

VERSION NEXT

NOW

TEKSYSTEMS NATURAL LANGUAGE PROCESSING ISSUE
FALL 2024



PAPERLESS PRECISION

UNLEASHING NLP IN BUSINESS MANAGEMENT

Paperless Precision:

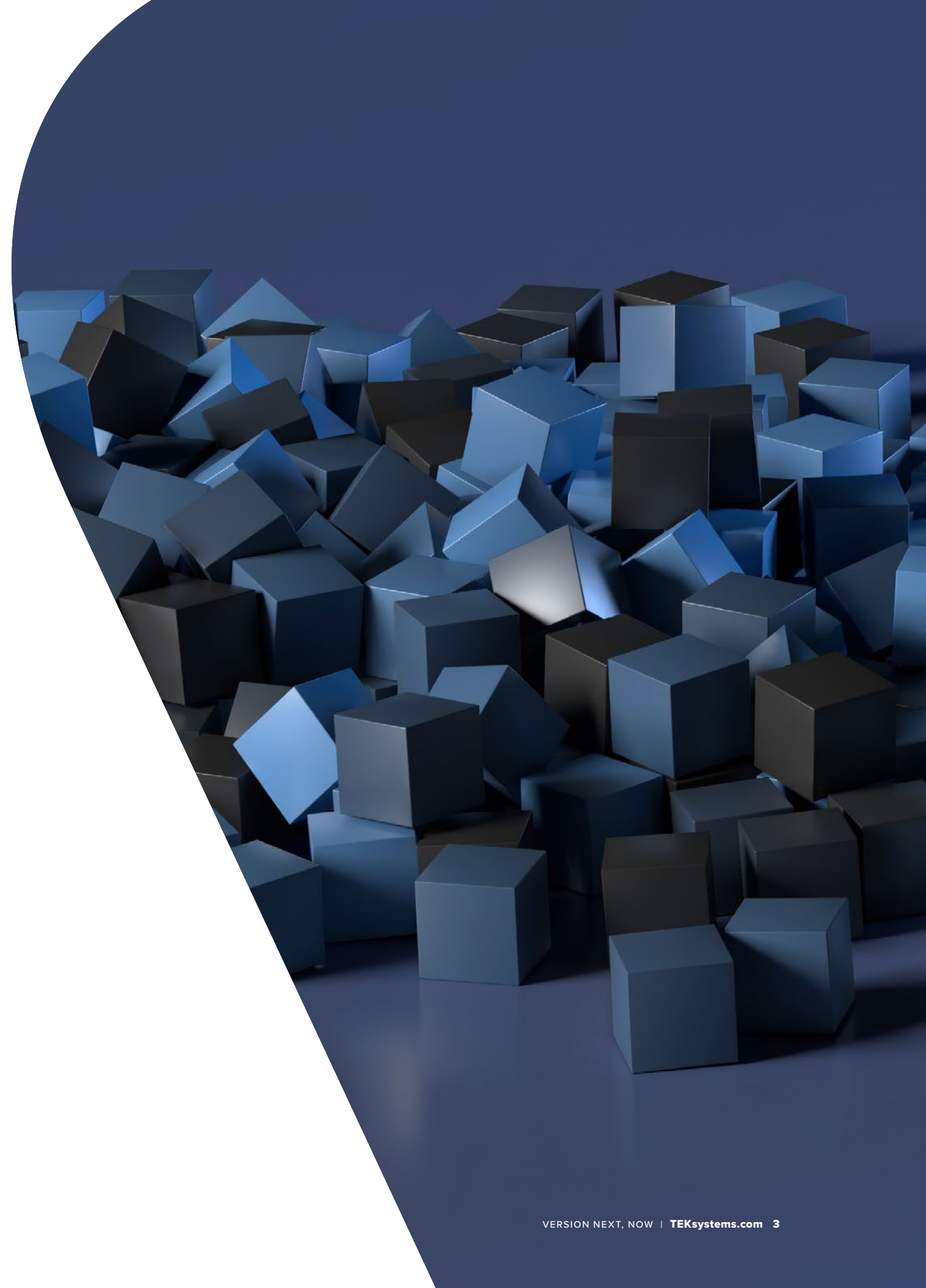
Unleashing NLP in Business Management

Paperwork—the bane of most knowledge workers' existence. But with the advent of natural language processing (NLP), it doesn't have to be such a tedious chore. NLP, a subfield of artificial intelligence, focuses on the interaction between computers and human language. By leveraging sophisticated algorithms and vast amounts of data, NLP enables machines to comprehend, interpret and generate human language in a meaningful and useful way—this is what digital transformation is all about.

