Monetizing Data Through Decision Intelligence

A White Paper for Financial Institutions

CIOs, CTOs & CDOs

quantexa

Connecting data | empowering decisions

Monetize your data and join the context-driven AI revolution

Contents

Data: The new boardroom agenda

The world is a very different place from what it used to be. Seven of the top ten companies are global Internet technology giants. These organizations focus on competing with low headcount, intellectual property rich propositions driven by data and a deep understanding of their customers. This evolution from traditional businesses (Figure 1) has been accelerating over the last ten years but is now set to become a new revolution that affects all organizations.

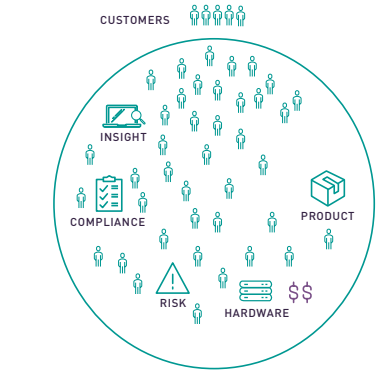
There is little choice for institutions:

Monetize your data and join the context-driven AI revolution or face asymmetric competition.

The response has been that board members are now turning to their CIOs and CDOs to ask questions such as:

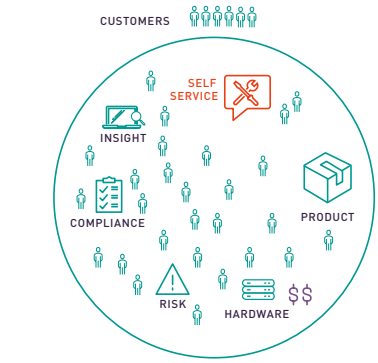
- How do we monetize our data?
- Are we adopting AI?
- Is our data unlocking our digital channels?

In this white paper, we share the essential building blocks IT or CDO must put in place to address these challenges.



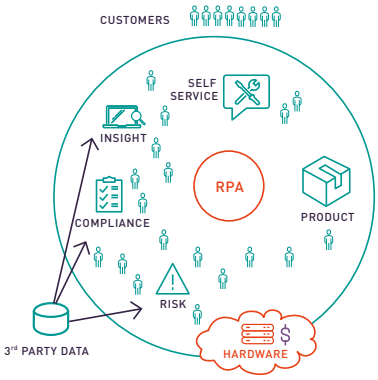
TRADITIONAL
1920 -2005

Organizations were people-intensive, limiting the product development and customers acquisition.



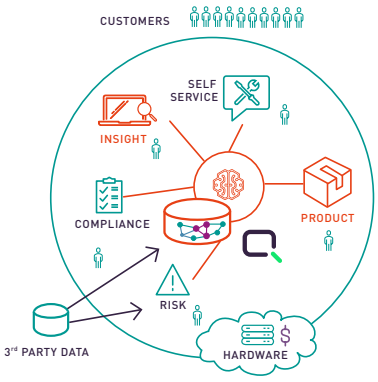
DIGITAL CHANNELS
2005

The advent of the Internet accelerated the adoption of digital channels, reducing customer service headcount and providing more data for customer insight. This allowed scope for more products and new models for customer acquisition growth.



DATA, CLOUD & RPA
2015

Third-party data, compute capacity and SaaS applications became more available, optimizing a range of back office functions. This was enhanced through quick-win tactical RPA deployments.



AI AUTOMATED DECISIONS
2017

Internet giants are driving standards for competition, forcing organizations to drastically cut costs. Data and context-driven automated decisioning is paving the way for a "white-collar worker revolution", impacting customer insight, compliance, product diversity and company growth.

Figure 1. *The evolution of organizations who monetize their data for competitive advantage.*

Decision intelligence driven by context

For many years business intelligence and analytical solutions have provided support to human decision makers for a range of matters: financial crime management, qualifying prospects, credit risk assessment, driving share of wallet, detecting fraud and customer onboarding. And it's not only the customers – there are also suppliers, employees and ultimately managing overall risk regardless of whether you work in B2C or B2B scenarios.

However, in a world of digital channels, many customers are expecting immediate responses and are turning to self-service approaches. Furthermore, increased competition is driving organizations to make better decisions, reduce their headcount and lower their cost base. These factors change the dynamic for financial institutions driving the need for more fully automated or human augmented decisioning.

Decision intelligence is the new business intelligence

Decision intelligence (DI) is the new Business Intelligence (BI) – the natural progression. It goes one step further, by not just providing the information for your staff to make better decisions, but it will also make a fully automated decision within seconds, 24/7, 365 days a year. If it cannot make the decision itself, it can refer to a human but provide all the gathered insight and highlight the key inputs for the decision, ensuring the human can focus on the decision-making as opposed to information gathering.

It may come as no surprise that this sounds like Artificial Intelligence (AI) – it's certainly part of the equation. However, AI alone has historically failed to achieve the level of automation in all but a few cases. AI in its simplest form is defined as the ability for a computer to act in a manner that is at least as good as, if not better than, a human in executing a given task (The Turing Test). Historically, analytical models have been used to pick out fraudulent claims with some level of success. However, they have often either missed the smart fraudsters or generated too many false positives for investigators to work through.

Today, there is renewed interest in advanced techniques such as "deep learning". These have showed signs on some small incremental gains in accuracy, subject to the use case and volume of data. The secret to changing the game and create accurate decisions that are good enough to deploy into production is **context.**

What is context?

A human will rarely make a decision without context, so why expect a computer to?

Imagine buying a house. You wouldn't simply peer through the letterbox and say, "I'll buy it". You would go inside, look at all the rooms and assess the condition, the plumbing, the electrics and heating – this is the **internal** data. You would then look around the neighborhood, the local schools, public transport, the value of properties in the area and crime statistics – this is your **external** data.

In order to bring data together, you first need to produce a single view or resolve entities of customers, suppliers, and employees, etc. You then need to understand the links between the data – i.e. who or which businesses are connected. This could be through transactions, common directors or shareholders and even addresses or correlated behaviors.

