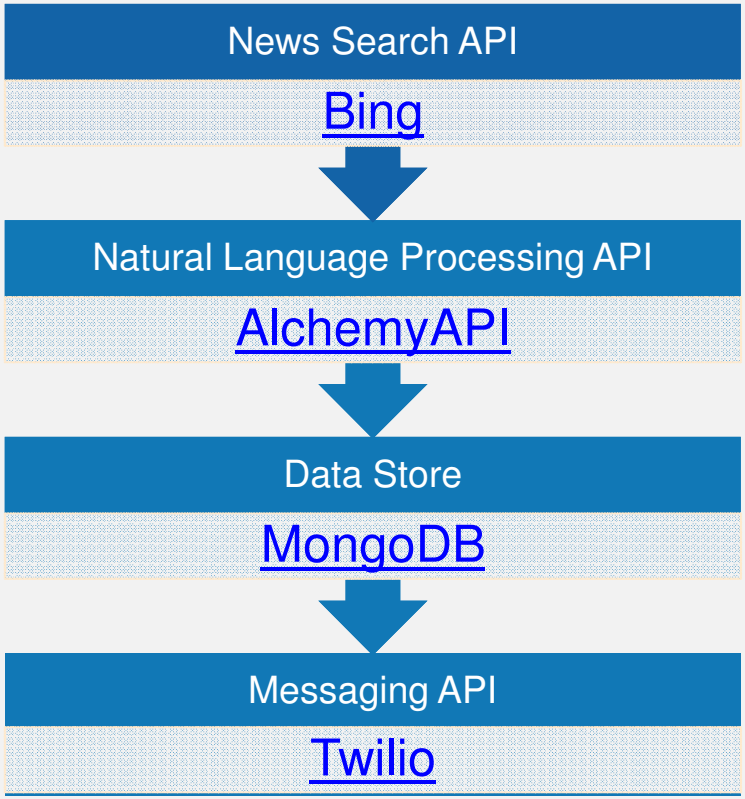
Gather intelligence from images and tag photos with the position, age, gender, and identities of people photos.

- **Examples:**

View Results



AlchemyAPI in the **Application Stack**



- Crawl the web with a Search API to find news articles relevant to your interests.
- Add intelligence and value with an NLP API. Filter search results to include specific actions (‘acquired’, ‘bought’, ‘sold’, etc.), or categorize the articles in a taxonomy.
- Store your enhanced search results in a database to access and analyze later.
- Send text alerts based on keywords, actions or trends found in your news article collection.

AlchemyData News

Query pre-enriched news articles to find trends, monitor events, and get up to date information

- Documentation
- Query builder with examples
- Sample application – Who's in the News
- Webinar:
  - 3 Ways to Enhance Your Product With Cognitive News

How others have used AlchemyAPI

Others have utilized AlchemyAPI to create enhanced solutions

- Watson News Explorer
  - Blog Post
  - Actual Application
- TED Talks Query Engine
  - Blog Post
  - Actual Application

