

Using Data and Analytics for Maximum Impact in Tax Authorities



About the authors



Imam Hoque
*Co-founder, COO and
CPO at Quantexa*

Imam Hoque and his team pioneered entity resolution and network analytics techniques that have been deployed globally across tax authorities, federal police forces, border controls, customs and intelligence operations. Working with HMRC, Irish Tax, New Zealand Tax, Canadian Tax and others, he has personally been involved in the transformation of many of these organizations to tackle tax evasion, VAT fraud and organized crime. As co-founder and Chief Product Officer at Quantexa, Imam empowers businesses to drive decision intelligence using new technologies.



Jerome Bryssinck
*Global Head of
Government at Quantexa*

Previously the Director of Government Solutions at SAS, Jerome spent 13 years working with several large government institutions globally in creating solution roadmaps and appropriate risk strategies to combat fraud and financial crime in the public sector. Relevant tax-related projects include the deployment of fraud detection systems covering multiple domains (including VAT carousel fraud, offshore fraud and profit shifting) for Belgium, Netherlands, Estonia, Spain and Singapore. Jerome also advises Tax, Customs and Welfare administrations in Hungary, Australia, Greece, Saudi, Poland, Slovakia, Austria, France and Belgium.



Valentina Ion
*Director of Government
Industry Solutions at Microsoft*

Valentina is in charge of driving customer digital transformation based on a deep understanding of their industry, its drivers and the critical solutions. She is also responsible for the development of the government industry solutions and business development model, including establishing strategic partnerships with ICT, advisory, academia and international organizations, such as OECD and IOTA. Valentina has more than 17 years' experience in ICT, business development and sales and marketing, with an education in business & ITC, finances & marketing, graduating from the National Economics Academy in Romania, Université de Sciences Sociales Toulouse and Université d'Orléans, France.



Steve Upton
*Director of Business
Strategy at Microsoft*

As a Director of Business Strategy within Microsoft's Worldwide Government Industry organization, Steve works to empower businesses to achieve more through digital transformation across public sector industry customers and partners. The organization is responsible for defining industry-specific needs with engineering teams, defining industry solution selling go-to-market, and helping cultivate a thriving ecosystem for partner-led industry solutions globally. As part of the Microsoft team, Steve has been driving solutions in Tax, Customs, and Global Trade to support all aspects of government. In the Digital Tax area, he has worked with Microsoft's partners to design solutions that improve the Tax Agencies' operational efficiency as well as to enable ease of paying taxes and improve citizens' ability to comply. This typically results in improved tax revenue, more satisfied taxpayers, reduced processing times or costs as well as potentially improved transparency.

Contents

1. Adoption of analytics and AI across the industry	3
2. Use cases, intelligence capabilities and data sources	10
3. Best practice approach: Using the intelligence system	11
4. Analytical capabilities	12
5. A platform for now and the future	13

1. Adoption of analytics and AI across the industry

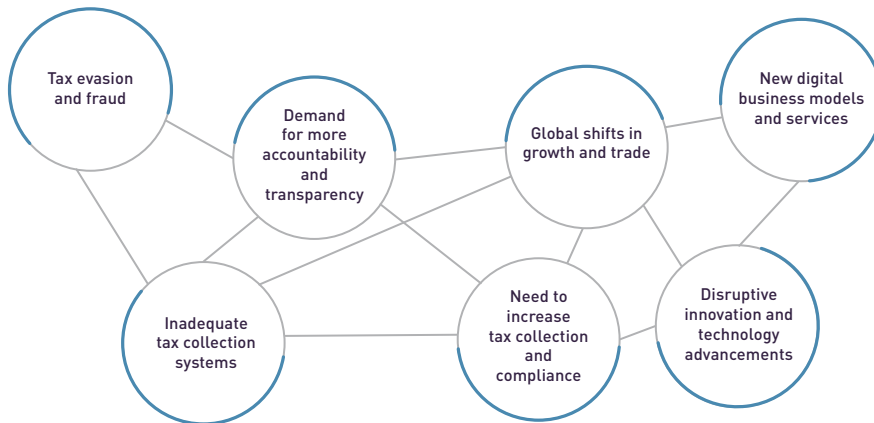
In an age when we can order clothing, hail a ride, track our fitness, book a flight, and perform multiple banking activities all from our smartphones, technology is shifting citizen expectations across the globe. These higher expectations directly translate to higher expectations for government services.