Using context to make better decisions

Let's consider using context in a commercial lending decision for a small to medium enterprise (SME). Without automated context, a traditional scorecard may look at a set of factors, such as the historic financial performance, and obtain a qualitative view of the SME's management team. The reality being that it is only likely to lend <$100k without the involvement of a credit risk decisioning team.

Alternatively, we can include a single view of the SME across internal and external data sources, as well as a network view of the ownership structure and payment data. The ownership data could tell us about the directors and shareholders, their history and track record. This brings to light board members who have come from serially successful backgrounds or, conversely, have a track record of failed businesses. The payment data has the potential to provide insight into the SME's customers and suppliers, how often they pay and how this has changed over time. For a bank with a large market share, it could also provide further insight into the supply chain within which the SME operates and any concentration or system risks that may exist. Using a single view of data including utilization of lending facilities or understanding any subsidiary structures with international footprints can also help as you move to the mid-market.

A second check is for the risk of **bust-out fraud.** Although this is technically fraud, it can often be the domain of credit risk. This is hard to spot without the presence of linked networks within the data. For example, a business may look normal until you detect cash being cycled through bogus employees and back to other companies cyclically. The telltale sign is when the companies start taking out credit and going into arrears one by one. At this point, it is time to stop lending to all businesses connected on this network. This is an example of how a knowledge graph within the data is also used as an input into a detection model to make it more accurate.

Finally, we can consider the challenge of identifying trade finance based money laundering, which is notoriously difficult to detect. The first step is to resolve entities within the internal and external data so that you can build a profile of each business – this is referred to as **entity resolution** and can happen in either real time or batch.

For example, in Figure 2, the banks internal data for "Business A" in Africa is enriched with external data to identify the company ownership structure and also the products and payment data associated with Business A. The next step, again possible in real time or batch, is for the system to automatically build the knowledge graph based on shared directorships, addresses, emails addresses, payment and trades. A detection model can now pick up on risks such as directors also owning companies in other territories (such as Business B and Business C) with which there are trades in place, with money ultimately flowing to other pseudo-businesses in further countries. This could potentially represent the laundering of corrupt funds from Africa into Europe.

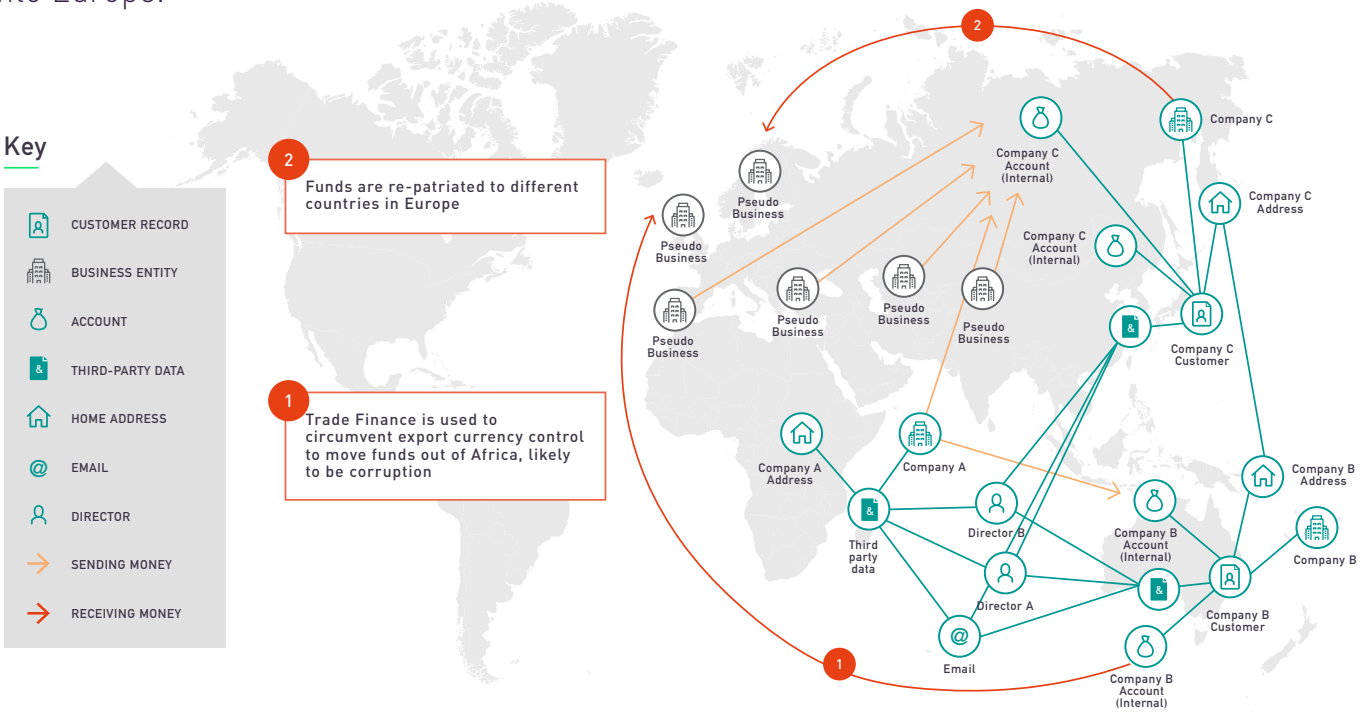


Figure 2. *An example of context automatically built using data.*

When combined with AI models, **context** is the essential ingredient that enables true decision intelligence, where models are suitably accurate to deploy in production environments. It does not have to be limited to customer data, suppliers, employees, IT infrastructure, directors, etc., as all interact with each other in graphs, as long as you can resolve the underlying entities across many internal and external data sets. Quantexa provides the platform that enables context-driven decision intelligence, solving the hard problems and letting you focus on using your preferred tools to create models will allow you to monetize your data.