IBM Watson

Backup

IBM

Alchemy's Platform: Transforming Data Into Value

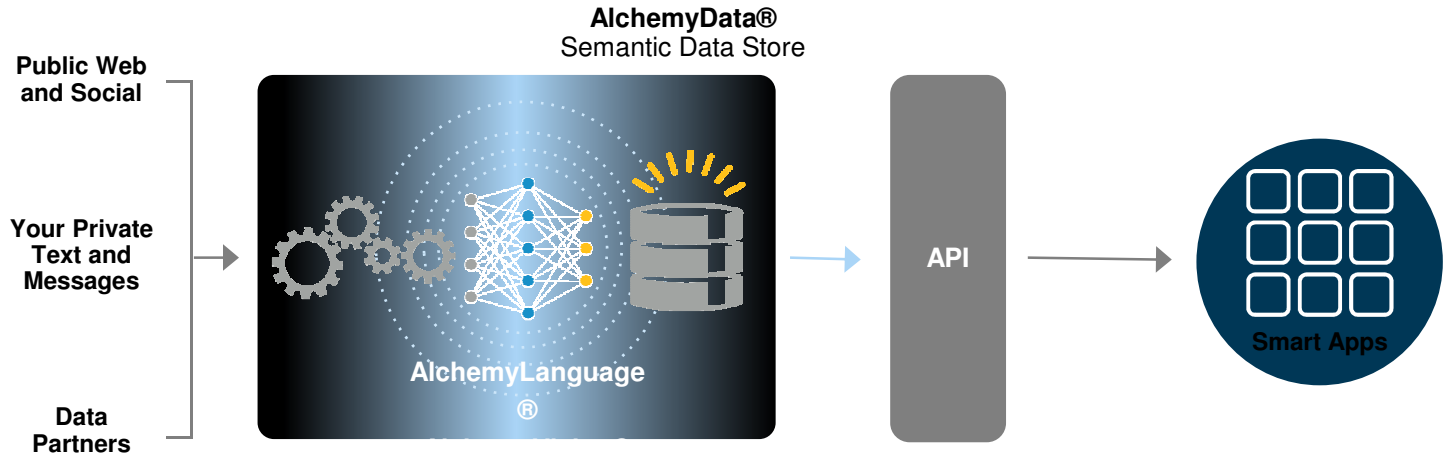
The world's unstructured data

Acquired and cleaned seamlessly

Stored and enriched with AlchemyAI

Accessed via simple REST APIs

For Smart Apps and Services



Alchemy's platform eliminates ***friction*** when extracting business value from unstructured data

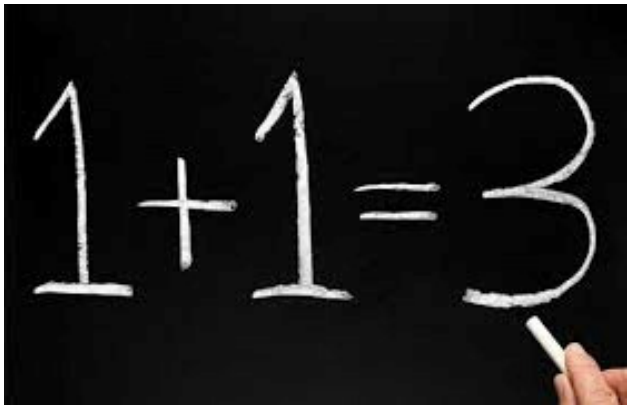
Service Accessibility

- Service is testable via Browser's URL bar
- Easy API keys, no complex authentication
- One key works across all services
- Common API parameters across all services
- All services return consistent XML and JSON
- Responses are easy to parse and understand
- Real-time usage tracking / API limits adjustment

Making It Easy

- Only platform to integrate real-time crawling
- Automatic content type detection
- Automatic language detection
- Automatic content cleaning
- Steerable page content extraction
- URL -> Metadata – no training, models, or anything
- Automatic entity linking + enrichment
- Individual + “Combined” API calls

Combinatorial Effect



- Dozens of enrichment functions
- Combining functions enables new use cases:
  - *Authorship + Entity + Sentiment = PR Targeting*
- More than any single algorithm
  - *Drives customer retention*
- Integrated data storage + analytics
  - *AlchemyData Buckets*
- Only full-stack solution
  - *NLP+Vision+Analytics*
- Only real-time solution
  - *Crawling+Cleaning+Enrichment*

AlchemyAPI Language

IBM Watson

Entity Extraction

Named entities specify things such as persons, places and organizations. AlchemyAPI's named entity extraction is capable of identifying people, companies, organizations, cities, geographic features and other typed entities from your HTML, text or web-based content.

Identifies named entities within subjects and objects of identified Subject-Action-Object relations, supporting advanced entity recognition capabilities including disambiguation, core-reference resolution and linked data output.

Capabilities

- Adds a wealth of semantic knowledge to your content to help you quickly understand the subject of the text.
- One of the most common starting points for using natural language processing techniques to enrich your content
- Retrieves a grouped, relevancy-ranked list of all named entities occurring within an HTML or text document
- * Linked Support Data

Key Value Proposition

- ✓ Add a wealth of semantic knowledge to your content
- ✓ Quickly understand the subject of the text
- ✓ One of the most common starting points for NLP

Use Case by Industry

- **Contextual Advertising -** Taboola uses our Entity Extraction and Concept Extraction APIs to have a contextual understanding of a URL. Based on that contextual understanding of the page this information is fed into the Taboola recommendation engine to place the most relevant piece of native advertising for the user.

Market References

- Spiderbook – CRM
- Waggener Edstrom – Public Relations
- Pocket – Content Discovery
- CEBRA – Public Health and Safety

Conversation Starters

- Identifying persons, companies, places and things in text is important
- To understanding the context of a piece of text whether it's a customer service email an internal document or an article on the web talking about your competition.