Clearly, **data** is a new type of economic asset, which is rapidly becoming vital in the global economy. Most economic activity will depend on data within a few years. Public sector bodies produce and collect huge quantities of data, a valuable raw material for the development of the economy to enable growth, create jobs, improve well-being, citizen safety and much more.

A **European data economy** is a required foundation to the EU's Digital Single Market strategy. In its European Data Market report (2017), the European Commission estimated that the value of the EU data economy was over €285 billion (2015), representing 1.94% of EU's GDP and that - with favorable policy and legislative conditions, as well as appropriate investments in ICT - the value of the European data economy could increase to €739 billion by 2020, representing 4% of the overall EU GDP.

Tax authorities are in the eye of the storm of these global forces; digital payments are growing in scale and significance, and data is becoming the currency of tomorrow. These and other changes are raising security and privacy questions and challenging the conventional role of tax authorities. These agencies, alongside the rest of the public sector and private businesses, are also facing structural changes as global growth is shifting east, global trade is coming under increased scrutiny, and employment patterns are being reshaped.

In every economy, these forces are requiring organizations to innovate rapidly, and tax authorities have a lot to gain or lose from these changes.



“Every year, cross-border VAT fraud costs our Member States and taxpayers about EUR 50 billion. At the same time, the administrative burden for small businesses is high and technical innovation poses new challenges...”

Valdis Dombrovskis
Vice-President, European Commission

“By applying advanced analytics techniques, tax administrations can begin putting their data to work to identify compliance and other risks, to tailor customer service, to design more effective treatment and intervention programs.”

Niall Cody
Chairman, Office of Revenue Commissioners Ireland

Data, analytics and artificial intelligence (AI) are powering a major revolution that is reshaping how organizations conduct people-intensive decision-making. This paper provides insight into what this revolution looks like in practice, the rate of adoption, answers the question “why now?”, examines a case study and provides practical guidance to the steps required to make the most of these disruptive new capabilities.

1.1 Decision Intelligence is the new Business Intelligence

Budget deficits around the world are driving a need for new revenue sources, and governments are looking to businesses to bridge the gap. Governments are focusing on real-time reporting at the transaction level to drive compliance and collection.