This convergence of NLP with Intelligent Document Processing (IDP) uses advanced technologies to automate the extraction, interpretation and processing of data from various document types across an organization, and it's especially useful in areas like healthcare, finance, legal and technology. NLP significantly enhances the capabilities of IDP solutions by improving text extraction, understanding, data extraction, document classification, search and retrieval, risk management, workflow automation, and user experience. It really has the potential to completely revolutionize the gathering, processing and analysis of data. Of course, there are still some challenges when it comes to language differences, data training, ambiguity, biases and more, but the benefits far outweigh these potential drawbacks. Addressing these challenges requires ongoing research, advanced techniques and comprehensive strategies to ensure that NLP systems are accurate, fair and robust to get the most out of IDP. But if you're willing to accept those challenges, with the right approach, these limitations can be addressed so you can make paperwork a thing of the past—or at least less dreadful.



Sharon Florentine
Contributing Editor



01

THE CHANGE AGENT

In an era where efficiency and accuracy are paramount, businesses are constantly seeking ways to streamline operations and reduce manual workloads. Increase productivity and revolutionize modern business operations with natural language processing (NLP).

[PAGE 7 >>>](#)

02

MARKET PERSPECTIVE

Amit V. Singh, AWS global head of GTM, generative AI and machine learning partnerships, shares his point of view on Intelligent Document Processing (IDP), highlighting IDP's AI-driven efficiency in automating document tasks, AWS's role in enhancing IDP with machine learning, and the diverse industries benefiting from these advanced solutions.

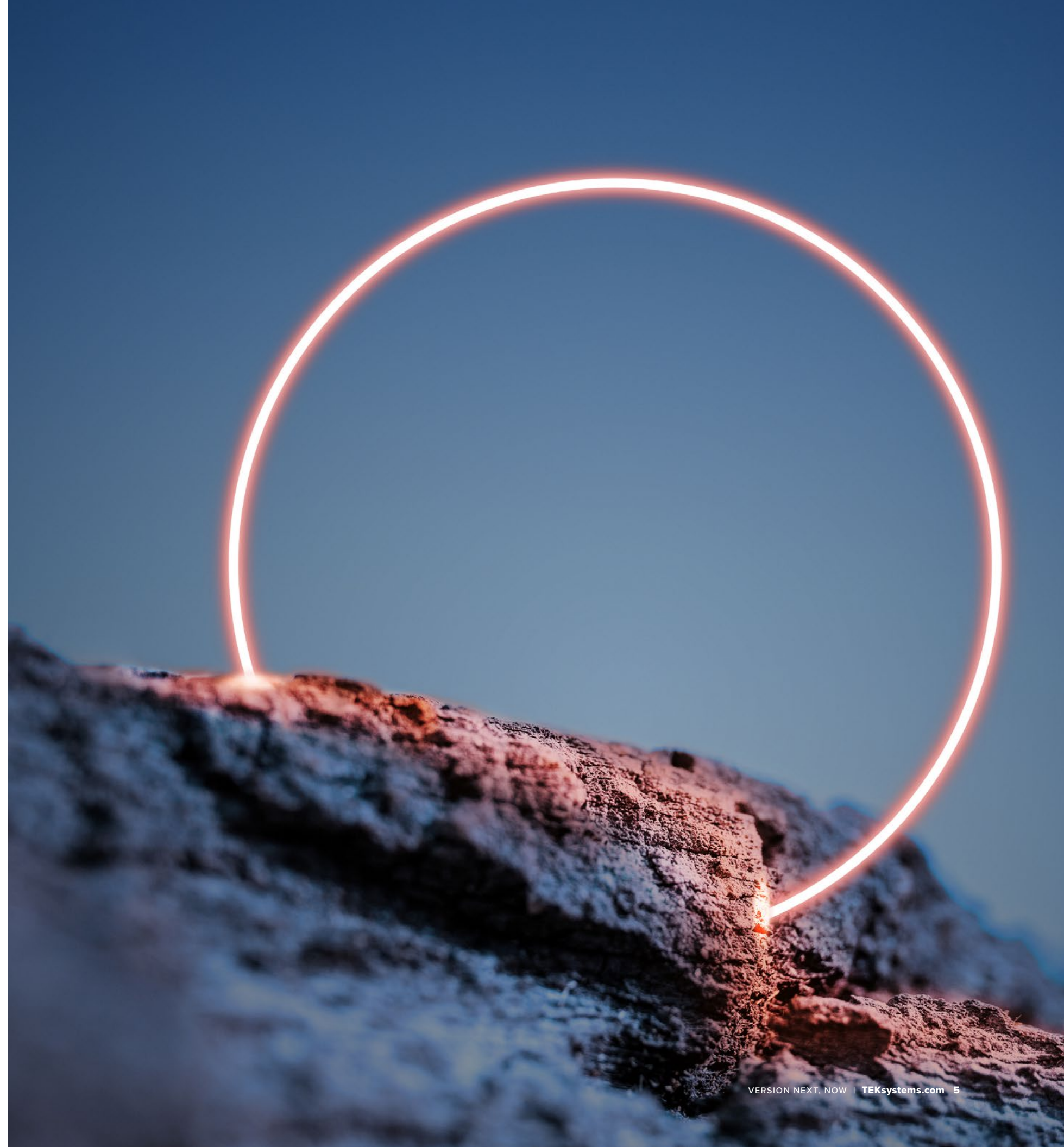
[PAGE 13 >>>](#)

03

TEKSYSTEMS' PERSPECTIVE

By integrating NLP into IDP solutions, organizations achieve higher accuracy, efficiency and scalability in document processing, leading to significant cost savings, improved compliance and better decision-making.