How is context used

The relationship between context and decision intelligence is outlined in *Figure 3.*

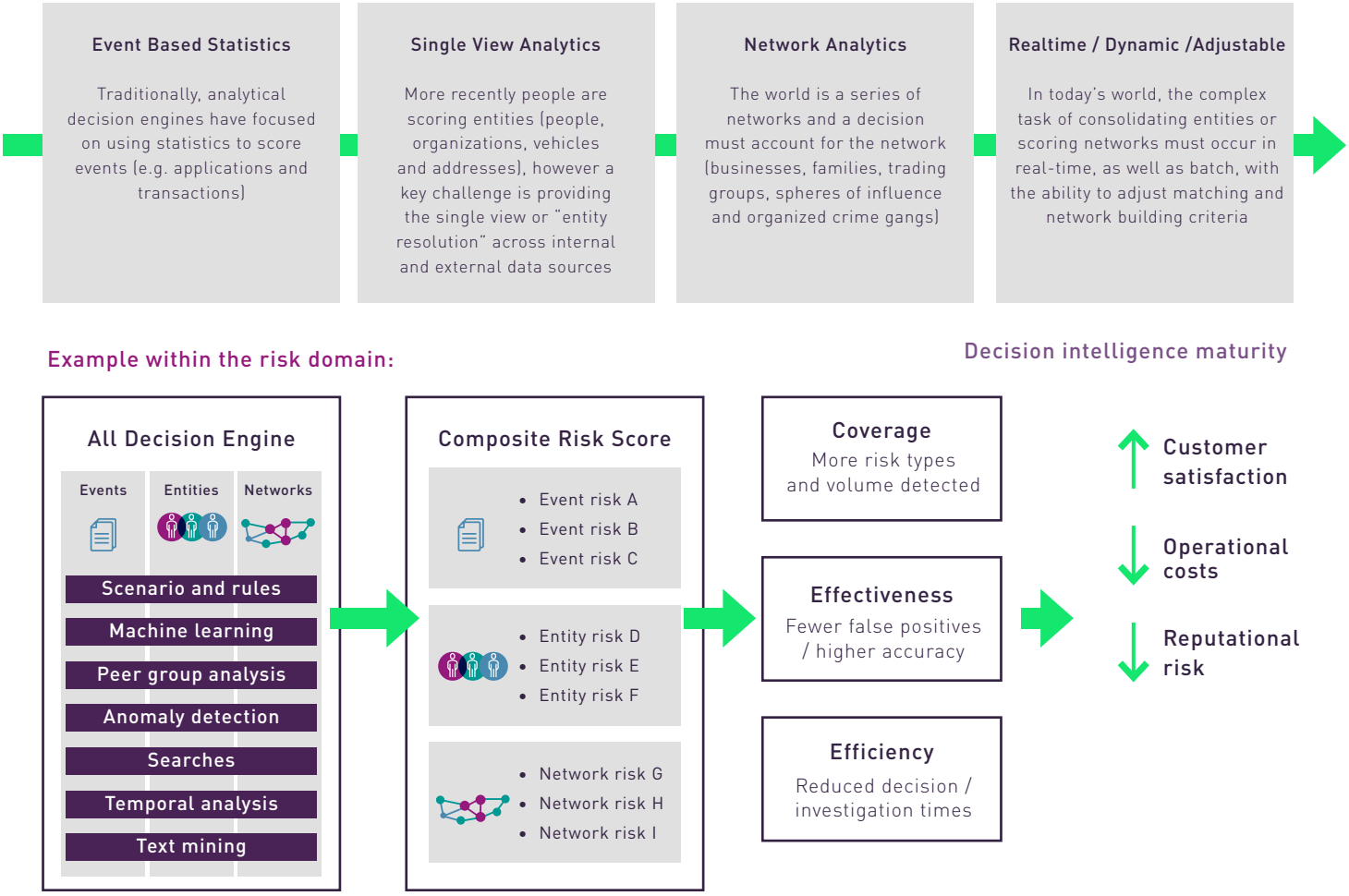


Figure 3. *How context enables better decision intelligence.*

Use cases for context-enabled decision intelligence

In reality, the options for optimizing decisions across the organization are almost limitless (Figure 4). Imagine a scenario where all of your organization's data is enriched and provided in a contextual view – customer, staff, finance and supplier decisions could all be optimized.

However, the reality is that some use cases deliver far more value for money and justify loading the data incrementally into a context-oriented platform within a data lake.

