cognaize™

ELEVATE YOUR FINANCIAL OPERATIONS WITH AI

Transforming Financial Document Processing with Hybrid Intelligence

Achieve Exceptional Accuracy, Efficiency, & Regulatory Compliance with Cognaize's AI Solutions

Executive Summary

In the rapidly evolving landscape of artificial intelligence, innovations like large language models, exemplified by ChatGPT, have redefined the boundaries of human-computer interaction with their emergent capabilities. However, within the finance sector, the real revolution is spearheaded by 'Hybrid Intelligence.' This innovative approach harmonizes AI's computational might with human insight, facilitating nuanced decision-making processes.

Financial institutions have long led the charge in the automation of document processing, quickly adopting AI technologies. Yet, they face persistent challenges in processing unstructured data due to the diversity of document types and formats, intricate details within various financial statements, and complex data management and privacy requirements.

To leverage their technological investments for significant financial returns, these institutions must look beyond simple automation. They need to embrace a transformative journey that includes both learning and organizational adaptation. This entails a collaborative synergy between human expertise and artificial intelligence, allowing for flexible interactions that adapt to changing conditions and requirements.

This whitepaper delves into:

- Common challenges encountered by our clients
- A case study showcasing a 66% reduction in operational costs for a US bank using our AI solutions
- Strategies for implementing AI in document annotation, advanced table detection, and data extraction for effective downstream applications
- Insights into the technological foundation of our solutions

Hybrid Intelligence as long-term competitive advantage

Banks still struggle with automation of unstructured data

Privacy becomes distinguishing factor

Introduction to Document Automation in the Financial Sector

The financial sector faces a significant challenge with the increasing volumes of unstructured data. This data, found in complex documents such as credit agreements and ESG reports, necessitates secure and compliant processing. Traditional methods, often manual and inefficient, struggle to extract and utilize relevant data from diverse formats for analysis and further downstream processing.

Artificial Intelligence emerges as a pivotal solution to these challenges, with Cognaize leading the charge. Our advanced AI models, pre-trained for a multitude of financial scenarios, redefine the standards of data automation. **Cognaize sets itself apart by:**

Speed

Rapid deployment through our AI-platform combined with the expertise of financial professionals, thus enhancing operational efficiency.

Deeper Insights

We provide profound insights that are critical for making informed decisions by constructing comprehensive knowledge graphs that extend beyond single documents.

Unparalleled Model Accuracy

Our unique approach integrates expert knowledge with AI in a meaningful way, ensuring that the AI is not only constantly learning but also delivering significant time and cost savings with unmatched accuracy.

Long-term Competitive Advantage for our Clients

Through on-premise installation clients retain ownership of their data and models.

Overview

Extraction and annotation of information from all documents specific to the financial industry, for example:

- Credit agreements
- SEC filings
- ESG reports
- Rent rolls
- Structured credit
- ISDA contracts
- Valuation reports

Cognaize's Distinct Advantages

- Rapid deployment
- Unparalleled accuracy
- Better insights
- Data ownership

Understanding AI-driven Document Automation

Our tailor-made solution for the financial domain, combines the power of 'Precise Language Models' and 'Advanced Layout Detection' to construct a context-rich 'Knowledge Graph'. This innovative approach harnesses the capabilities of language models to reason along the graph, delivering unparalleled transparency and explainability in a use-case-agnostic manner. Our transformative technology empowers financial institutions to navigate complex data landscapes with ease, unlocking valuable insights and driving informed decision-making.

This approach not only understands complex financial terminology but also respects privacy and the nuances of compliance while ensuring accurate, actionable insights, leaving the final decision to the financial experts through state-of-the-art user interfaces.