AlchemyAPI Language

IBM Watson

Sentiment Analysis
"Voice of the Customer"

Sentiment is the attitude, opinion or feeling toward something, such as a person, organization, product or location

Provides easy-to-use mechanisms to identify the positive or negative sentiment within any document or webpage.

Capable of computing document-level sentiment, sentiment for a user-specified target, entity-level sentiment, quotation-level sentiment, directional-sentiment and keyword-level sentiment (enables social media monitoring)

Capabilities

- Adds a wealth of semantic knowledge to your content to help you quickly understand the subject of the text.
- One of the most common starting points for using natural language processing techniques to enrich your content
- Retrieves a grouped, relevancy-ranked list of all named entities occurring within an HTML or text document

Key Value Proposition

- ✓ Understanding customers
- ✓ Trend Analysis
- ✓ Social Media monitoring

Use Case by Industry

- **Consulting –** Prior to using our Sentiment Analysis service Gallup was hiring a small army of temporary employees to manually read and score hundreds of thousands of customer survey responses for their customers. This process regularly took up to several months and was quite expensive due to its manual nature. Simply put, this process doesn't scale. Upon evaluating Alchemy's Sentiment Analysis tool they found that it was close to 80% accurate. The processing time to apply a Sentiment score to went from months to days with Alchemy's service. With mountains of data located in CRMs, Email, Chat, Social Media and voice conversations that are then transcribed to text the sources of customer feedback are many and the amount is vast and overwhelming for most enterprises.

Market References

- Pulsar – Marketing
- AdTheorent – Advertising

Conversation Starters

- How well is your organization understanding the voice of your customers? Today's consumers are using Social Media, Email, Text Message, Blogs, Community Forums, Customer Service Chats and telephone conversations to voice their opinion. All of this is unstructured data with valuable insights waiting to be extracted so you can act on them in real-time.

AlchemyAPI Language

IBM Watson

Keyword Extraction

Keywords are the important topics in your content and can be used to index data, generate tag clouds or for searching.
Capable of finding keywords in text and ranking them– a relevance score based on statistical analysis.
Determines the negative or positive sentiment associated with each identified keyword.

Capabilities

- Employs sophisticated statistical algorithms and natural language processing technology to analyze your content and identify the relevant keywords.
- Supported in over a half-dozen different languages, enabling even foreign-language content to be categorized and tagged.
- Works on URLs, HTML documents and plain text.
- Automatically detects the language of the content and then performs the appropriate analysis.

Key Value Proposition

- ✓ Index data
- ✓ Generate tag clouds
- ✓ Searching

Use Case by Industry

- **Tax –**Intuit company had 20M documents that were primarily User Generated Content or "UGC". The story around this use case is not knowing what assets an enterprise has for either internal search or SEO enhancement for public search.
- **Market research** - Consumers reveal their preferences for products or services based on their browsing behavior and in online dialogues (downloading and reading content, search history, topics pursued in social channels etc.)
  Finding the actionable information in the "noise" and volume of data. Identify conceptual relationships. Using Keyword Extraction and Relation Extraction API calls from AlchemyAPI consumer intent is far more accurately determined than traditional methods such as Surveys and focus groups.
- **Voice of the Customer / Customer Satisfaction–** Use Alchemy's Keyword Sentiment Analysis to identify which keywords in a document have negative sentiment associated with them. We ran 5,000 app reviews on the consumer app Mint by Intuit and found keywords with negative sentiment to be a strong signal for ways the app developers could improve the application.

Market References

- Pulsar - Marketing
- Spiderbook - CRM
- Waggener Edstrom – Public Relations
- AdTheorent – Advertising
- Triberr - Publishing
- CrisisNET – Crisis Data Platform
- BrainJuicer – Market Research

Conversation Starters

- Similar to Entity Extraction, identifying keywords in text is important to understanding the context of a piece of text whether it's a customer service email an internal document or an article on the web talking about your competition.