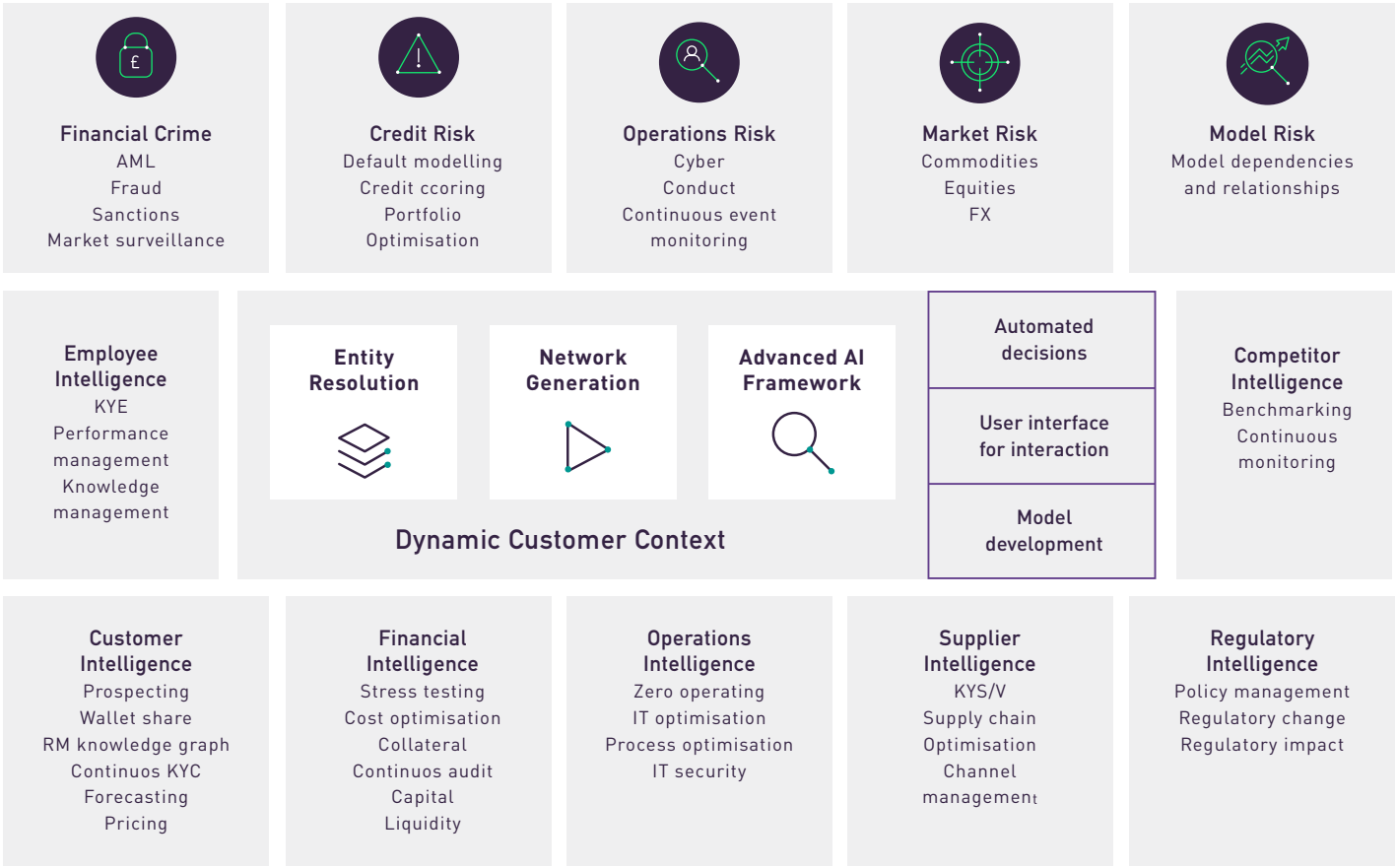


Figure 4. *Context can drive value across multiple use cases.*

Listed over the next few pages are examples of use cases where context has been used to deliver significant fiscal value.

Customer intelligence

Organizations that deal with customers, whether individuals or businesses, often struggle to achieve a "single customer view" across their entire customer base – this requires **entity resolution**.

But why stop at the single customer view? A dynamic contextual view of your customer brings together internal and external relationship and behavioral data that is always up-to-date. A complete understanding of the customer is needed to underpin traditional sales models as well as digital channels, where the real time nature is critical. It unlocks many business opportunities:

B2B Prospecting:

If selling to businesses, by enriching your current customer context with third-party data, such as company and ownership data, it is possible to find links from existing customers to potential prospects. This could be a board or shareholding relationship, or a connection through a common address. This can be used by relationship managers or brokers to approach current customers for referrals, especially if a model is used to predict the likely adoption. A contextual view is extremely valuable where organisations have complex structures. By adding in your existing clients' payment data, you will also be able to extend the referral relationships to your clients' customers or their suppliers.

Wallet share:

If you have multiple products that could be sold to an individual or business, a contextual view of data silos will help you understand what products you have sold and to whom. Based on peer group analysis, you can determine other products that may be of interest to particular customers and support a "next best action" execution method – digital or human.

Churn prevention:

By aggregating data across different channels and subsidiaries, it may be possible to pre-empt churn using a model. As a minimum for complex multi-product businesses, this enables you to orchestrate better automated outreach at the point of renewal across, what could be, multiple siloed products.

Compliance

The cost of compliance in banking has been escalating, with evermore regulations and increasing sophistication among the criminals.

KYC

When onboarding business customers, you must ensure that opportunities are not being provided for known money launders or sanctions are not being breached. Third-party data, or even internal data from different lines of business, can be used to pre-populate or pre-screen onboarding KYC (Know Your Customer). Businesses are not static, so a continuous or trigger-based KYC approach will use the dynamic customer context to automatically alert if there are new risks raised based on changes in ownership structure or transactional behavior. An obvious extension to KYC is KYS (Know Your Supplier).

AML

AML (Anti-Money Laundering) is a continuously evolving responsibility driven by regulation in response to evermore ingenious criminals seeking to break the audit trail of their money movements. This has led to sophisticated schemes, such as the use of markets instruments and trade finance. These are notoriously difficult to detect without the use of context. However, by adopting a context-based approach it is possible to reduce false positives one hundred fold, but creating more legitimate alerts of suspicious activity than ever before.

Fraud

Banks face a raft of different fraud types. Fraud can affect a company in multiple ways, including financial losses, impacting customer service, reputational risk and even compliance issues.

First party fraud

Whether at point of application or post-application customers (either retail or SME commercial) can defraud the bank. This is often as part of organized criminality targeting lines of credit. Having a full picture of the context is often the only way these types of fraud can be interdicted before they escalate.

Internal fraud

As a form of conduct risk, these can take many forms: first party collusion, unauthorized trading, erroneous product creation, bribery or corruption. The effective detection of these types of behavior requires joining the dots in the data to produce a single view of the employees, as well as a network view of their related interactions.

Credit risk

Although the credit risk environment within the banking domain is heavily regulated, there is still the opportunity to take some innovative approaches to improve the P&L by better managing risk and improving efficiency. For example, there is the initial commercial decision to put in place a facility or extend a facility. In an ideal world, if risk were better understood you could lend more with confidence. If the threshold could be increased for which a model-based decision were made, this can improve customer experience as well as efficiency.

Decisions for a digital channel could be automated and only more complex cases referred for manual decisions. If the full picture is presented, and data is pre-aggregated, this saves time and effort. Once the facility is in place, it would be beneficial to monitor and automatically trigger alerts as an early warning if the data indicates any issues.

Using context as an input for commercial lending models can ensure that as much of the data is under consideration for models making decisions. This includes the corporate ownership structures, payments and their indication of a business's customer history and supplier payments. It can also be extended to any relevant adverse media and other data feeds that could be relevant.

Taking a platform approach

Although the use cases discussed are often considered as their own silos, there is broadly the same underlying approach – ingest data (often the same data points), enrich it, run a model and make a decision. Therefore, it is not unreasonable to consider the same underlying technology platform supporting the use cases working in unison, as indicated in Figure 5.

