UniQreate - NEO xTract

Empowering domain experts to unlock the value of data in documents through AI and automation
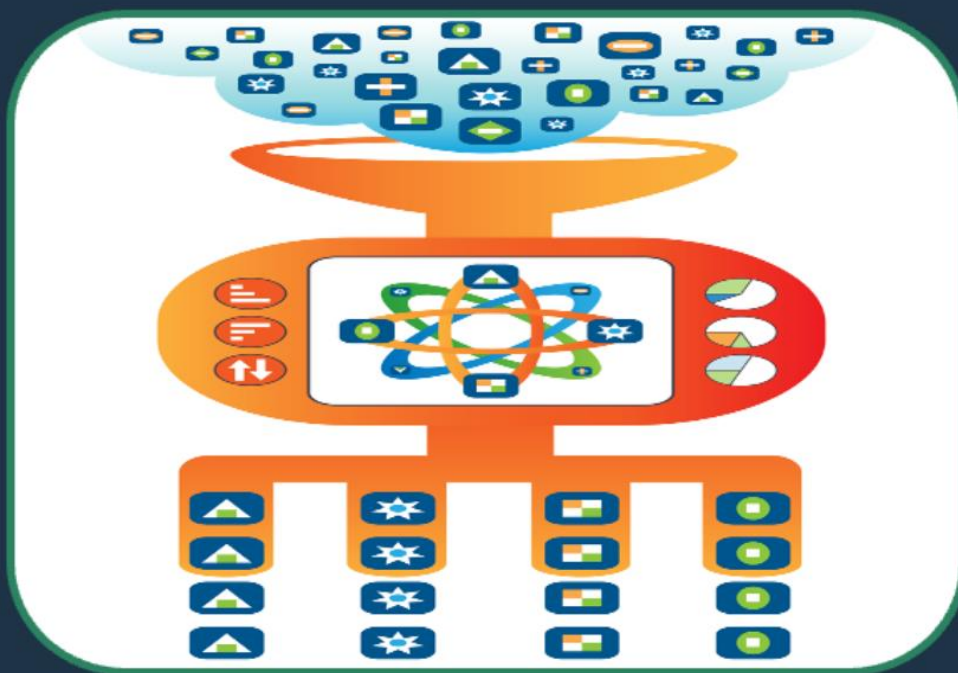
Jeremy Cox, Founder CX-Create
Sponsored by Oracle

Table of Contents

About this report

Based on interviews with CEO and co-founder Nitin Gupta, this report introduces innovative startup, UniQreate and its automated data extraction product, NEO xTract, that harvests critical data from unstructured sources such as PDF documents. It enables business domain experts such as analysts, asset managers, portfolio managers, operations teams, etc., to transform the data they need from raw and unstructured formats into a data-centric model integral for critical and timely investment advice to their clients.

The company, headquartered in New York, launched its **"SaaS first"** product, NEO xTract, in August 2022.

The theme for this month is around startups in the Finance sector and how they are innovating, changing the competitive landscape, and contributing to growth.

CX-Create is an independent IT industry analyst and advisory firm, and Oracle sponsors this report.

Highlights:

- Key drivers – domain experts know what data they need but struggle to get it quickly and efficiently.
- NEO xTract puts the power of data in the hands of the business users
- NEO xTract simplifies the complexity by reducing the process to three simple steps
- Oracle Cloud Infrastructure and Oracle support prove their value to UniQreate

The business context for UniQreate

Key drivers – Domain experts know what data they need but struggle to get it.

Domain experts such as data and market analysts, operations middle and back-office users, asset managers, data scientists, and others need a simple way to get relevant, accurate, and timely data from a vast array of documents from multiple sources. They know what information they need, can apply their expertise, and advise clients accordingly when they manage to get it. Asset managers, for example, may want to draw on non-financial and financial data from websites, news items, and social media, combining this with data from prospectuses, and annual reports, in their search for alpha, the ultimate investment.

High cost of 'document to data' lifecycle – in manual effort, time, accuracy, and missed opportunities

Unfortunately, data is rarely stored for convenient extraction, synthesis, and use. In many organizations, data is locked away in departmental silos and buried deep within large documents that come in different formats, both structured and often unstructured. It is not uncommon in large companies to rely on teams of researchers (internally or through a 3rd party such as a BPO/BPM) pouring over sizable documents and manually extracting relevant data and transposing it into spreadsheets for analysis. This approach takes considerable time and effort and is prone to inaccuracies. The researchers may lack business domain expertise and inadvertently overlook critical data. Decision-makers and advisors then have a problem trusting the results, causing decision delays and missing time-sensitive opportunities, and sullying the firm's reputation. Today's clients expect asset and wealth managers to provide insightful and timely advice, particularly given the post-pandemic market volatility.