[PAGE 19 >>>](#)





THE CHANGE AGENT

In an era where efficiency and accuracy are paramount, businesses are constantly seeking ways to streamline operations and reduce manual workloads. Increase productivity and revolutionize modern business operations with natural language processing (NLP) and generative AI.

01

A New Era in Document Management

Imagine a world where mountains of paperwork are handled effortlessly, where contracts, reports and emails are meticulously organized and understood in mere seconds. In this world of advanced document management and generative AI, businesses operate with unprecedented efficiency thanks to the seamless integration of cutting-edge technology. This digital transformation is not the stuff of science fiction, but a present-day reality made possible by the advancements in natural language processing (NLP).

NLP, a subfield of artificial intelligence, focuses on the interaction between computers and human language. By leveraging sophisticated algorithms and vast amounts of data, NLP enables machines to comprehend, interpret and generate human language in a way that is meaningful and useful.

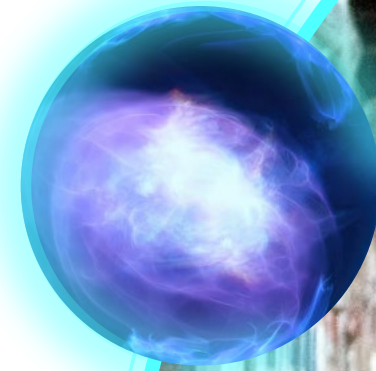
From automating customer service responses to analyzing vast troves of legal documents, natural language processing use cases, enhanced by generative AI, are revolutionizing business operations across various industries, making tasks that once seemed insurmountable now manageable with the click of a button.

As we delve deeper into the NLP applications and capabilities, we uncover the true potential of this transformative technology. Traditionally, document management used to involve time-consuming manual processes that were prone to human error. Employees had to sort, categorize and file documents by hand, often leading to inefficiencies and bottlenecks. Searching for specific information within these documents was akin to finding a needle in a haystack, and the physical storage of paper documents required considerable space and resources.

In contrast, modern Intelligent Document Processing (IDP) solutions, powered by NLP and enhanced by generative AI, have redefined document management for businesses. IDP automates the extraction, classification and validation of information from a variety of document types, both structured and unstructured. These systems can swiftly process invoices, contracts and forms, accurately pulling relevant data and integrating it into digital workflows.

The shift from traditional methods to IDP solutions exemplifies the profound impact of technology on modern business operations. By adopting IDP, organizations can streamline their workflows, reduce operational costs and allocate human resources to more strategic tasks, ultimately driving productivity and innovation.

The global NLP market is projected to increase from **\$36.42 billion** in 2024 to **\$156.80 billion** by 2030, reflecting a compound annual growth rate (CAGR) of **27.5%**.



HOW DOES NLP WORK?

NLP works by combining computational linguistics with machine learning and deep learning techniques. It is a transformative technology that enhances the ability of machines to understand and interact with human language. Its integration into Intelligent Document Processing significantly improves efficiency and accuracy in handling large volumes of documents, enabling businesses to automate and streamline various processes.

The process typically involves several key steps:

1. Text preprocessing

- Tokenization: splitting text into words, phrases or symbols
- Normalization: converting text to a standard format (e.g., lowercasing, removing punctuation)
- Stop words removal: removing common words that may not be useful for analysis (e.g., “and” or “the”)
- Stemming and lemmatization: reducing words to their root forms

2. Feature extraction

- Bag of words: representing text by the frequency of each word
- TF-IDF (term frequency-inverse document frequency): weighing the importance of a word in a document relative to a collection of documents
- Word embeddings: representing words in continuous vector space (e.g., Word2Vec, GloVE)

3. Model training

- Supervised learning: training models on labeled data for tasks like sentiment analysis or named entity recognition
- Unsupervised learning: extracting patterns from unlabeled data, such as topic modeling

3. Evaluation

- Assessing model performance using metrics like accuracy, precision, recall and F1 score

