Top Natural Language Processing Applications in Business

UNLOCKING VALUE FROM UNSTRUCTURED DATA
For years, enterprises have been making good use of their structured data (tables, spreadsheets, etc.). However, the larger part of enterprise data, nearly 80 percent, is unstructured and has been much less accessible.

From emails, text documents, research and legal reports to voice recordings, videos, social media posts and more, unstructured data is a huge body of information. But, compared to structured data, it has been much more challenging to leverage.

Traditional search has done a great job helping users discover and derive insights from some of this data. But enterprises need to go beyond search to maximize the use of unstructured data as a resource for enhanced analytics and decision making.

This is where natural language processing (NLP), a field of artificial intelligence (AI) that’s used to handle the processing and analysis of large volumes of unstructured data, can be a real game changer.

While AI describes a broad range of technologies, NLP is one of three AI-driven capabilities that enterprises can readily harness to create business value and competitive advantage:

1. Internet of Things (IoT): applying technologies, such as real-time analytics, machine learning (ML), and smart sensors, to manage and analyze machine-generated structured data
2. Computer Vision: using digital imaging technologies, ML, and pattern recognition to interpret image and video content

As it powers document understanding applications that unlock value from unstructured data, NLP has become an essential enabler of the AI evolution in today’s enterprises.

This white paper discusses the emergence of NLP as a key insight discovery technique, followed by examples of impactful NLP applications that Accenture has helped clients implement.

---

**Natural Language Processing Market Size Estimate**

$16B by 2021

**Source:** www.marketsandmarkets.com/PressReleases/natural-language-processing-nlp.asp
Emergence of NLP applications in the enterprise

As AI technologies like machine learning (ML), deep learning, and cognitive search have entered the mainstream, we’ve seen a rapid evolution in NLP applications. This has been driven by two key factors. First, having grown accustomed to virtual assistants like Siri and Alexa, as well as sophisticated search like Google, in their daily lives, users expect the same experience in the workplace. Second, today’s NLP no longer relies on manual, rules-based approaches: integrating ML with NLP allows for far greater automation, scalability, and accuracy.

Enterprises can leverage NLP applications in two ways – query understanding and content understanding – to improve both user experience and insight discovery:

- Provide better, more targeted responses by understanding the user’s questions and intent
- Identify out-of-scope requests and present intelligent alternatives

Understanding content (mining insights from vast amounts of unstructured data)

- Extract business entities from text documents to identify employees, customers, products, procedures, etc.
- Identify and understand the meaning of natural language content – documents, reports, emails, etc. – to provide natural language answers

NLP enables improved understanding of user queries and enterprise content. This ensures that every user is connected to the most relevant, helpful resources which would otherwise remain hidden within vast quantities of data.

Today’s NLP applications – spanning industries and ranging from chatbots and virtual assistants to intelligent business solutions – are supported by a robust NLP technology ecosystem.
NLP technology ecosystem

Rising demand for NLP-powered applications is fueled by a growing range of technologies, from open source frameworks and evolving vendors to cloud-based APIs. Commonly-used NLP technologies include:

**Open source frameworks**
- Python NLTK
- Python SpacY
- Python GenSim
- UIMA
- GATE
- Apache OpenNLP

**Industry specific platforms**
- Quid (market and competitive intelligence)
- Twiggle (e-commerce)

**Cloud API providers**
- Google Natural Language API
- Amazon Comprehend
- Microsoft LUIS (Language Understanding Intelligence Service)

**Evolving vendors**
- Pool Party
- SmartLogic
- Cotical.io

**Other:**
- SAS NLP offering
- IBM Watson

While NLP is a relatively new enterprise technology, it’s being enhanced every day. The ML algorithms supporting NLP are improving all the time, with industry giants like Google, Microsoft, and Amazon all making strides to improve accuracy.

With this growing ecosystem of NLP solutions, enterprises have increasing flexibility to select appropriate approaches and toolsets for their specific use cases. Accenture leverages this robust ecosystem and our own technology assets to orchestrate the different components of NLP applications – making them easily maintainable and scalable.

We’ve helped many clients implement NLP applications. Some of the most impactful are highlighted in the next section.
Enterprise NLP applications in action
Chatbots and virtual assistants

CHALLENGE

Chatbots are becoming ubiquitous. All around us, Siri, Alexa, Cortana, and Google Home are incorporating natural language conversations between humans and AI into everyday interactions. As consumers get accustomed to these virtual assistants, they expect the same experience in the workplace. The emphasis is on how well computers talk to humans, interact with them, understand text and so on.