IBM Watson

Spiderbook – CRM

IBM

Spiderbook

AlchemyAPIs Used

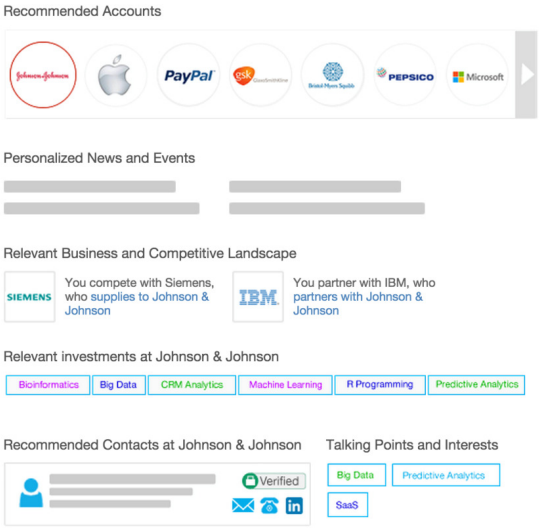
- Keyword Extraction
- Language Detection
- Entity Extraction

What it is:

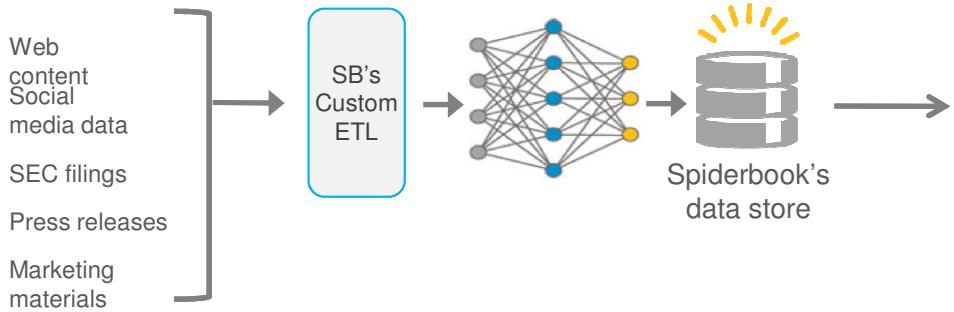
- Challenge: processing all of the content available to provide deal-closing business intelligence.
- Go beyond traditional CRM by using NLP and entity recognition– Empowering sales teams to close more deals by predicting prospect behavior
  - Analyzes: 1,000+ photos / day, 40,000 users interacting with site daily, 100,000+ registered restaurants
  - Types: public online data (website content, PR< marketing materials, social media, SEC filings, LinkedIn, SlideShare), business profiles through partnerships with data services providers
- Solution: a predictive customer relationship discovery app that mines the entire Web

ROI

- Creates 10x more accurate customer relationship predictor using AlchemyAPI
- Value: speed is great, accuracy is paramount – AlchemyAPI delivers both
- Outperformed its closest competitor by 20-30%
- Results: SpiderGraph (unique business intelligence service), exponential user adoption ($1M seed funding)



Additional Notes



IBM Watson

Pocket – Content Discovery

IBM

pocket

AlchemyAPIs Used

- Entity Extraction
- Concept Extraction

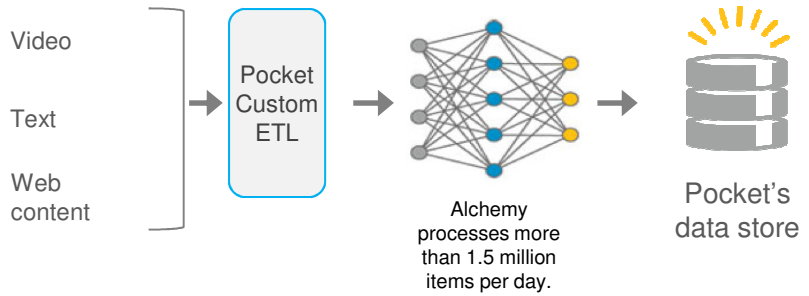
What it is:

- Challenge: pinpoint concepts and entities within saved content to enable accurate classification of text, video and web pages (save for later capabilities)
  - Analyzes: 500,000 unique items/ day, 1.5 million items/ day total
  - Types: HTML, text
- Pocket: cloud-hosted APIs decipher unstructured data and HTML; actionable data retuned
  - Broad language support
  - Discovery of articles worldwide for everyday users to interact with
- Solution: AlchemyAPI's entity and concept extraction APIs

ROI

- Results: classification accuracy for winning user experience, higher performance of in-house developed classification and recommendation APIs

Additional Notes



IBM Watson

Pocket – Content Discovery

IBM

Simply Measured

What it is:

- Monitor social media to help clients understand their customers' needs
  - Who should we engage?
  - What do our followers care about?
  - What content and promos should we offer?