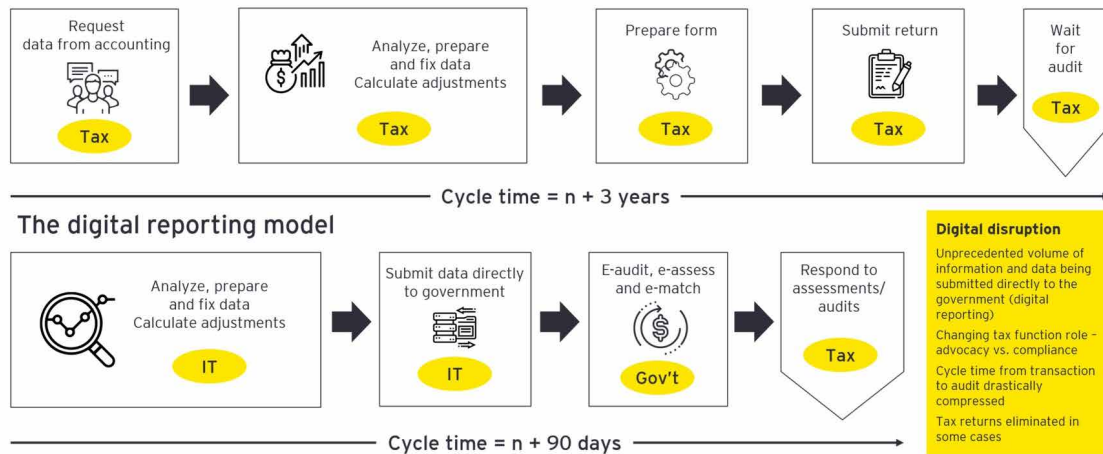


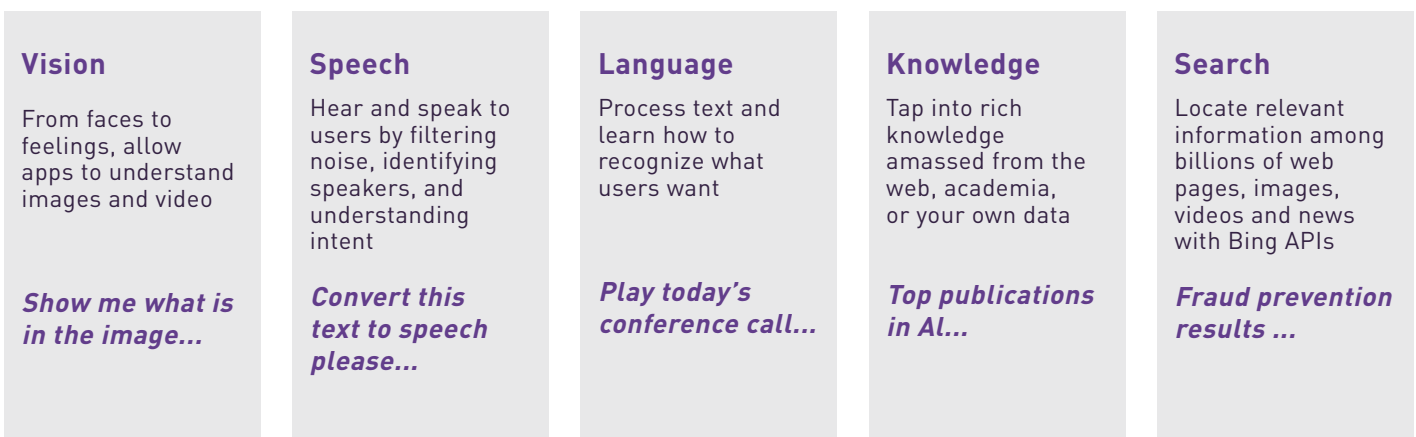
Figure 1. The traditional compliance model. Source: EY, May 2019

The traditional compliance model is changing, and both companies and tax administrations are looking into proactive measures, including:

- Anticipate tax risk and optimization models (what-if analysis – modeling capabilities)
- Connected financial and tax insights (on demand, at scale)
- Automated tax filings and real-time invoice validation and correction
- Appropriate detail and frequency of e-invoicing for regulatory authorities
- Data standards and Tax taxonomies
- Connected forecast and use of predictive models

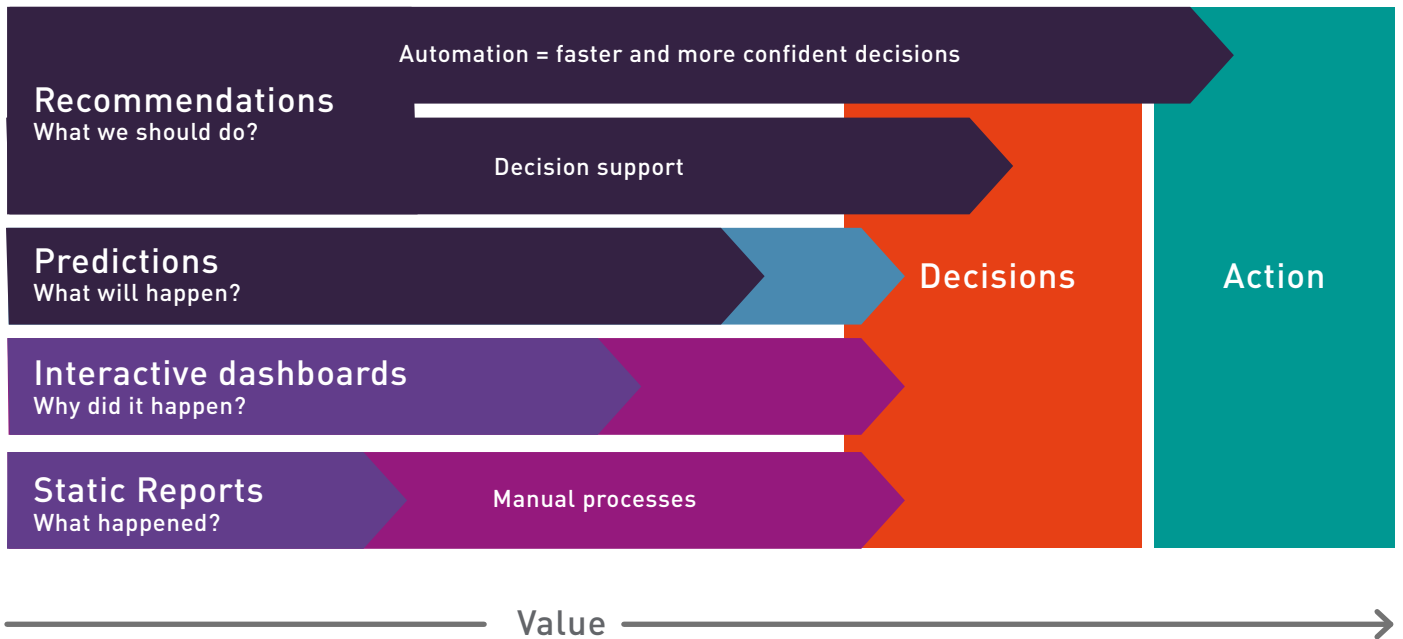
There has been a clear and significant trend within the commercial sectors of rapidly accelerating adoption of analytical techniques and AI in helping to automate people-intensive decision-making. The objective is simple and clear: where possible, automate the preparation of the data, the analysis and ultimately as many decisions as possible, more accurately than ever before.