Profound Layout Understanding provides crucial context to language models

1. OCR: Fine-tuned models for challenging languages such as Chinese
2. Page Layout Understanding: Identifies sections, headers, captions and other structures that are key for comprehension through the language models downstream
3. Enhanced Table Detection: Detects tables including merged cells for table structure-based querying

Precise Language Models dedicated- and open domain- models specifically trained or fine-tuned for the financial industry

1. Dedicated High Performance Models Combined with Powerful Chat/ Instruct Models achieves highest efficiency for known tasks while allowing generic querying
2. Supervised Fine-Tuning and Reinforcement Learning from Human Feedback allows for better results with smaller models that can operate with less computational resources
3. Generative Insights: harness our precise language models for instant data comprehension and deeper insights

Hybrid Intelligence fostering meaningful cooperation between subject matter experts and AI – judgement layer, UI, interaction

1. Seamless Experience for Subject Matter Experts
2. Large Library of Use Case- Dedicated Applications
3. Creating Hybrid Intelligence through meaningful cooperation of experts and AI such as teaching mode or sporadic validation - depending on the best cooperation model for a specific task

Overview

+ Designed for the financial industry
+ Use-case-agnostic
+ User interfaces for AI-human expert interaction

Profound Layout Understanding

Precise Language Models

Hybrid Intelligence

Advancing AI Reasoning: The Synergy Between Knowledge Graphs and Language Models

By **Vahe Andonians**: Founder, CPO & CTO, Cognaize

The advent of transformers has revolutionized text vectorization, achieving unprecedented precision. These embeddings encapsulate profound semantic meanings, surpassing previous methodologies, and are why Large Language Models are so convincingly good at generating text.

Large Language Models further demonstrate reasoning capabilities, albeit with limitations; their depth of reasoning tends to diminish rapidly. However, integrating knowledge graphs with these vector embeddings can significantly enhance reasoning abilities. This synergy leverages the inherent semantic richness of embeddings and propels reasoning capabilities to unparalleled heights, marking a significant advancement in artificial intelligence.

In the finance sector, Large Language Models are predominantly utilized through Retrieval Augmented Generation, a method that infuses new, post-training knowledge into Large Language Models. This process involves encoding textual data, indexing it for efficient retrieval, encoding the query, and employing similarity algorithms to fetch relevant passages. These retrieved passages are then used with the query, serving as a foundation for the Large Language Model to generate the response.

This approach significantly expands the knowledge base of Large Language Models, making it invaluable for financial analysis and decision-making. While Retrieval Augmented Generation marks a significant advancement, it has limitations.

A critical shortcoming lies in the passage vectors' possible inability to fully grasp the semantic intent of queries, leading to the vital context being overlooked. This oversight occurs because embeddings might not capture certain inferential connections essential for understanding the query's full scope.

Moreover, condensing complex passages into single vectors can result in the loss of nuances, obscuring key details distributed across sentences.

Additionally, the matching process treats each passage separately, lacking a joint analysis mechanism that could connect disparate facts. This absence hinders the model's ability to aggregate information from multiple sources, often necessary for generating comprehensive and accurate responses requiring synthesizing information from various contexts.

Efforts to refine the Retrieval Augmented Generation framework abound, from optimizing chunk sizes to employing parent chunk retrievers, hypothetical question embeddings, and query rewriting. While these strategies present improvements, they don't lead to revolutionary outcome changes. An alternative approach is to bypass Retrieval Augmented Generation by expanding the context window, as seen with Google Gemini's leap to a 1 million token capacity. However, this introduces new challenges, including non-uniform attention across the expanded context and a substantial, often thousandfold, cost increase.

Incorporating knowledge graphs with dense vectors emerges as the most promising solution. While embeddings efficiently condense text of varying lengths into fixed-dimension vectors, enabling the identification of semantically similar phrases, they sometimes fall short in distinguishing critical nuances. For instance, "Cash and Due from Banks" and "Cash and Cash Equivalents" yield nearly identical vectors, suggesting a similarity that overlooks substantial differences. The latter includes interest-bearing entities like "Asset-Backed Securities" or "Money Market Funds," while "Due from Banks" refers to non interest-bearing deposits.

Advancing AI Reasoning: The Synergy Between Knowledge Graphs and Language Models (cont)

By **Vahe Andonians**: Founder, CPO & CTO, Cognaize

Knowledge graphs also capture the complex interrelations of concepts. This fosters a deeper contextual insight, underscoring additional distinct characteristics through connections between concepts.

For example, a US GAAP knowledge graph clearly defines the sum of "Cash and Cash Equivalents," "Interest Bearing Deposits in Banks," and "Due from Banks" as "Cash and Cash Equivalents."