ROI

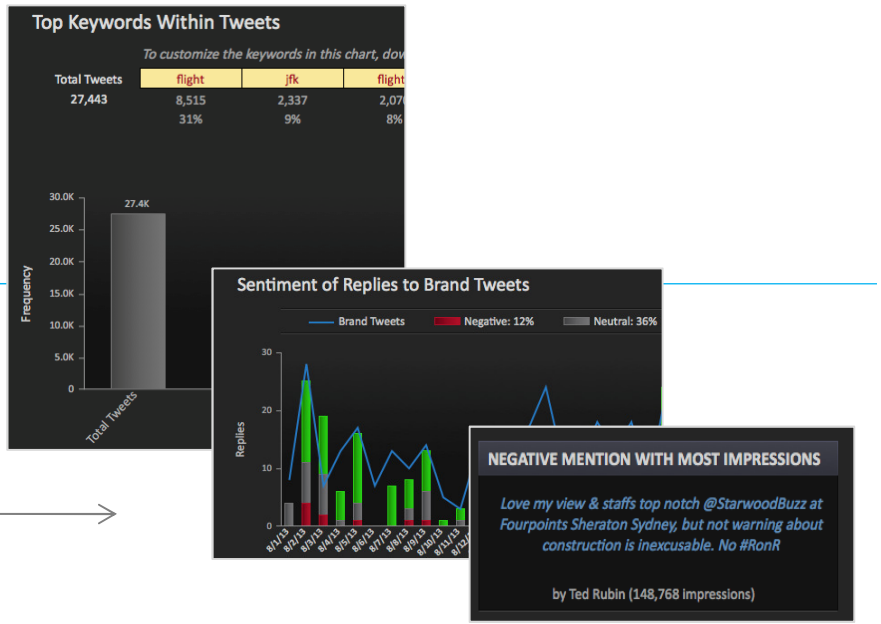
Results

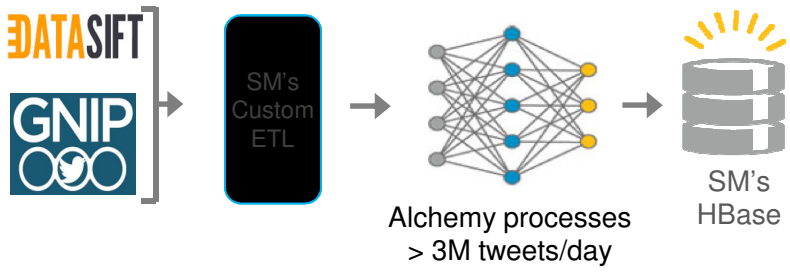
- Real-time trending of topics followers care about
- Increased engagement by targeting content and offers relevant to users' interests

AlchemyAPIs Used

- Concept Extraction
- Entity Extraction

Additional Notes





IBM Watson

TABElog – Retail

AlchemyAPIs Used

✓ Image Analysis

What it is:

- Challenge: finding images that help foodies virtually experience restaurants
- Serving up accurate restaurant reviews with Image analysis
  - Combines food reviews, photos and guest blogs with AlchemyAPI's computer vision services to enable foodies to discover the best dining experiences across the US.
  - Analyzes: 1,000+ photos / day, 40,000 users interacting with site daily, 100,000+ registered restaurants
  - Types: images and photos
- Solution: Image Analysis REST API that finds the best images to represent a restaurant

ROI

- Reduces the amount of time needed for the mundane, repetitive task of separating useful images from mountains of unrelated or inappropriate content.
- Value: increased productivity, dramatic time savings, desired results
- Results: Tabelog.us ("specialized social networking site"), viral adoption by restaurants and users, rapid application and site development

Additional Notes

- Why the Image Analysis API?
  - "As an experienced developer, one of the first things that impressed me was how quickly I implemented the Image Analysis API."
  - "It only took an hour to get a custom library built and working with it."
    "The simplicity of the REST API makes it language agnostic and gives us the cross-domain integration we need."
  - Other factors: performance based– competitive services did not reliably filter images like they wanted.