AI is enabling new capabilities:



From data to Decision Intelligence

At the heart of the approach is being able to do this as early in the cycle as possible, by leveraging all the internal and external data available. The output typically takes the form of an automated answer: yes, no or if an automated decision cannot be made, then refer to a person but assist them in making the decision as accurately and quickly as possible. This capability represents the next generation of Business Intelligence, which the industry calls **Decision Intelligence**.



The crucial component of a decision intelligence capability is the *model* that makes the decisions – this is where the analytics and AI is embedded. The model drives the key metrics for any unit that is making decisions:

- Does it find more of what you are looking for, regardless of whether this is opportunity, risk or compliance issues?
- How effective and accurate is the model? Does it generate too many false positives to review, creating significant manual effort?
- Does it make the staff dealing with referred decisions more efficient? For example, can they halve the amount of time a review or investigation takes?
- Does it make it easier for taxpayers to comply and pay their taxes?

Within the tax domain, agencies are asking the following questions that are easily addressed by Decision Intelligence using AI and machine learning (ML):

Taxpayer experience

- Could tax administrations make it easier for the taxpayers to comply and pay taxes?
- Could tax administrations enable a personalized real-time experience?
- Could the tax administration role evolve towards tax advisory?
- How can artificial intelligence help to transform the tax contact centers towards a personalized omnichannel engagement?
- Could we make the tax process an inclusive process through real-time translation and voice recognition?