In the McKinsey & Co report: Analytics transformation in wealth management, relationship managers spent between 60%-70% of their time in non-advisory tasks:

- They lack a unified platform for client information or research across asset classes leading to high levels of manual administration.
- They also lack pre-advisory customer risk analytics, putting the firm at risk of regulatory non-compliance.

Generalist automated data mining systems are inflexible.

Automation attempts by investing in data mining technologies or developing homegrown systems are only partially successful. Automated data mining systems are often inflexible or generalist. To work, they rely on good master data management (MDM) processes and systems and ad-hoc requests to the IT department to siphon off relevant data for a specific business purpose. Homegrown systems fail to keep up with change and are expensive to maintain.

It is these complex challenges that UniQreate's SaaS platform, NEO xTract, addresses.

UniQreate - the story so far

UniQreate was co-founded by three data and tech enthusiasts, Nitin Gupta (CEO), Kiran Kumar (CTO), and Rakesh Srivastava (COO), with the shared belief that data, AI, and automation, used wisely, can improve our lives. The idea behind what would become NEO xTract developed from Gupta's text mining project while completing an MBA at Oxford University. The project involved automation and AI to help academic and healthcare research. Before that, Nitin was a global product manager at Thomson Reuters (now a part of LSEG), where a team of hundreds would manually sift through mountains of documents to extract risk and compliance information for capital markets.

Nitin's prototype automation project convinced him that a similar approach if applied in a commercial setting, would massively improve efficiency and turn unstructured data buried in

documents into value without the business user needing help from IT departments or running the operations risk of sharing the knowledge with a 3rd party.

Kiran Kumar, with over twenty years as a systems engineer, developed solutions to tackle complex problems for clients, including the National Stock Exchange and the Government of India but had always hankered to create something original and was attracted to Nitin's idea. His long-time friend, Rakesh Srivastava, again a long-time veteran within IT services and recently, after seven years as Global Competency Manager at CGI, would also bring in-depth delivery, cloud, and project management skills to the team.

After some initial mentoring support, the three co-founders decided to focus on financial services and capital markets. They knew data and domain experts were heavily reliant on manually intensive processes and where time and speed were critical to success.

The fundamental design criteria were to develop an automated and intelligent document > data lifecycle requiring no technical skills that could be rapidly deployed with minimum training and provide domain experts with the information they needed. Unlike in-house development projects designed for a specific use case, UniQreate, with a product-centered focus, developed a platform that could be adapted to different use cases. After a few years and input from early clients, NEO xTract was launched in August 2022.

NEO xTract solution overview

Documents can be easily uploaded to the SaaS platform by a business user, who is presented with an intuitive screen and can specify the type of data that is relevant. After a few documents have been added, the machine learning algorithms rapidly learn what data is essential and predict relevant data found in other documents irrespective of structure, layout, context, or format. It also provides indications of provenance so the user can verify the selection. This further allows for the models to re-calibrate in an incremental learning framework. Each document improves the accuracy of predictions. When the user is satisfied, they can upload more documents without worrying about formats, confident that the platform will automatically extract, verify, synthesize and present the required data. Links to source documents enable the user to drill down to more detail if needed or to satisfy them that the predictions are still accurate.

Figure 1 outlines the simple value proposition and process.

xTract process

FIGURE 1: NEO XTRACT AUTOMATED PROCESS

NEO xTract puts the power of data in the hands of the business domain experts

Manual approaches are slow, expensive, and prone to inaccuracy.

The NEO xTract platform massively reduces the effort and time in extracting, verifying, and exporting "database worthy data" for use by an analyst or other domain expert. Manual approaches typically involve teams of researchers looking for data in multiple documents and formats. Domain experts must specify what data they seek and hope those performing the manual tasks fully understand the brief. As data is often copied and reformatted to fit a spreadsheet before it can be analyzed, it is prone to error, exposing the domain expert to the risk of giving poor advice. Time and accuracy are of the essence, and any delay reduces the value of a perishable information asset.

Home-grown automation attempts are inflexible and involve many steps in a complex process.

Even large organizations have limited development resources, and business users in different departments may flood the development team with requests to support individual use cases. It is not uncommon to involve a dozen steps or more just to satisfy a single use case.

The domain expert must articulate their requirements and perhaps educate a team of developers to ensure their request is fully understood. A feasibility study and a business case might follow before any decision to proceed. Sample documents will have to be marked up and relevant data highlighted. The developers then write scripts and create templates for gathering data. Any data scientist will need to train ML models to produce predictable results. Once the decision to develop the solution is approved, the application will be developed and then deployed. Domain experts must verify the results before they have confidence in the solution. Data from different sources and formats will need to be transformed, cleaned, and normalized into a standardized readable format before being exported for use by the expert.