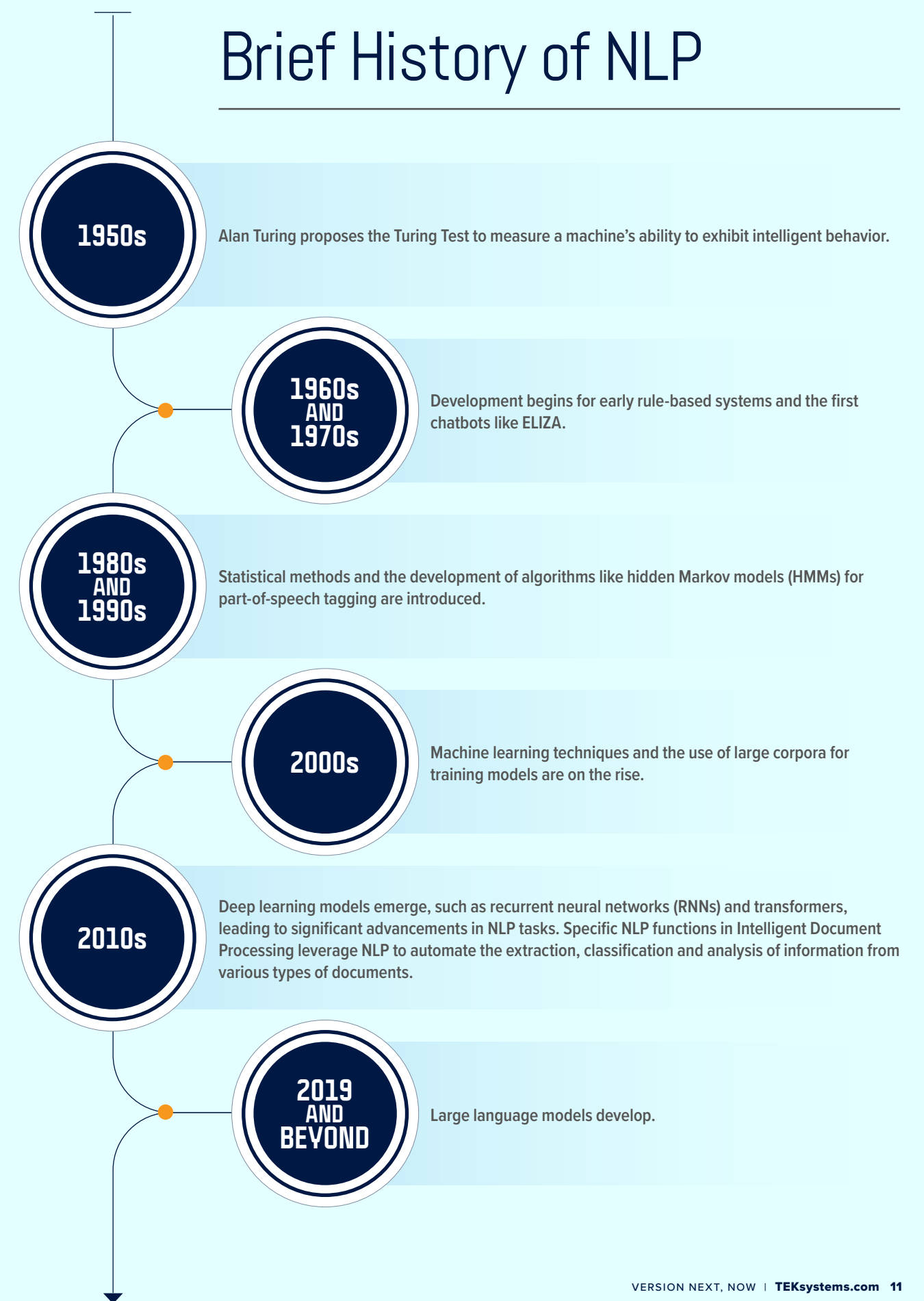
4. Deployment

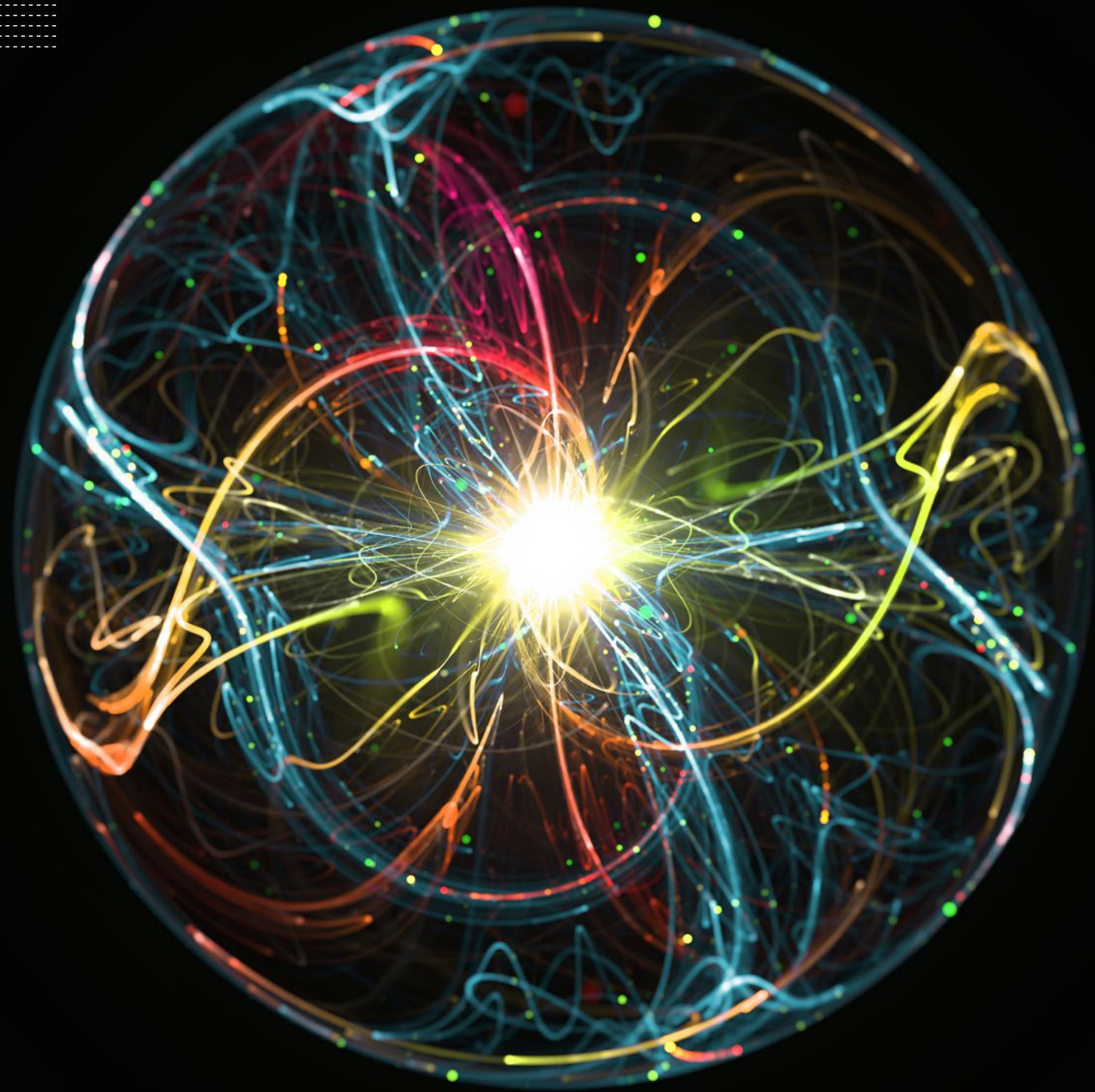
- Implementing the model in applications such as chatbots, search engines and recommendation systems

To effectively reap the benefits of natural language processing, organizations need a robust data strategy that ensures the availability, quality and relevance of data.

Organizations must think about how to combine the best of both worlds, providing specific information about guidelines and expectations for a particular domain or industry rather than just generic responses. This is where NLP, fueled by Gen AI, becomes a differentiator.

Brief History of NLP





MARKET PERSPECTIVE

Amit V. Singh, AWS global head of GTM, generative AI and machine learning partnerships, shares his point of view on Intelligent Document Processing (IDP), highlighting IDP's AI-driven efficiency in automating document tasks, AWS's role in enhancing IDP with machine learning, and the diverse industries benefiting from these advanced solutions.

02

AI-Driven Efficiency in Document Processing: Insights From AWS



Can you explain what Intelligent Document Processing (IDP) is and how it differs from traditional document processing methods?



Amit V. Singh (AWS): [Intelligent Document Processing](#) (IDP) is a widely recognized industry term used across various sectors to accelerate paper-based or digital business processes and minimize costs by leveraging AI for automatic extraction, classification and analysis of information with high precision.



How does AWS leverage machine learning and artificial intelligence to enhance IDP solutions?