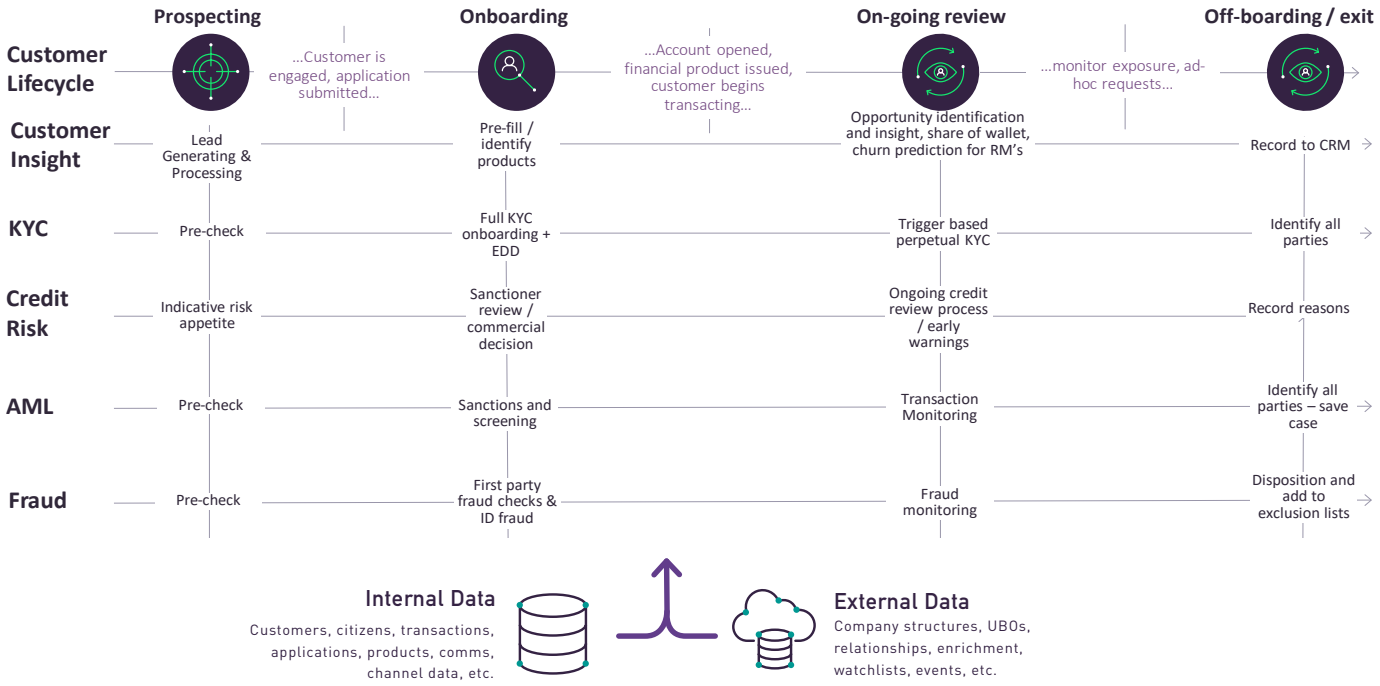


Figure 5. *Taking a platform approach for commercial banking customers.*

The concept being that the data is held in a single platform providing all the context that is required. This context can be called in real time or in batch to deliver a score at different stages in the customer life cycle. The way the data is ingested and the platform remain consistent. However, the model is specific to the given application, such as a fraud check or KYC check and can request the data as required by each use case. The user interface or the integration of the outcome can be customized to the different user sets. Performance enhancements across the different use case silos can be broken down. For example, it makes no sense to provide an RM with a potential lead if the fraud or KYC check indicates that the prospect would not pass the risk profile.

Architectures for context-enabled decision intelligence

Organizations often consider that creating an architecture for single customer views, model development or automated decisioning is a highly complicated undertaking that can take years and is plagued by challenges such as data quality. However, this doesn't need to be the case. Large complex international institutions have proven the concept in weeks and had product deployments live within six months.

This section outlines the requirements, approach and the technologies involved.

Creating a Dynamic Customer Context for decision intelligence

Context can be built around any type of data, suppliers, employees or customers. Figure 6 represents a conceptual view of the context-driven decision intelligence approach that focuses on customers – we call this a Dynamic Customer Context (DCC).