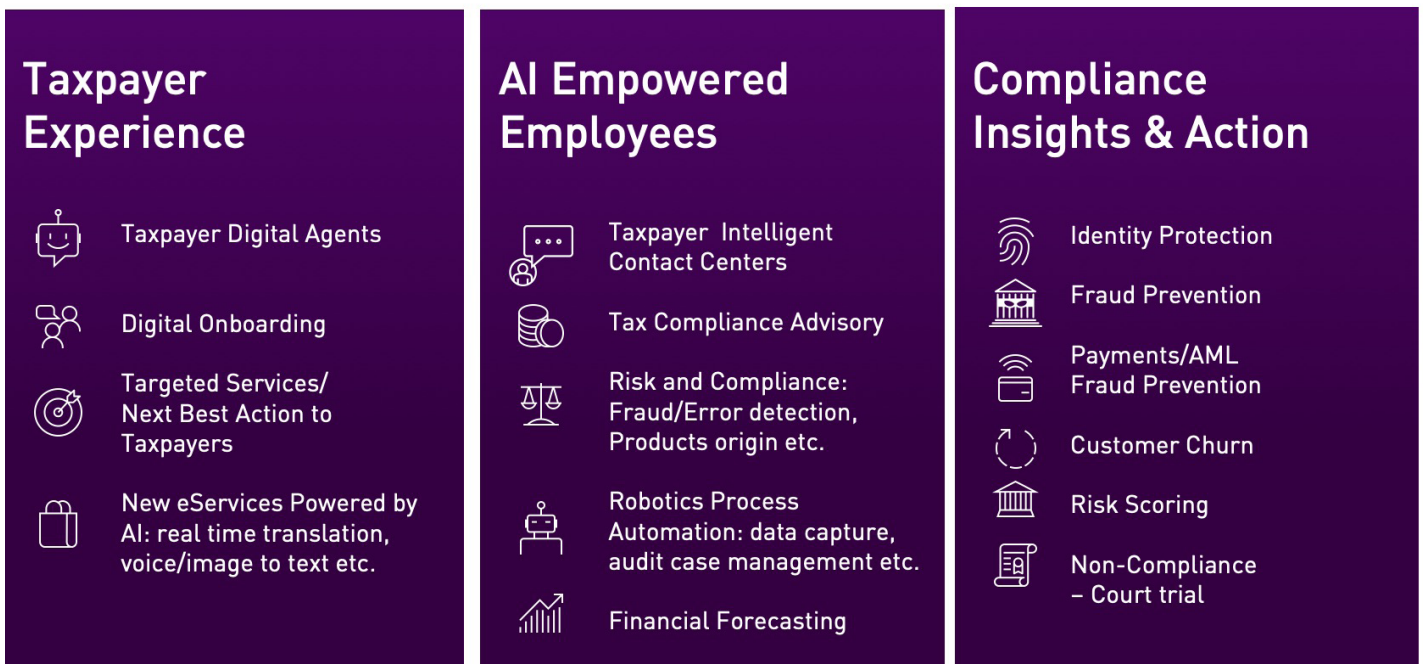
Tax employee experience

- Could employees improve response time to tax inquiries and requests?
- How can big data compute to build a 360° view of the taxpayer (using both structured and unstructured data)?
- Could advanced analytics improve transparency by allowing employees to easily view and track the progress and their interactions in real time?
- Would data-enabled visibility improve the decision-making capabilities for tax agency management with real-time dashboards and analytics?
- How can Robotics Process Automation improve employee efficiency via data capture and audit case management?
- How can technology improve risk assessment accuracy and help keep information secure?
- How could we improve on tax revenue forecasting and budgetary execution?
- Would automation cover, at least partially, the talent gap tax administrations are facing lately?
- How can tax administrations move from a reactive to proactive view of the economy evolution and be able to take timely actions?

Tax compliance

- Is an individual attempting to evade tax?
- Is a business evading corporate tax?
- Is a VAT return accurate or part of an organized VAT carousel fraud network?
- Is there evidence of base erosion and profit shifting?
- Are industry specific schemes being abused?
- Which treatment should be used for a particular business to collect unpaid taxes?
- Which inspections should be carried out in order to have maximum deterrent effect?
- Are there offshore risks?

In this paper, we will use a VAT carousel use case to illustrate what can be achieved.



1.2 Rates of adoption

As tax administrations transform their operations with technology to automate and embed tax compliance in day-to-day business operations with taxpayers, they also require strategic risk-management skills for profiling taxpayers for audits and efficiently allocating the appropriate resources. Many tax administrations are already re-evaluating their current talent needs against these additional capabilities.

For tax authorities, required data skills are less about gathering and auditing data and more about using data analytics, data mining, forensic auditing, risk management and tax technology skills to make informed decisions. When tax administrations explore how to build the necessary skills for their employees, they should focus on several strategies:

- Complex scenarios combining data from various sources and using AI and machine learning tools to identify fraud, forecast revenues and maintain taxpayer compliance
- Data analysts who use a data warehouse and analytics to produce insight for the management team of KPI metrics, assist in the planning of audits to maximize their efficiency, implement transparency and open-data initiatives, and consolidate and combine data from different sources to build a 360° taxpayer view

Tax administrations and businesses compete for tax professionals with skills in new technology, big data management and analytics, and tax risk management. Private organizations often have an advantage in this area over public administrations. To combat this, tax administrations can set up a strong ecosystem of cooperation among themselves, accountancy firms and academia, as well as IT vendors in exchange for experience sharing, idea sharing and mutual support in building a sustainable way forward.