Amit V. Singh (AWS): IDP utilizes a combination of AWS services, including Amazon Textract, Amazon Comprehend, Amazon Comprehend Medical and Amazon Bedrock. Amazon IDP is powered by a combination of pretrained AWS AI services APIs, providing industry-leading accuracy for extracting data from unstructured and semi-structured documents such as forms, notes, invoices, receipts and tables.

“Intelligent Document Processing (IDP) is a crucial technology used across various industry sectors to streamline the processing of a large volume of documents, often reaching millions per year.”



Amit V. Singh
AWS



What are some common use cases and industries that benefit the most from IDP solutions provided by AWS?



Amit V. Singh (AWS): IDP is a crucial technology used across various industry sectors to streamline the processing of a large volume of documents, often reaching millions per year. This technology finds significant application in sectors such as financial services, insurance, the public sector, healthcare, life sciences, legal and manufacturing industries. AWS IDP places particular emphasis on key focus industries including financial services (i.e., mortgage and insurance sectors), as well as the public sector and healthcare. Some specific use cases include [the following]:

- **A leading mortgage lender** with over 4 million lifetime customers offers a wide range of mortgage options. They reduced document processing time from hours to minutes using AWS IDP.
- **A financial services-oriented technology company** for mortgage and home equity lending and services achieved an accuracy of more than 90% leveraging AWS IDP.
- **A major health insurance company**, which provides health coverage for over 47 million members through affiliated health plans, utilized AWS IDP to automate 90% of their document classification for claims processing.