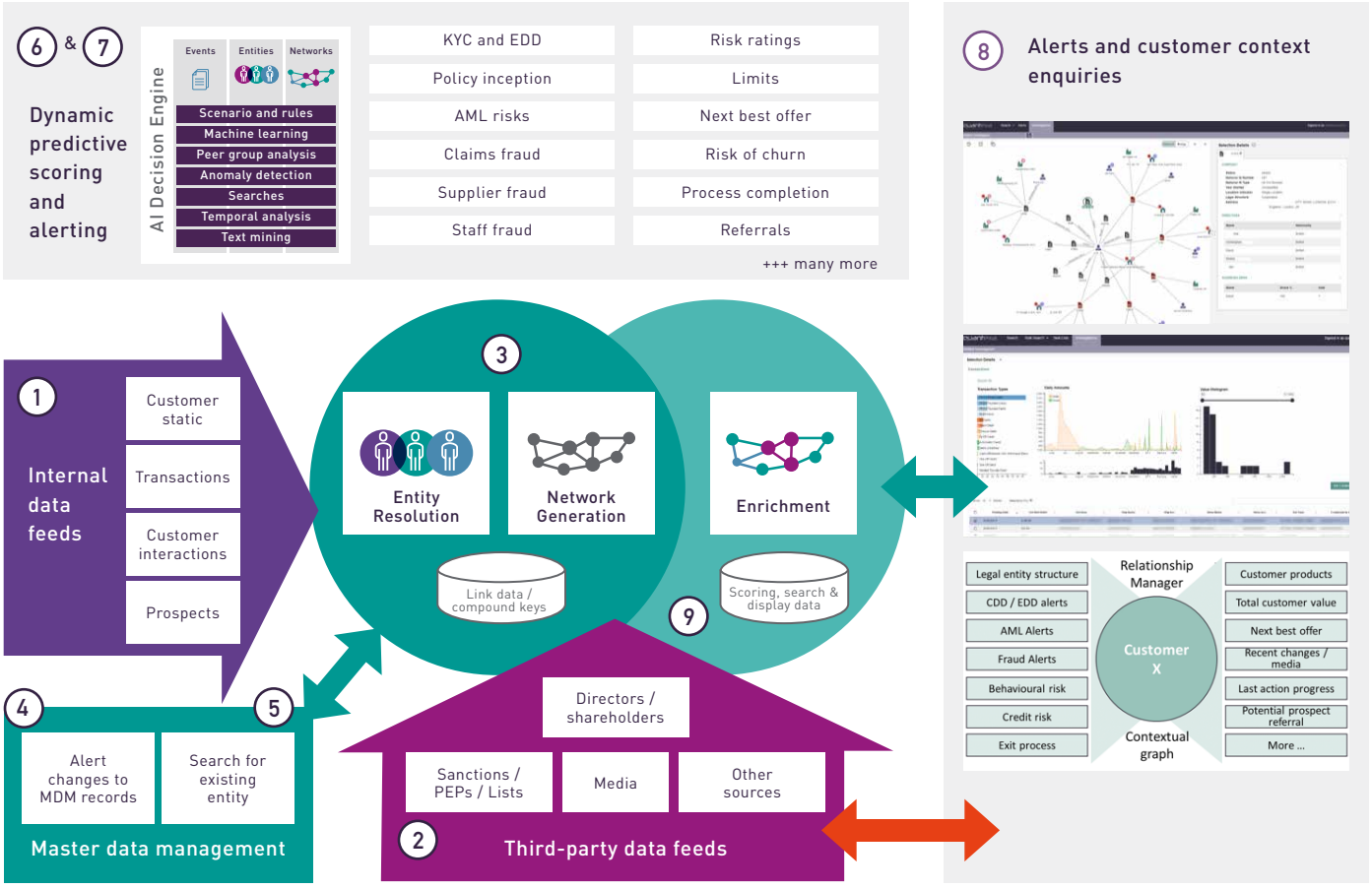


Figure 6. *Context-driven decision intelligence approach.*

The key building blocks of a context-driven decision intelligence platform are described below for an instance focused on customer decisions – we refer to this as a Dynamic Customer Context (DCC):

1. **Real-time ingest of internal data:** This data needs to be added to the DCC as soon as it's available. Although this can be a daily batch, as organizations evolve this can be either micro batches or even streamed.

2. **Real-time ingest of external data:** This may be uploads from multiple providers, in batch or on-demand. It can include a whole range of data, including media where entities have been extracted to convert the data from unstructured to structured, traditional company ownership information, risk data and other lists.

3. **Dynamic entity resolution and network generation:** This allows APIs to be called in order to assemble resolved entities in near real-time using templates for varying degrees of fuzziness. This enables organizations to enrich their internal data with third-party data sources to provide a more comprehensive single view of the customer while overcoming data gaps or data quality issues. However, this is not limited to simple entities, hierarchical views of businesses can also be generated and enriched on demand. Additionally, the engine needs to be able to provide networks based on a range of different templates. In both instances the platform can obey the security constraints according to the requestor's credentials for both entity resolution and network generation down to field level, row level or data source level.

4. **Interaction with master data management:** Given the DCC represents a central point of aggregation for changes in the customer context, it should interact closely with any master data management environment. This involves both a feed of known and confirmed facts regarding the customer and their single view, as well as raising alerts and suggestions to changes in the customer information based on inbound data.

5. **Search for the existence of a customer:** In most institutions there are multiple touch points where an entity could be created, whether directly as a customer or a counterparty. It is important to be able to quickly search on multiple criteria to narrow down an entity to one which may already be defined within the institution. Rather than creating a duplicate, a link is created to an existing DCC record.

6. **Continuous update of the DCC:** For a range of changes to the customer based on internal or external data changes, different categories of profile can be updated through the triggering of defined decision models that monitor for particular actions to be taken. These apply AI techniques across the full DCC.
7. **Real-time alert generation:** As well as changes in the data updating profiles or generating automated actions, they may also generate alerts that need to be actioned by a human operative. This scoring capability can also be event-driven, where an API can request a score on demand and the system can respond in seconds based on the most up-to-date customer data.

8. **Support for DCC enquires:** If a human operative wants the very latest customer information, they can request a full picture that is specific to their need at that point in time. This could be to get a referral from an existing customer to a new prospect, understanding the next best offer for a customer or the investigation of a potential fraud alert.

9. **Scoring, search and display data:** With the DCC being focused on the linking and metadata that drives entity resolution and network generation, there is also the need for full contextual data that can be used when searching for customers, display details or generating scores from models.

Define your own trigger thresholds

The DCC is fed with data at multiple frequencies, but the institution can decide the frequency at which profiles are recalculated or actions are generated, for example, in near real-time, on-demand, intra-day, daily or lower frequencies.

This definition of a consistent and organization-wide control of profile events, assessment timing, and trigger events has three key advantages. Firstly, it is a consistent approach that is not open to the interpretation of individuals. Second, it can be at higher frequencies than traditional approaches, thus reducing lag times that could miss a customer cross-sell opportunity or an escalating risk. Finally, it is relatively easy to centrally change the mandate based on geo-political trends, new competition or risk management emphasis.

Benefits of the Quantexa's decision intelligence approach

IT operational benefits

Open architecture. Quantexa has adopted a fully open architecture approach. This ensures the DCC can be easily integrated with any other systems within the organization. This co-existence strategy ensures that existing investments are embraced and a best-of-breed, future-proof approach where new innovative open technologies and tools can easily be introduced.

Alignment with private or public cloud approaches. An open approach ensures that deployments can scale out across any form of public or private cloud architecture. Deployments can be native or containerized using systems like Docker or Kubernetes. This allows small to huge deployments to easily be deployed through horizontal scaling, as well as ensuring any on-premise investments are smoothly migrated to future cloud aspirations. It also provides far simpler models for high availability arrangements at significantly reduced costs using commodity arrangements. Coupled with the support for use of agile approaches to development and deployment, time to value is reduced and updates are easily managed.

Build once, use many. The same underlying repository of data can service multiple use cases, creating resolved entities and networks for different purposes, without multiple replications of data repositories. The platform is proven to scale to over fifty billion records with thousands of end users.