An ecosystem approach could potentially address the bottlenecks in adopting AI by tax administrations.

Eighteen bottlenecks could limit artificial intelligence's (AI's) benefit to society.

Four categories of limitations to AI use

Critical barriers for most domains

- Data accessibility
- Data quality
- High-level AI-expertise availability
- High-level AI-expertise accessibility
- Regulatory limitations
- Organizational-deployment efficiency

Critical barriers for select cases¹

- Data volume
- Data labeling
- AI-practitioner talent availability
- AI-practitioner talent accessibility
- Access to computing capacity

Contextual challenges

- Data availability
- Data integration
- Access to technology
- Privacy concerns
- Organizational receptiveness

Potential bottlenecks

- Access to software libraries and other tools
- Organizations able to scale AI deployment

Figure 2. Examples of bottlenecks limiting AI benefits. Source: McKinsey Global Institute analysis 2019

Adoption rates are accelerating at a blistering rate; this is evidenced by the vacuum forming within the skills market where data scientists are hugely in demand. Another indicator is a recent study into the adoption of AI within the financial services (a leading indicator of technology adoption) conducted by Finextra and Quantexa in October 2019. The graph below indicates that only 11% are not currently engaging in some form of AI adoption.

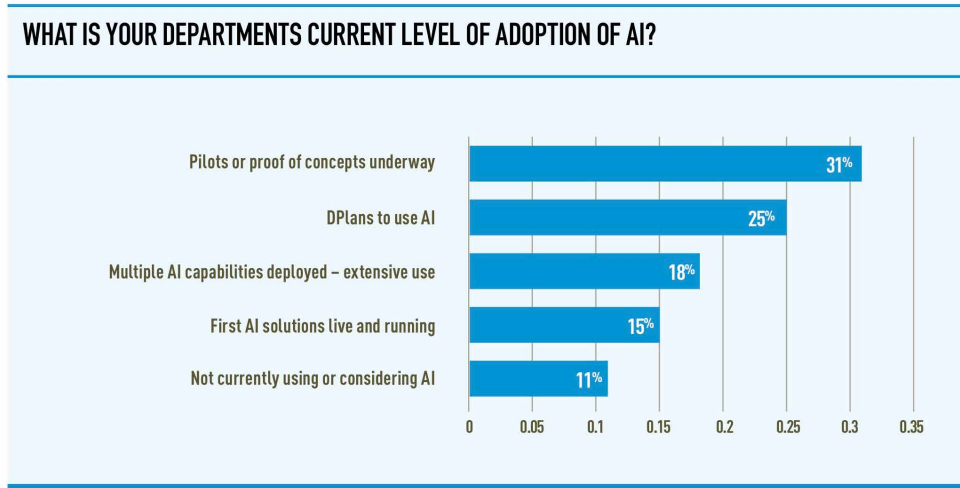


Figure 3. Levels of AI adoption. Source: Finextra and Quantexa "The state of AI in financial services" October 2019

Further analysis indicates that it is not always the more obvious AI applications, such as chatbots and Natural Language Processing, that are being adopted but instead the automation or assisted decision making, i.e. Decision Intelligence.