The challenge for enterprises? Providing their people with sophisticated natural language understanding applications that meet their expectations – and improving information discovery and collaboration.

Sources:
1. Accenture Research: Chatbots are Here to Stay
2. IDC FutureScape: Worldwide IT Industry 2018 Predictions
3. Gartner: Conversational AI to Shake Up Your Technical and Business Worlds
Chatbots and virtual assistants (contd.)

SOLUTION

Powered by NLP, enterprise chatbots and virtual assistants can enable conversations between humans and computers in everyday business processes. Bringing deeper natural language understanding, they enhance search as well as providing an entirely new way for employees, customers and partners to interact with enterprise data and work more productively. For example, an enterprise chatbot that can hold a dialog with the user to acquire targeted information (dates, times, order details, etc.) and/or integrate with business systems to complete common office tasks quickly (reservations, orders, etc.).

By combining content processing, ML, NLP, and voice recognition, enterprises can transform simple search-based intranets or support portals into AI-powered virtual assistants. Whether it’s a question about technical support on a consumer electronics website, pricing and product details on a partner or e-commerce website, investment products on an investor portal, or management advice on a hotel franchisee site, AI-powered chatbots and virtual assistants can play key roles in process automation and user experience improvement.

BENEFITS

- Improve business processes and reduce support costs by enabling AI-driven self-service virtual assistants for customers and partners
- Enhance search and knowledge-seeking experiences for customers, partners and employees
- Improve brand reputation through more efficient processes and more engaging customer experiences
Intelligent document analysis

**CHALLENGE**

With overwhelming volumes of incoming documents and correspondence every day, large organizations often struggle to properly analyze and derive useful insights from this content. Whether it’s legal documents, medical records, policies, or contracts, without a centralized and automated analysis approach, it may be difficult for organizations to effectively understand and make use of the document content to support decision making and operational efficiencies.

**SOLUTION**

Intelligent document analysis uses AI techniques including NLP, entity extraction, semantic understanding, and ML to analyze content, extract meaning, and reliably aid process automation and decision making. These applications can identify specific items of information in documents – like date, order number, and policy number – so they can be categorized and analyzed.

Core functionalities of this solution can include:

- **Optical character recognition (OCR)** – converts different types of documents, such as scanned paper documents, PDF files, or images into editable and searchable data
- **Text analysis** – analyzes documents to identify specific language or terms and extract linguistic meaning
- **Deterministic classification** – uses a pattern-based classifier to look for sequences of terms that indicate a specific sort of document
- **Machine learning** – trains a ML model with example datasets to predict document type or extract and classify text (e.g. learning aircraft component names)
BENEFITS

- Improve compliance and risk management
- Drive internal operational efficiencies
- Enhance business processes
Document search and match

**CHALLENGE**

Enterprises store vast quantities of documents, so it’s critical to be able to identify and match documents of related themes – not just keywords – quickly and accurately.

In the recruiting and staffing industry, for example, recruiters are in a race to find the right candidate before competitors identify the same person in their candidate databases. Given the large amounts of candidates that most recruiters deal with, it’s easy to lose track of the best matches.

There are many candidate search and match solutions delivered as standalone products or embedded into applicant tracking systems which can provide convenience, but they’re limited to basic resume parsing and matching corresponding metadata to a job posting. As a result, most of these embedded search systems fail to significantly improve fill rates or deliver a better experience to recruiters.

**SOLUTION**

The ideal solution is an application that can indicate people’s fit for jobs by accurately and automatically comparing job descriptions to people’s resumes, CVs or other documents.

The solution we’ve developed is bidirectional and works like this: a job description (the whole document) is submitted as a search request and the comparison system automatically returns a short list of the best-qualified candidates from a database of CVs. Alternatively, a job seeker (or professional recruiter) submits a CV and the system returns a list of the most appropriate, currently-available vacancies.

It’s an approach that combines advanced search techniques, analytics, NLP, and ML to provide statistical and linguistic capabilities for understanding applicant profiles and identifying the best candidates. Predictive analytics uses previously successful candidates’ placement histories to find additional candidates ahead of a new job posting – giving recruiters a head start against the competition.
ML leverages feedback from recruiters to improve results over time.