OUR PERSPECTIVE

By integrating NLP into IDP solutions, organizations achieve higher accuracy, efficiency and scalability in document processing, leading to significant cost savings, improved compliance and better decision-making.

03

The Synergy Between NLP and IDP

IDP uses advanced technologies to automate data extraction, interpretation and processing from various document types. **NLP** significantly enhances IDP by improving text recognition, understanding, data extraction, document classification, search and retrieval, risk management, workflow automation, and user experience.

NLP enhances Optical Character Recognition (OCR) for better text recognition and correction, especially in noisy or low-quality scanned documents. Beyond text extraction, NLP provides semantic understanding, allowing systems to grasp content meaning. Techniques like named entity recognition (NER) identify and classify entities within documents.

Automated data extraction, powered by NLP, accurately pulls key data points from documents like invoices, contracts, forms and emails, reducing manual effort and errors. NLP also automates document classification into predefined categories, improving organization and routing within workflows. Clustering and tagging enhance document management by grouping similar documents and adding relevant keywords for easier search and retrieval.

Enhanced search capabilities, driven by contextual and semantic search, allow users to perform more intuitive and accurate searches, particularly useful in domains like legal, medical and technical documentation. NLP improves data normalization and standardization, ensuring consistent formatting and enriching extracted data by linking related information.

NLP enhances compliance and risk management by identifying and extracting information required for compliance reporting, auditing documents for standards, and flagging noncompliant content. It automates end-to-end document processing workflows, integrating with systems like ERP, CRM or RPA platforms, reducing manual intervention and increasing efficiency. Intelligent routing capabilities direct processed documents based on content and context.

Lastly, NLP enhances user experience by providing intuitive interfaces for interacting with document processing systems through natural language commands and queries, making systems more user-friendly and accessible.

The convergence of NLP with IDP uses advanced technologies to automate the extraction, interpretation and processing of data from various document types across an organization. Generative AI enhances these capabilities further by interpreting complex documents such as doctors' notes and technical diagrams, providing summaries and insights that older AI models couldn't achieve.



Industries most in-demand for NLP are financial services and healthcare.

IDP—a Gen AI solution—will be a game changer.

CHALLENGES OF NLP-DRIVEN IDP

NLP faces challenges like language differences, data training, ambiguities, misspellings, bias and phrases with multiple meanings. These challenges impact its effectiveness and efficiency.

Languages vary in syntax, grammar and semantics, making it difficult for NLP systems to perform well across multiple languages. Training NLP models requires large volumes of high-quality, annotated data, which is resource-intensive and time-consuming. The quality of training data directly impacts model performance, with inaccuracies or biases leading to poor results.

Developing NLP systems involves complex processes requiring expertise and computational resources. Rapid advancements in NLP mean keeping up with the latest techniques can be challenging. Natural language is inherently ambiguous, with words and sentences often having multiple interpretations, requiring sophisticated contextual understanding.

Misspellings and typographical errors in real-world text data challenge NLP models' accuracy. Bias in NLP models, learned from training data, can lead to unfair outcomes, requiring careful selection and preprocessing of training data. Many words and phrases have multiple meanings depending on context, necessitating large amounts of contextual data and sophisticated modeling techniques for effective disambiguation.

Addressing these challenges requires ongoing research, advanced techniques and comprehensive strategies to ensure NLP systems are accurate, fair and robust.

BENEFITS OF NLP-DRIVEN IDP

NLP offers numerous benefits in business operations, customer interactions and data analysis. One primary benefit is enhanced customer service. NLP-powered chatbots and virtual assistants provide 24/7 support, handling routine queries and freeing human agents for complex issues, resulting in faster response times and higher customer satisfaction. Sentiment analysis tools monitor customer feedback in real time, allowing prompt issue resolution.

NLP improves efficiency and automation by handling repetitive tasks like data entry, document processing and information extraction, minimizing errors and increasing operational efficiency. In finance, it automates information extraction from financial documents, speeding up loan approvals and compliance checks. In the legal field, NLP streamlines contract analysis and legal research.

NLP enhances data analysis by extracting insights from unstructured text data, informing strategic decisions, improving marketing campaigns and providing deeper customer understanding. It helps organizations maintain a competitive edge by monitoring and analyzing competitor activities and market trends.

NLP contributes to better decision-making and knowledge management by improving search result accuracy, aiding quick information retrieval and enhancing productivity. Intelligent document processing solutions categorize and tag documents, making retrieval efficient. This is beneficial in fields like healthcare, improving access to relevant medical literature and patient records.

Despite challenges, implementing generative AI and NLP in IDP solutions offers significant benefits, including higher accuracy, efficiency and scalability, leading to cost savings, improved compliance and better decision-making.