Meeting security needs. Traditional batch-based environments that resolved entities and built networks would do so using "admin rights". Even if data was redacted, users could often infer the existence of sensitive data. Moving to an on-demand approach means the security credentials of the user or process requesting the entity will be used in their generation, removing security risks. There is also full support for a granular security model, supporting more demanding environments, as well as a fully audit trail of any actions taken by the users of the system.

Improvements to master data management (MDM) and data quality. It is possible to have a bi-directional relationship with MDM, where any changes in internal or external data that updates the DCC, alerts the MDM environment. In the same way, any changes within the MDM environment as relayed to the DCC. The result being overall improvement in data quality.

Use any modeling approach and technology. The same open approach means a huge range of analytical tools can be used by data scientists to develop the automated decisioning models.

Supporting customer self-sufficiency. The Quantexa Technology Academy is a self-service training environment that enables our customers, or their preferred integration partners, to upskill their staff in network analytics approaches and development of models leveraging our platform.

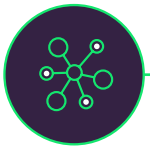
Business operational benefits

Generate more opportunities or risks. The use of networks to underpin predictive models leaves fewer stones unturned. For example, ensuring that even the most complex risk situations, money laundering schemes or fraud can be identified.

Improved effectiveness. Networks, when used to underpin models, also significantly reduce the false positives. Less time and resources are wasted on following up on unfruitful opportunities.

Faster and more efficient human interventions. The combination of drawing together all the data sources into a single customer view, augmented by the networks, cuts through the data preparation effort faced by employees tasked with making a decision. With this approach, they immediately focus on the decision itself, guided by the AI models, cutting efforts by up to 70%.

The use of networks drives attention to higher value. Smaller risks and opportunities generally do not represent themselves as large networks of activity. By using networks, an organization can focus on the high value areas.



Aggregation of third-party data providers. Organizations often find themselves using several third-party data providers for different use cases. The result can be diminished purchasing power or the use of non-optimal providers for particular jurisdictions. By aggregating multiple providers at an institution-wide level, individual departments and use cases can leverage and pay for the most appropriate source with confidence of purchasing power.

Support for digital transformation. As institutions are moving to digital channels, it becomes imperative that an accurate view of a customer is available and decisions can be made based on context in real-time.

Immune to data quality issues. The advanced entity resolution approach uses a multitude of key combinations, data statistics and models to transparently resolve entities and match data despite missing or poor-quality data. This statistical single view of customer does not require extended data quality programs to complete.

Explainability and transparency. The approaches used within Quantexa's platform are fully transparent, even where advanced AI techniques are used. This ensures the resulting implementation for entity resolution and network build remains fully transparent. If appropriate approaches are taken to building the detection models, these can have the "model lift" associated with networks. Additionally, there is a full audit trail of the data and models that were used at a given point in time.

Deploying Quantexa's decision intelligence platform

Quantexa's Contextual Decision Intelligence is the only near real time, proven, scalable (tens of billions of records and thousands of users) solution using context and a full security model. It provides the basis for the majority of your automated AI decisioning needs. Although Quantexa offers several detection models, our customers' own teams and their partners can also develop models, and we incorporate extensive self-service training to support upskilling.

Our belief is that the decision intelligence platform should be an integral part of any organizations' data fabric and should operate in both batch and, crucially for today's digital world, in real time too.

Figure 7. **illustrates our approach to deploying the decision intelligence platform.**

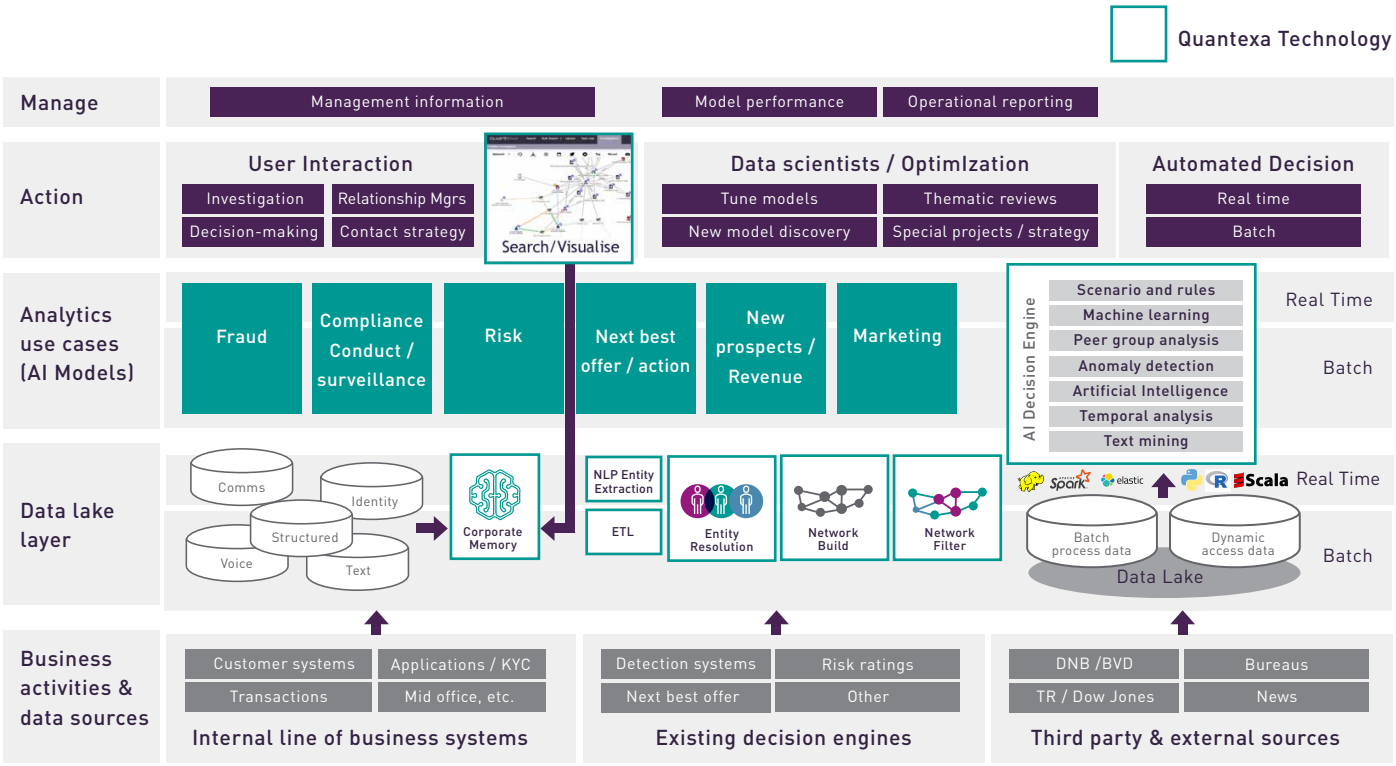


Figure 7. *Architecture of Quantexa's Contextual Decision Intelligence platform.*

Quantexa's decision intelligence platform components

Many organizations consider taking a batch-only approach, however, as outlined on the left this has numerous shortcomings. Batch environments should only be required to provide a play-pen in which data scientists can develop and test their models using their favorite tools at full data scale.