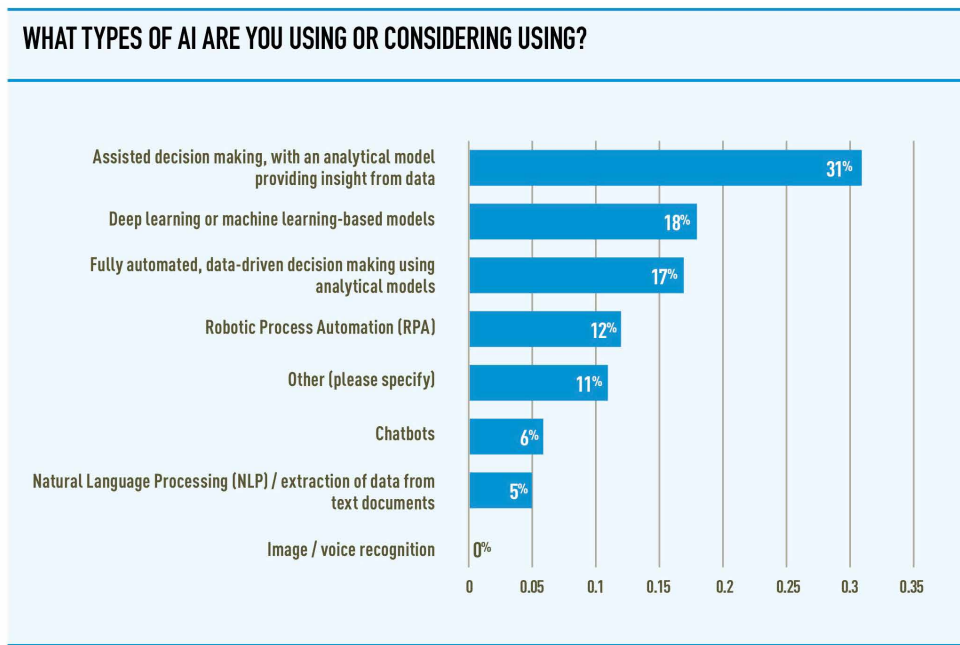


Figure 4. Types of AI business are using or are considering using. Source: Finextra and Quantexa "The state of AI in financial services" October 2019

Adoption within Tax Authorities has been mixed, with some countries having more sophisticated adoption of model based decisioning. However, there have been varying levels of penetration across the different potential use cases and rarely has it been applied earlier in the cycle, through upstream decisioning, to minimize losses and effort to recover any losses.

1.3 Why now?

Model-based decisioning or AI is not a new concept; it has been around for half a century. So, why is adoption accelerating at such a pace?

To understand this, you need to consider the three core ingredients that drive AI:

1. **Data is critical** – There a huge volume of data in a digital format. Studies now estimate that global data is doubling every two years. For example, some tax authorities are collecting invoice data from companies.
2. **Compute capacity** – AI is very compute-intensive, it would have been unthinkable to have thousands of users on a system running models against more than 50 billion records with response times in seconds. Now this is a reality. This has been made possible by cloud-based architectures and containerized compute capacity.
3. **Digital channels** – This allows the models to interact directly with individuals and organizations at the point the decision is required, empowering immediate event-based decisioning and eliminating the workflows associated with batch processing. New Internet-oriented businesses are also driving the expectation of immediate answers on demand, making a human in the decision loop less desirable.

This perfect storm is driving rapid adoption as the business benefits can be significant. The availability of cloud compute capacity at the push of a button, such as Microsoft Azure, makes time to deploy significantly shorter than procuring hardware and infrastructure. Furthermore, analytical platforms as a service such as Microsoft Azure HDInsight creates ready-to-go capability without having to hire teams of people to integrate and manage analytical environments.

2. EU Tax department use case example: VAT Carousel Fraud

Background

- EU Tax department reduced its VAT carousel losses by 98%. This means a 3.5% increase of the collected VAT.
- This is achieved by implementing an ultra-early detection using contextual decision intelligence and analytics.
- **Contextual decision intelligence** is a holistic approach to data that involves focusing on network patterns and not on a confined area.
- Pilot was delivered in three months and production system in one year.