Real-World Application:

McDonald's

McDonald's is leveraging NLP as part of their data-driven transformation to stay ahead in the competitive quick-service restaurant industry. Craig Brabec, former McDonald's chief data analytics officer, highlights the critical role of data in enhancing customer and crew experiences. The "Accelerating the Arches" growth plan includes initiatives like Automated Order Taking (AOT) technology, developed in partnership with IBM.

NLP is at the core of the AOT system, enabling it to automate customer interactions by accurately processing spoken orders. This technology improves order accuracy and speed, allowing employees to focus on other responsibilities. By minimizing human intervention, AOT reduces errors and enhances the efficiency of the ordering process.

The adoption of NLP and AI in AOT is part of McDonald's broader digital transformation strategy, which saw over \$18 billion in digital channel sales in 2021, a significant increase from the previous year. This shift underscores McDonald's commitment to integrating advanced technologies to create personalized and efficient customer experiences.







Brabec emphasizes the importance of aligning data initiatives with business performance metrics. McDonald's use of NLP and other data-driven technologies aims to deliver superior customer service and operational efficiency, ensuring that every interaction is tailored to individual preferences. This approach not only enhances the customer journey but also demonstrates the real impact of technological investments on the company's strategic objectives.

All information shared herein was accessed from public sources as indicated and is not indicative that the named entity is a TEKsystems client nor that the work was performed by TEKsystems.



TEKsystems' Tips

Best Practices for Implementing NLP

-  **Identify business needs and goals:** Determine specific areas where NLP can add value, such as customer service, market analysis or operational efficiency. Define clear objectives, such as improving customer satisfaction, gaining insights from social media or automating repetitive tasks.
-  **Collect and prepare data:** Gather relevant textual data from sources like customer reviews, emails, social media and internal documents. Clean and preprocess the data by removing noise, normalizing text and tokenizing sentences and words.
-  **Choose the right NLP tools and techniques:** Select appropriate NLP techniques, such as sentiment analysis, entity recognition, machine translation or topic modeling, based on the identified goals.
-  **Build and train models:** Develop and train NLP models using historical data. Fine-tune pretrained models for specific tasks to save time and resources. Ensure the models are trained in diverse and representative data to avoid biases.
-  **Integrate and deploy:** Integrate NLP models into existing systems, such as customer service platforms, CRM systems or data analysis tools. Deploy models in a scalable and secure environment, ensuring they can handle real-time data and large volumes of text.
-  **Monitor and improve:** Continuously monitor the performance of NLP applications to ensure they meet business objectives. Collect feedback and update models to improve accuracy and relevance. Stay updated with advancements in NLP to leverage new techniques and tools.

By systematically implementing NLP and focusing on these applications, organizations can significantly improve efficiency, decision-making and the customer experience.

TEKsystems Portfolio

- As an [AWS Premier Tier Services Partner](#), TEKsystems covers the full spectrum of Amazon Web Services (AWS) initiatives. From design, migration and implementation to adoption and improvement, continuous integration and delivery (CI/CD), infrastructure as code, Lean-Agile, and more—we're there.
- As a [Google Cloud Premier Partner](#), we support the full spectrum of delivering Google Cloud initiatives, from design, migration and implementation to adoption and improvement, covering CI/CD, infrastructure as code, Lean-Agile, data analytics, AI, machine learning and Gen AI.
- As a [Microsoft Solutions Partner](#), we bring qualified expertise and deep experience to help you maximize ROI and achieve real value. From discovery and design to adoption and improvement—we'll tailor our solutions to meet your needs and help you stay ahead of what's next.
- As a [Red Hat Premier Business Partner](#), we provide qualified technical leadership, open source expertise and scale to help you get the most out of your Red Hat products—no matter where you are in your modernization journey.
- With 30-plus SnowPro certified architects, our [Snowflake Elite Partner](#) status highlights our proven skills and experience to help you leverage Snowflake's innovative technology and achieve data-driven results.
- As a [ServiceNow Elite Partner](#), we bring experience and subject matter know-how to help you drive your ServiceNow initiatives. From implementation to optimization—we'll tailor our services to help you stay ahead of the curve and accelerate the adoption of ServiceNow solutions.
- [In Good Company](#)
Transformational technologies demand equally transformative partnerships. We offer full-stack capabilities coupled with depth and diversity of experience in leading platforms that help organizations grow, innovate and thrive.

The views and opinions expressed in this publication are those of the authors and do not necessarily reflect the views of TEKsystems, Inc. or its related entities.