By contrast, most of the work is done by the API-driven, micro-services scale out architecture that provides key capabilities of entity resolution, network building and filtering. These services can be orchestrated by the "Decision Layer" that uses graph scripts to create the required scores or provide data through further APIs. Finally, there is a browser-based extensible interface that allows users to interact with scores, entities and graphs or networks.

Figure 8. **provides an overview of Quantexa's Contextual Decision Intelligence platform.**

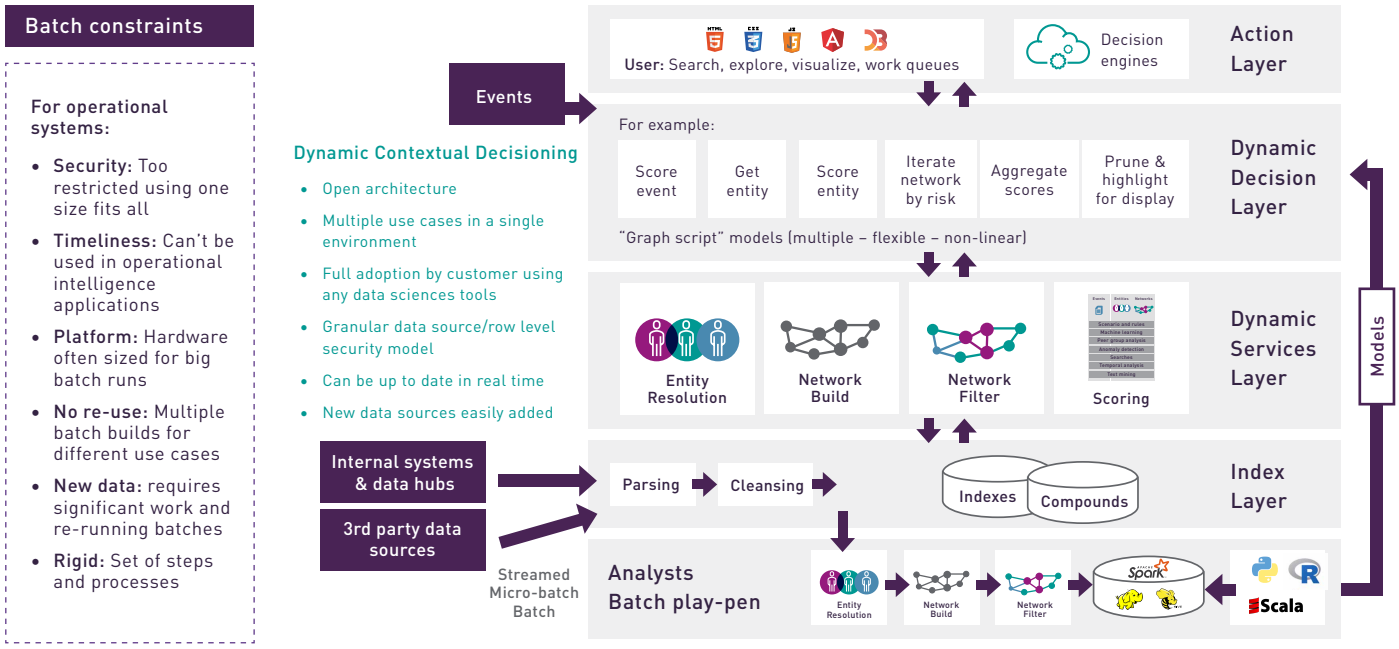


Figure 8. *Components of Quantexa's Contextual Decision Intelligence platform.*

Steps to success

Although many senior management teams intuitively believe that there is significant value to becoming a data-driven organization through decision intelligence, it is not always clear how to make this happen in practice.

Here are seven critical steps to ensure success:

1 VISION

Ensure the organization understands what could be achieved and why a decision intelligence approach can be transformational. Focus on concrete outcomes.

2 DESIGN

It is crucial to shift to an architecture that is layered horizontally. Too often there are silo solutions where the data is vertically integrated into a specific application, each of which require duplicated data and dedicated niche skills, with no flexibility to easily add new models and capability.

3 TEAM

Establish a shared resource of data science skills that can assist different business functions in getting the most out of the organization's data.

4 TRANSFORMATION PLAN

Few organizations offer a green field opportunity, you have to assume a migration. Ensure the CDO approach is embedded within change programs and that it becomes the "context data fabric" though which all data is plumbed. Measure this progress through KPIs for management across the organization.

5 QUICK WINS

These are critical to the success of the program. Choose compelling opportunities with a proven track record where analytics can provide benefit, such as within the fraud or customer insight domain.

6 COMMUNICATE AND SELL

This ensure the CDO program maintains momentum. Ensure all the departments understand how data is providing significant benefit.

7 EXPLOIT

Once established, the "context data fabric" will quickly contain most of your data. As long as there is embedded capability for entity resolution and network analytics, new models will be trivial to deploy. Then, you can efficiently create new business applications using this context and make this an ongoing program.

There is no reason why you cannot have your first decision intelligence successes using context in production in under a year.

Find out how you can use your data to make a difference.

SPEAK WITH AN EXPERT

Visit our website to learn more about Quantexa

www.quantexa.com

quantexa

Connecting data | empowering decisions

10 York Rd, London SE1 7ND

E. info@quantexa.com