IBM Watson

Pulsar – Marketing

PULSAR Social Data Intelligence

AlchemyAPIs Used

- ✓ Language Detection
- ✓ Keyword Extraction
- ✓ Sentiment Analysis

What it is:

- Challenge: going beyond metrics to reach insight because ineffective in collecting and interpreting streams of social data
- Helping brands understand customers with social media intelligence
  - Analyzes: 20-30 million documents/ month
  - Types: Unstructured data across social platforms
- Social intelligence platform
  - Integrates AlchemyAPI's services in every product on social media monitoring platform
  - Tailors outputs to meet clients' specific objectives through data surfacing
  - Better understand what customers want in products and demand in their experiences

ROI

- Value: easy-to-integrate NLP capabilities enable rapid monetization
- Result: superior social data intelligence, accelerated R&D at significantly lower cost, expanded product lines and innovation of social media intelligence products, optimizing social channel

Additional Notes

- **Social Panels** (newest development): designed to listen to the online conversations of specific audiences
- Ability to apply metadata that distinguishes relevant versus irrelevant content
  - Designate an audience and listen to everything that group publically discusses through all social channels
- On R&D savings:
  - "AlchemyAPI's services do a lot of the legwork, allowing us to slice, dice and structure the data so that our clients can understand their audience and differentiate based on what they learn from the conversations happening around them"

IBM Watson

Waggener Edstrom – Public Relations

IBM

WAGGENER EDSTROM COMMUNICATIONS

AlchemyAPIs Used

- ✓ Keyword Extraction
- ✓ Concept Tagging
- ✓ Named Entities
- ✓ Author Extraction
- ✓ Text Extraction

What it is:

- Challenge: Analyzing real-time data to determine if the content is relevant to a specific customer (website content, social media, news coverage)
- Enabling agile PR strategies by gathering intelligence and identifying trends
  - Analyzes: 1.5 million tweets/ day, 2.5 million posts/ day
  - Types: Website content, social media, news coverage
- Data mining and analytics platform (WE Infinity) that identifies trends and discovers competitive intelligence showing where and how performance improvements can be made
- WE Infinity – Near real-time analysis using 5 AlchemyAPI services

ROI

- WE Infinity: real-time data mining and analytics platform, 80% automation, real-time communication and performance analytics
- Value: data-driven decisions create a bigger bang for every campaign dollar
- Result: prototype app developed in one week, WE infinity (real-time data mining and analytics platform)

Additional Notes

- **5 APIs used:**
  - **Keyword Extraction:** finds and ranks specific keywords in page content and index.html
  - **Concept Tagging:** enables abstraction, understanding how concepts relate, and tags accordingly
  - **Entity Extraction:** identifies people, companies, organizations, cities, geographic features and other entities from HTML, text or web-based content
  - **Author Extraction:** identifies the source for news articles and blob posts
  - **Text Extraction:** retrieves the text from web pages and strips the 'chrome' while allowing return of embedded links; Infinity also extracts content titles using Text Extraction
- Analysis services are 'turned on and off as is appropriate'
- Content is processes through AlchemyAPI REST services and then sent to an editing application that matches articles and discussions to concepts that WE's account teams and clients want to further review

IBM Watson

CEBRA – Public Health and Safety

cebra Centre of Excellence for Biosecurity Risk Analysis

AlchemyAPIs Used

- ✓ Named Entity Extraction
- ✓ Text Extraction
- ✓ Language Detection

What it is:

- Challenge: real-time discovery, sharing of biosecurity information
- Identifying and sharing data on infections requires daily Google queries and rigorous monitoring for mentions of tainted products/ outbreaks in RSS feeds, search engines, industry journals and Twitter
  - Analyzes: 6,000 unique text items/day
  - Types: text
- International Biosecurity Intelligence System (IBIS)
  - Web 2.0 intelligence network for plant and animal biosecurity devoted to collecting and organizing information used to track and forecast diseases/ follow emerging diseases
- Solution: AlchemyAPI's text extraction, language detection and named entity extraction

ROI

- Value: excellent situational awareness and competitive edge
- Results: 206 RSS feeds monitored/day, 283 separate Google queries/ day, real-time information collection and archiving, reliable resource for proactive disease control, increase automation

Additional Notes

- Top 20 results for each query and any new URLs are added to the results queue for processes. Processing is executed by calling:
  - URLGetAuthor
  - URLGetLanguage
  - URLGetRankedNamedEntities
  - URLGetRawText
  - URLGetText
  - URLGetTitle
- Each processed article becomes a new 'node' in IBIS database
- Node includes original title and text, then translated title and text, the author and the location
- Geonames are used to convert place names found by AlchemyAPI into lat/long coordinates