> 'We want to put domain experts close to the data but far from the technology, which they shouldn't have to worry about."
>
> Nitin Gupta, CEO, and co-founder

NEO xTract simplifies the complexity by reducing the process to three simple steps

UniQreate has managed to eliminate complexity, time, and difficulty in harvesting valuable data from documents in any form. The solution involves three simple steps and can be deployed across different use cases within minutes. The three steps outlined in Figure 1 involve the following activities:

1. Upload: Documents will be ingested by the product, and it will understand the structure and content of the document to enhance the user experience in the following steps.
2. Extract / Verify: The document will be presented on a web page with an easy-to-use yet comprehensive set of features that help quickly select and extract the data of interest. Once the engine has learned enough about the data, it pre-selects the values and only requires verification.
3. Export: The data that has been extracted can be sliced, combined, normalized and formatted in multiple ways to suit the exact needs of each business and its teams. This step transforms the source data and gives it to you in easy-to-use CSV or JSON formats.

The platform is scalable across multiple use cases without having to write a single line of code. This provides another advantage over home-grown developments. The ML algorithms understand documents and different formats and can be trained by the user within hours to harvest relevant information. As Nitin said, we want to put domain experts close to the data they need but far from the technology, which they shouldn't have to worry about.

Low cost without the need for business cases or procurement sign-off.

The cost to subscribe is very low. Individual subscriptions covering up to 300 documents a month are currently $83.33 per month billed annually. A team of three and up to 900

CX-Create

documents costs $208.33 per month, billed annually. Five users and up to 1500 documents cost $291.33 monthly, billed annually.

Current position

The NEO xTract SaaS platform runs on MS Azure serverless technology and will be leveraging Oracle Cloud Infrastructure (OCI)and its autonomous database and serverless tech very soon. Initial clients include Capital Markets, Insuretech, Legaltech, etc.

The current go-to-market model is direct; however, recent partnerships with clients and ecosystem partners like Oracle are expected to add substantial market bandwidth to the company. On June 13, 2022, Nitin gave a short introductory presentation at the S3 AIR Summit 6.0 at the New York Stock Exchange. This event focused on innovation in investment management and introduced innovators providing new methods to help the investment community search for alpha. (The UniQreate pitch can be accessed via this www.airsumm.it/uniqreate).

Financial services today broader market tomorrow

Many industries face similar challenges maximizing value from unstructured document-driven data, and UniQreate has attracted interest from several prominent organizations, including brokerages, asset managers, consultants, and insurance companies.

Future direction - what to expect

The pressures to automate and seek sources of growth have significantly increased following the Covid-19 pandemic, global volatility exacerbated by Russia's invasion of Ukraine, and expanding regulatory pressures on financial reporting, including information on ESG performance. While automation of operational systems and customer engagement has been the central area of focus, data buried in documents has had less attention. UniQreate and the NEO xTract are likely to be of increasing interest in any industry where market analysis and investment decisions are critical for growth.

Oracle Cloud Infrastructure and Oracle's support prove their value to UniQreate.

The initial reason for selecting Oracle Cloud Infrastructure (OCI) was that many of UniQreate's clients and potential clients use Oracle products and services in the financial sector. Oracle's favorable terms for startups provide OCI with a low introductory discount of 75% for the first two years. Before signing up, the team explored technical benchmarking data to satisfy that OCI offered scalability, security, and flexibility, to meet different client needs and locations.

When asked about the experience, Nitin said that the Oracle startup support team helped navigate Oracle and provided access to technology experts and advice. The experience felt like a true partnership, and future introductions to existing Oracle customers will help grow the company.

CX-Create's viewpoint

UniQreate solves a common problem.

NEO xTract solves a significant challenge for asset managers wishing to deliver their best investment advice to clients and meet regulatory requirements. Reducing the time from document to value gives advisors a competitive advantage and contributes positively to the client experience, fostering trust and enhancing the firm's reputation. While most automation initiatives have been geared towards operations, supply chains, and customer engagement, support for domain experts has been largely overlooked. UniQreate fills that gap, and its highly flexible and adaptable platform offers support for any industry where domain experts need to extract valuable data buried in documents.

Their partnerships with various ecosystem organizations, including ISVs, Cloud, Data Warehouses, Consulting, BPO/BPM, etc., will significantly enhance market coverage and likely accelerate UniQreate's growth.

Summary details

Table 1: Fact sheet

Solution name	NEO xTract	Solution category	Automated document to data lifecycle
Key industries	Consulting, financial services, healthcare, insurance and other knowledge intensive industries	Geographies	Global
Deployment model	SaaS	Licensing basis	Subscription
Size of organizations served	Potentially any size	Go-to-market model	Direct and with partners
URL	https://www.uniqreate.com/	HQ	New York, US

Appendix

Further reading

- June 13, 2022, introductory presentation at the S3 AIR Summit 6.0 at the New York Stock Exchange. This event focused on innovation in investment management and introduced innovators providing new methods to help the investment community search for alpha www.airsumm.it/uniqreate
- For more reports on innovative startups go to: https://cxcreate.io/blog/ and select the category Startups and Scale-ups for innovation.