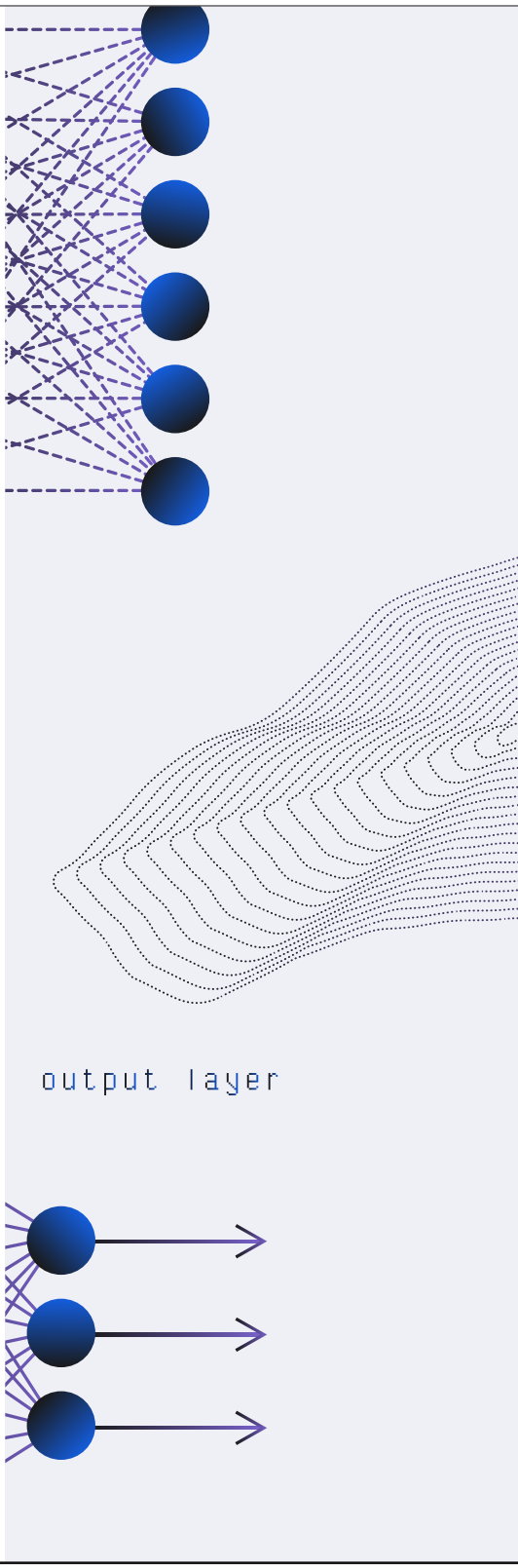
By integrating these detailed contextual cues and relationships, knowledge graphs significantly enhance the reasoning capabilities of Large Language Models. They enable more precise multi-hop reasoning within a single graph and facilitate joint reasoning across multiple graphs.

Moreover, this approach offers a level of explainability that addresses another critical challenge of Large Language Models. The transparency in how conclusions are derived through visible, logical connections within knowledge graphs provides a much-needed layer of interpretability, making the reasoning process not only more sophisticated but also accessible and justifiable.

The fusion of knowledge graphs and embeddings heralds a transformative era in AI, transcending the limitations of individual approaches to achieve a semblance of human-like linguistic intelligence. Knowledge graphs introduce previously gained symbolic logic and intricate relationships from humans, enhancing the neural networks' pattern recognition prowess and finally resulting in superior hybrid intelligence.

Hybrid intelligence paves the way for AI that not only articulates eloquently but also comprehends deeply, enabling advanced conversational agents, discerning recommendation engines, and insightful search systems.

Despite challenges in knowledge graph construction and noise management, the integration of symbolic and neural methodologies promises a future of explainable, sophisticated language AI, unlocking unprecedented capabilities.



CASE STUDY

AI-driven Credit Automation in a US Bank

A leading US bank aimed to transform its credit lending operations by automating the analysis of unstructured data, with the ultimate goal of enhancing risk assessment models and boosting operational efficiency, all while reducing costs.