This solution also adds value to human resources teams in global enterprises, enabling them to identify the right people from hundreds of thousands of geographically-scattered experts. For example, a project manager can submit a project requirement document as a search request or ask a question like “Who are our NLP experts in North America?” The system will return a list of relevant experts for the project by automatically analyzing and identifying matches from employee databases.

### BENEFITS

- Reduce the time to fill a job position or to identify experts for specific skillsets within an organization
- Scale to millions of jobs postings and resumes (in recruiting) or multiple employee and project databases (in an enterprise “Find the Expert” application)
- Increase revenue and reduce costs – candidates’ matches can be found within the recruiter’s database rather than via paid job boards

### Delivering results across multiple use cases

Document search and match can be applied to various types of enterprise documents. Examples include analyzing legal documents for variations of risky contractual languages or, in the financial services industry, cross-analyzing loan or mortgage documents with borrowers’ profiles.
Data storage optimization analytics

CHALLENGE

Corporations routinely maintain petabytes of content in expensive, on-premise storage, so it’s no surprise that CIOs want to reduce costs by deleting or migrating stale or obsolete content. Standard storage management software applications can count the bytes, but they can’t compute the business value of content or the risk of losing data. Organizations need an approach for automating the migration of content to lower-cost storage, while providing file-level traceability and real-time tracking of ROI.

To reduce storage, organizations must answer these FAQs:

• When was the content created? Last accessed?
• Who created the content? Are they still employed?
• Which business unit owns the content?
• Which products are associated with it?
• How much does it cost to store?
• Is it covered by a data retention policy?
• Can we move to low-cost cloud storage?

SOLUTION

With a data storage optimization analytics solution, IT departments, CIOs, and content owners get the visibility they need to identify content eligible for deletion or migration to lower-cost storage.

First, organizations should use enterprise data connectors and search engines to help them understand the enterprise context for their data. Connectors help join content source metadata with other corporate sources, with source metadata and context indexed into a search engine and made accessible for ad hoc search and visualization.

Business rules should also be developed by stakeholders to help drive data movement to optimize storage utilization and achieve program ROI. Once the business rules are laid out, organizations can identify files eligible for lower-cost storage, archiving, or deletion, if obsolete. Where possible, this should be an automated process that provides traceability.

With NLP and ML, documents and their content can be more accurately read and identified. This drives storage cost savings through rapid detection of duplicate or near-duplicate content, as well as providing a 360-degree view of enterprise data.
Data storage optimization analytics (contd.)

Example business rules

- Retain active files on primary storage
- Archive content inactive for two or more years
- Archive content authored by contractors and terminated

BENEFITS

- Provide actionable visibility into content in high-cost storage
- Identify content provenance, ownership, and frequency of use
- Apply domain-specific business rules to identify eligible content and migrate it to lower-priced storage
- Enable cloud-readiness
- Track savings (potentially millions per year) and compute ROI in real-time
Sentiment analysis, also known as opinion mining, uses AI to automate the process of identifying opinions about a specific subject from a piece of content.

As mentioned earlier, it’s estimated that 80 percent of an organization’s data is unstructured in the form of emails, chats, articles, documents, web content, and social media. Manual analysis of this information for customer, product, or employee sentiment would be practically impossible. With AI-driven sentiment analysis, an organization can detect people’s opinions or feelings about a topic and uncover actionable insights that would otherwise be unobtainable.

In its simplest form, sentiment analysis categorizes a sentiment as positive or negative. It could also quantify the sentiment (e.g. -1 to +1) or categorize it at a more granular level (e.g. very negative, negative, neutral, positive, very positive). By combining NLP technologies, text analytics, linguistic analysis, and ML, sentiment analysis applications can cope with the complexities of language.
Example use case:

**Voice of the customer** – sentiment analysis is widely applied to voice of the customer materials like reviews and survey responses, online and social media, and healthcare materials for applications that range from marketing to customer service to clinical medicine. By using sentiment analysis across this data, organizations can better understand customers (and possibly change how they engage with them), predict demand, and improve overall company performance.