Meet Our Contributors



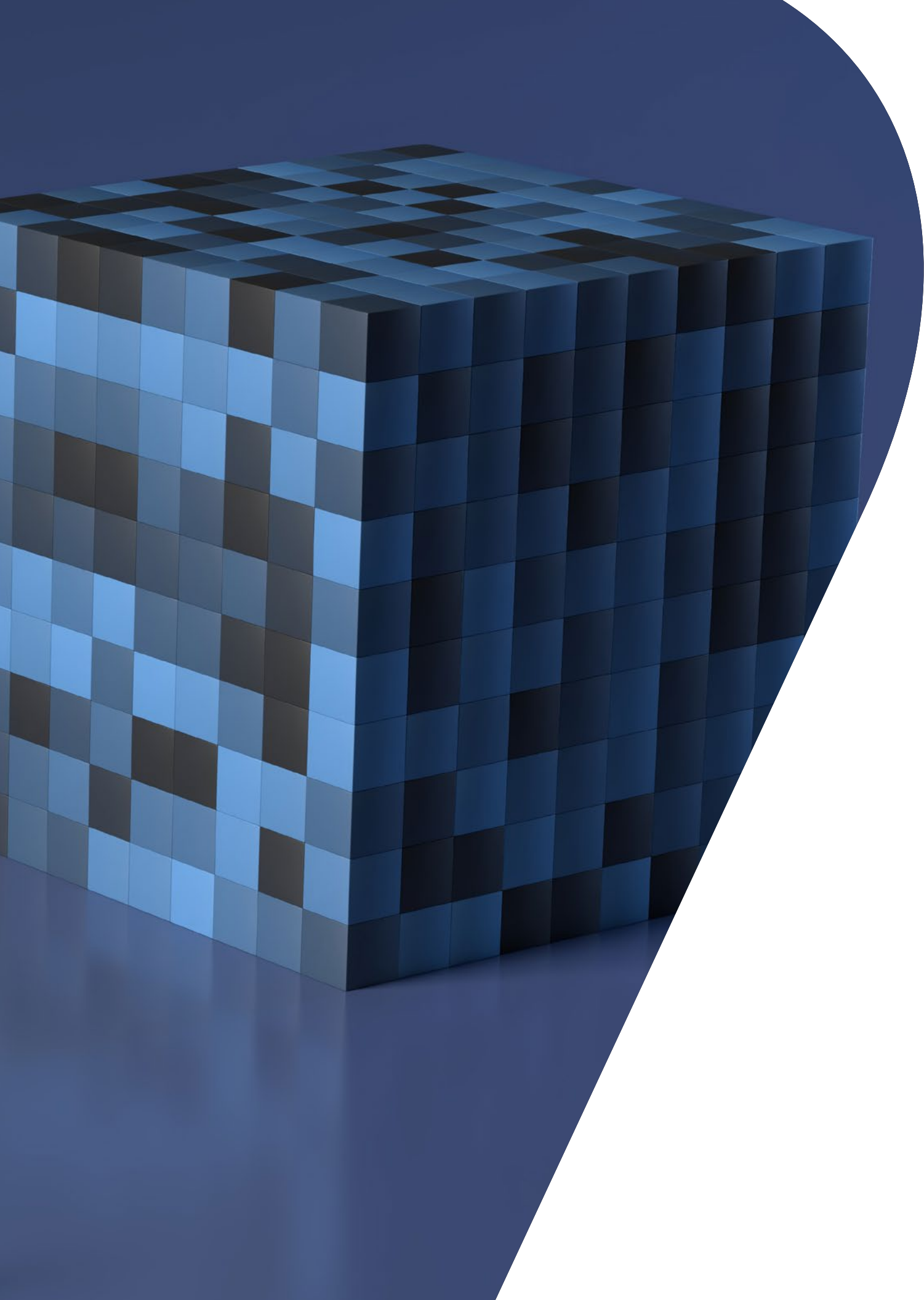
Srinivasan Swaminatha

Managing Director, TEKsystems Global Services



Amit Singh

Global Head of GTM, Generative AI and ML Partnerships, AWS



About TEKsystems®



We're TEKsystems. We accelerate business transformation for our customers. We bring real-world expertise to solve complex technology, business and talent challenges—at global scale. We're a team of 80,000 strong, working with over 6,000 customers, including 80% of the Fortune 500 across North America, Europe and Asia, who partner with us for our scale, full-stack capabilities and speed. We're strategic thinkers, hands-on collaborators, helping customers capitalize on change. We're building tomorrow by delivering business outcomes and driving positive impacts in our global communities. TEKsystems is an Allegis Group company [TEKsystems.com](https://www.teksystems.com).

Sharon Florentine, Contributing Editor



[Sharon Florentine](#) is the contributing editor for Version Next, Now, TEKsystems' quarterly publication. She is an award-winning independent writer and editor with more than 20 years of experience in the tech industry. Her work has appeared in Computerworld, PC Magazine, CRN and eWEEK, among others, and she is a passionate advocate for equity, diversity and inclusion in tech and beyond. Most recently, Sharon was a senior writer for CIO.com, where she covered software development, Agile, IT careers, learning and development, and DEI. She lives near Philadelphia.

Listen Now



Don't miss TEKsystems' leaders on The Agile Brand podcast. In a special series, host, author and business expert Greg Kihlstrom sits down with leaders from TEKsystems to discuss how natural language processing enables IDP solutions.



Be in the Know

Check out previous issues and know what's next: [Version Next, Now](#).

Follow Us



Sources

1. [Natural Language Processing—Worldwide](#), Statista