Solution

The bank adopted Cognaize's AI Platform as a private Software-as-a Service (SaaS) solution, designed for efficient processing and analysis of vast amounts of unstructured data. The platform's 'Enhanced Layout Understanding' capability enables it to interpret complex financial documents such as credit agreements with remarkable accuracy and speed.

Outcomes

Cost Savings

The bank achieved a remarkable 66% reduction in operational expenses, significantly exceeding the initial target of a 25% reduction.

Processing Capacity

The bank's data processing capacity expanded by 150%, showcasing the platform's ability to handle increased workloads without compromising quality.

Data Accuracy & Customer Satisfaction

The bank observed substantial improvements in data accuracy and client satisfaction through faster processing, which in turn, strengthened its overall business strategy.

Conclusion

The strategic implementation of Cognaize's AI Platform resulted in considerable cost savings, enhanced processing capabilities, and improved customer service, marking a pivotal shift in the bank's credit lending operations.

Use Case

+ Automation of credit lending operations

Document Types

+ Credit agreements and 10k reports among others

Objective

+ Enhance risk assessment models & increase efficiency

66% Cost Savings

Capacity Increase

Accuracy Improvement

Better Decisions

Implementation of AI-driven Document Automation Projects

Implementing AI-based document automation in the financial sector offers a significant opportunity to boost operational efficiency. However, this transition is often met with apprehensions and concerns. Financial institutions are tasked with navigating data privacy issues, ensuring regulatory compliance, and seamlessly integrating AI technologies into their existing legacy systems.

Moreover, there is an essential requirement to ensure these AI-driven solutions align with the bank's ethical standards and meet customer expectations as well as internal process requirements. In the digital banking landscape, where trust and integrity are paramount, the potential advantages of AI must be judiciously weighed against these critical considerations as well as adhere to high standards of integrating effectively with downstream systems.

Cognaize employs a structured six-step approach for client solutions:

Overview

+ Fine line between capturing AI efficiencies and managing data privacy and regulatory compliance
+ Scalable project approach to reduce implementation time
+ KPI-driven implementation
+ Downstream usability is key

1 Assessment & Planning

Evaluate current document processing workflows for AI integration

- Document types & workflows
- Data points need to be extracted
- Fully automated or with human interaction

2 Technology Selection

Choose AI technologies for document structures and languages

- Document structure to determine use of base models such as table detection / page layout detection
- Languages
- Precise Language Models

3 Model Development

Defining a pipeline tailored to specific data extraction needs

- Fine-tuning
- RAG optimization
- Customization based on annotated samples

4 User Training & Testing

- Definition of human and AI interaction
- Workflow set-up

5 Success Measurement

- Time: extraction per document
- Accuracy: F1 score/precision/recall or any custom metrics

6 Integration

Integrate Cognaize solution with existing downstream systems

AI Risks in Banking: Navigating the Digital Frontier

The adoption of AI in the banking sector marks a significant leap towards improving efficiency, customer experiences, and financial innovation, albeit with notable challenges. These include potential biases in AI decision-making, the risk of operational and reputational damage due to lack of human oversight, and the complexities of maintaining data integrity.

Furthermore, the opaque nature of AI systems raises issues with regulatory compliance and trust preservation, while the need for clear AI decision explanations and strong data security measures, particularly under GDPR, highlights the importance of robust AI and data protection strategies. As Cognaize advances AI document automation in finance, addressing these challenges with effective risk mitigation strategies becomes crucial for responsible and secure AI application.

Retaining Intellectual Property for Competitive & Security Edge

Downsized models for private cloud or on-premise installation preserve the intellectual property rights over AI models and their corresponding datasets is crucial for maintaining a bank's competitive edge and operational security. By retaining ownership of the IP, banks can ensure that their bespoke AI models, fine-tuned to their specific operational requirements and customer needs, remain exclusive and out of reach from competitors or unauthorized entities. This exclusivity not only bolsters competitive standing but also minimizes the risk of illicit exploitation or duplication of a bank's AI-powered capabilities.

Privacy By Design

'Privacy by Design' ensures GDPR compliance by promoting on-premise or private/hybrid cloud deployments, allowing banks to securely manage data within compliant jurisdictions. This enhances privacy, security, and enables GDPR-compliant AI application management and scaling.