**BENEFITS**

- Provide marketing and competitive intelligence
- Enhance product development
- Improve customer retention
- Analyze the impact of an event (e.g. a product launch or redesign)
Insider threat detection

CHALLENGE

Costly data breaches may originate from external culprits or employees. But in many cases, an insider (employee, former employee, contractor, business associate, etc.) who has authorized access to valuable data can intentionally or inadvertently cause even greater damage. And as data volumes and accessibility grow, organizations – government and commercial, large and small – need to guard their data against these incidents.

Insider threats could involve fraud, the theft of confidential or commercially valuable information, the theft of intellectual property, the sabotage of computer systems, or the disclosure of information damaging to a company’s brand or reputation.

However, when organizations have large volumes of both structured data (documents, spreadsheets, transaction records) and unstructured data (social media, emails, voice recordings, notes), it becomes very difficult to conduct accurate data classification, monitoring, and analysis. Integrating search, analytics, and NLP can help to solve this challenge.

ON THE RISE – Insider threats account for nearly 75 percent of security breach incidents.

According to the Ponemon Institute’s report, “2018 Cost of Insider Threats: Global Organizations,” the average cost of an insider threat annually is $8.76 million.

SOLUTION

With an effective insider threat detection solution in place, organizations can avoid the lengthy legal battles – as well as monetary and reputational losses – resulting from insider trading, non-compliance, leaks of trade secrets, data breaches, and government intelligence leaks.

This is where integrating NLP-based insider threat applications can help determine if there is any illegal or nefarious intent within communications and detect threat patterns for rapid risk mitigation.

Source:
Solution features:

- Search engines can scale to billions of records
- Threat investigation using data available via the search engine
- Integration with third-party archiving solutions and analytics dashboards
- Greater flexibility and customization for in-depth analysis and reporting
- A 360-degree view across organizational data from multiple sources

BENEFITS

- Mitigate risk by detecting patterns and identifying red flags
- Keep a better pulse on the organization by customizing a “risk” score for any employee
- Faster time to incident and response with comprehensive enterprise visibility
- Deeper insights with a full complement of analytics frameworks for all threat detection workloads
NLP in your enterprise

Analyzing structured data alone is no longer enough. Sophisticated business analyses, predictions, and decision making all need more. The use cases we’ve provided show how unstructured content can be used to unlock tremendous new insights. With a well-implemented NLP solution in place, organizations can enable a deeper understanding of unstructured content, providing enhanced BI and analytics.

As your organization starts designing and building your NLP applications, it’s essential to ensure that your IT staff and/or implementation partners have the bandwidth and expertise required to conduct a thorough assessment for aligning NLP technologies with your business objectives. Whether your organization is only in the initial phase of evaluating various NLP technologies or has already identified the preferred solution for implementation, we can help – from evaluating and selecting the best-suited solution to implementing, tuning, and managing the application.

Looking for support and want to leverage Accenture’s NLP expertise? Contact us to discuss your goals and start defining a strategy and implementation roadmap for your NLP applications.
About Search & Content Analytics

Search & Content Analytics, formerly Search Technologies, is part of Accenture Applied Intelligence. We live in a data-driven world. But not everyone is making the most of their data. 80 percent of all data is unstructured – imagine the hidden insights trapped within unstructured enterprise content such as voice, images and emails. At Search & Content Analytics, our mission is to help enterprises unlock the full value within their unstructured and structured data. We combine innovative technologies such as machine learning and natural language processing with search and big data analytics to transform the way people work. Whether it’s improving intranet and website search, monitoring internal communications to detect insider threats, helping recruiters match jobs to résumés, analyzing oil wellhead reports, or exploring molecular data, we bring comprehensive search and analytics services to clients across industries. Clients include organizations in e-commerce, media, healthcare, financial services, recruiting, manufacturing, and the government sector.

What knowledge and insights are trapped in your data? Let us help you find better answers.

Visit our [webpage](#) to learn more about our capabilities and client work.

About Accenture

Accenture is a leading global professional services company, providing a broad range of services and solutions in strategy, consulting, digital, technology and operations. Combining unmatched experience and specialized skills across more than 40 industries and all business functions – underpinned by the world’s largest delivery network – Accenture works at the intersection of business and technology to help clients improve their performance and create sustainable value for their stakeholders. With more than 469,000 people serving clients in more than 120 countries, Accenture drives innovation to improve the way the world works and lives.

Visit us at [www.accenture.com](http://www.accenture.com).