Auditability

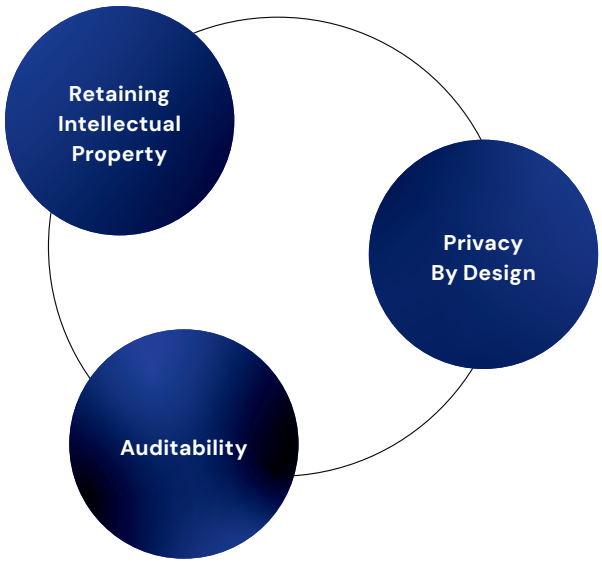
Ensuring AI systems are auditable is key to risk mitigation and compliance with the Federal Reserve's SR Letter 11-7 and Datenschutzkonferenz (German Data Protection Conference) guidelines. By adopting a transparent AI governance framework with thorough documentation, regular assessments, and clear AI decision tracing, banks can maintain regulatory compliance and build trust with customers and stakeholders.

Overview

- Banks as guardians of customer data
- Complex regulatory environment
- Pressure to improve efficiency

Risk Mitigation is Driven By:

- Intellectual property safeguarding
- Data Privacy by design
- Audit trail for AI solutions



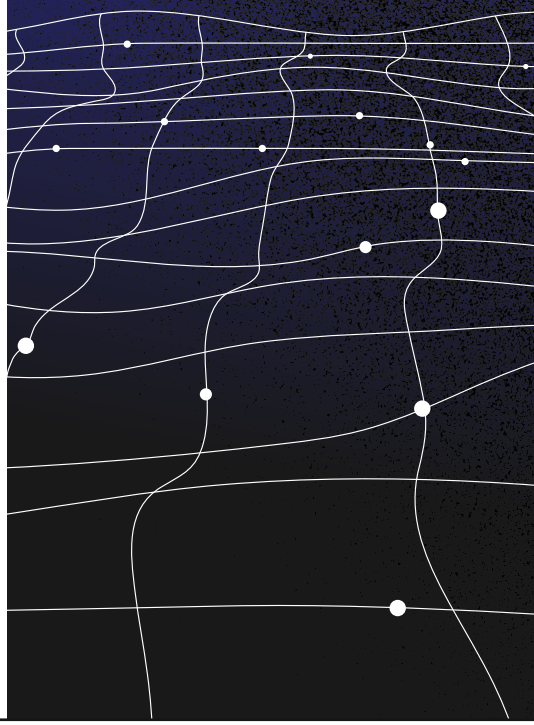
Conclusion

In conclusion, this paper has explored the pivotal role of Artificial Intelligence in transforming the financial sector, specifically through the lens of document automation. Emphasizing 'Hybrid Intelligence,' 'Profound Layout Understanding,' 'Precise Language Models,' and strategic 'Knowledge Graph' integration, Cognaize has established a unique vantage point.

This approach not only enhances data automation but also ensures compliance, privacy, and operational excellence in document processing. Our solutions, underscored by advanced AI models fine-tuned for the financial domain, offer unparalleled accuracy, speed, and scalability, facilitating a seamless transition towards digitization and automation.

Cognaize's commitment to privacy by design, intellectual property retention, and rigorous risk mitigation frameworks further solidify its position as a leader in transforming how financial services leverage AI. As we navigate the complexities of unstructured data and regulatory demands, our dedication to innovation, ethical AI use, and collaborative human-machine intelligence paves the way for a future where financial institutions can thrive with confidence and competitive edge.

Get in touch for a demo or a free consultation:
solutions@cognaize.com



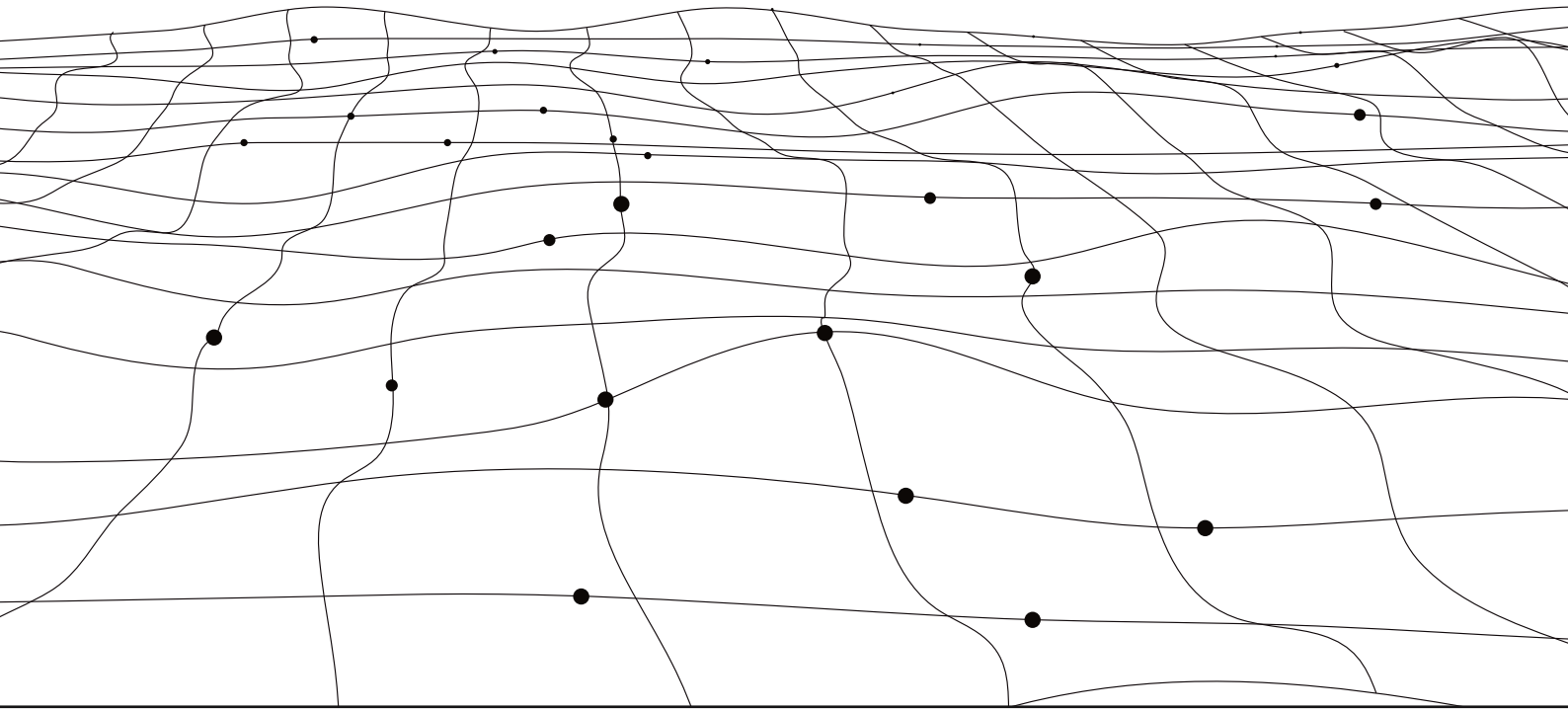
Cognaize Office Locations

115 Broadway
Floor 5
New York, NY 10006
USA

Taunusanlage 8
60329
Frankfurt am Main
Germany

41 Arshakunyats
0056
Yerevan
Armenia

611 North Brand Blvd
Floor 13
Glendale, CA 91203
USA



cognaize™

Automating Unstructured Data with Hybrid Intelligence

www.cognaize.com
solutions@cognaize.com