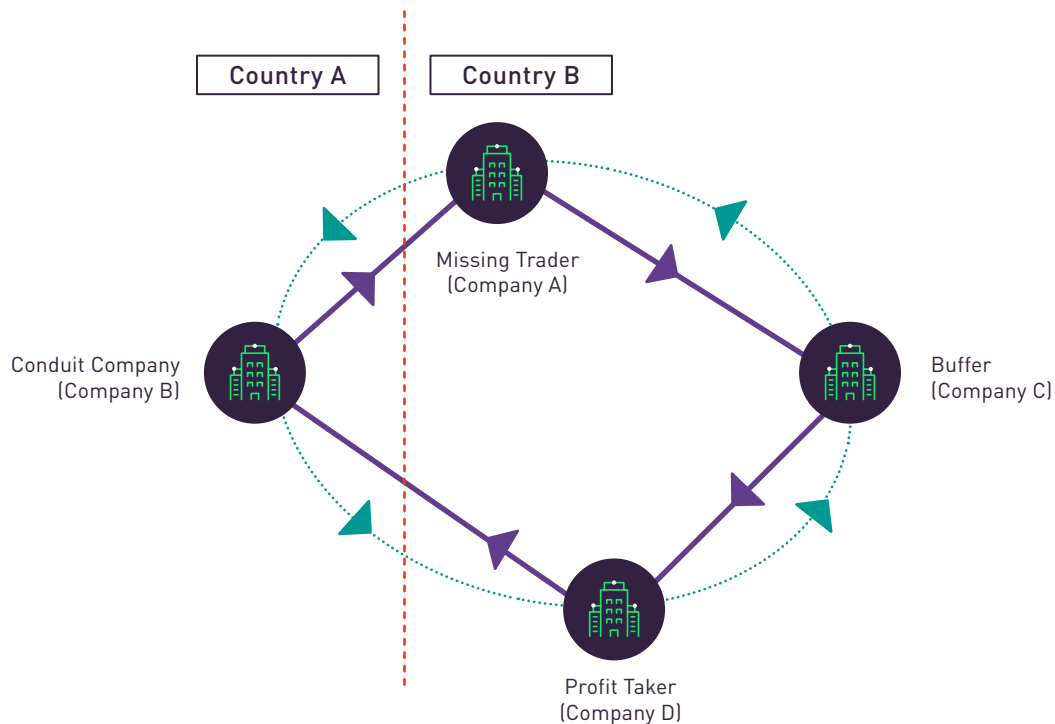


Figure 5. Example of VAT carousel fraud

To create a **carousel**, a criminal will create or buy companies to acquire goods – in this example, this is *Company A*. These companies will then buy goods from a company in the EU (*Company B*). *Company A* will then purchase goods from *Company B*, with no VAT being due, and then sell these goods to a different company (*Company C*). This company could be a legitimate one, or a company managed and set up by the criminals. Here, VAT will be added to the sale price.

Since the goods have been sold domestically, *Company A* must legally declare the VAT it has included on the goods and pay this to the Tax department. However, in the case of a missing trader intra-community (MITC) fraud, *Company A* will vanish and the VAT will not be paid. The goods will then be passed through many other companies which function as a buffer so as to complicate any potential investigation. Next, a company (*Company D*) will send these goods in another EU country, reclaiming the VAT that was not in fact ever paid to the Tax department. This loss of tax revenue negatively affects the provision of crucial public services for the country including education, policing and healthcare, and also increases the tax obligations of honest taxpayers.

3. Best practice approach: Using the intelligence system

This section touches on some of the learnings that have enabled organizations to be successful in harnessing the value of data and analytics. The first step is to analyse the challenges that hamper projects. Then, to understand the latest learnings, approaches and best practice, consider the value of cloud-based approaches and finally some guidance to the steps that allow organizations to truly become data-driven.

3.1 Challenges and pitfalls

One of the major stumbling blocks for decision intelligence approaches has been the failure to truly operationalize solutions. Too often models are built within a lab setting and never see well-integrated deployment. Frequently, the models only generate flat files of insights that are used as an input into some complex process. This is the **data-decision gap**. Below is a breakdown of the common challenges that can be solved by adopting the right approach.

People

- How to ensure adoption of what staff could consider as “threatening AI”
- When there is a “human in the loop”, how to make them more efficient and remove the laborious tasks?
- Are the skills required to deliver the operational AI available and transferable?

Process

- How to operationalize AI models created by data scientists
- How to make a significant impact and create compelling business benefit?
- How do we avoid large complex ETL type projects and achieve quick wins?

Models (AI)

- Poor model accuracy means benefits are not achieved
- Will the model be transparent, explainable and non-biased?
- Can a model make use of internal and external data and other context?
- Is it possible to co-exist with existing models?
- What if we do not have enough consistent known outcomes to train a model?
- Can a model understand and respect data privacy consent?

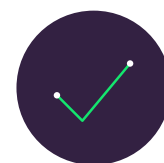
Data

- Data quality impacts model performance – can we deploy effective AI?
- No single view across internal and external customer data
- No full networked context across customers / suppliers / employees
- Data lakes have struggled to deliver value – became data swamps
- Duplication of data for multiple use cases is very expensive – can this be avoided?
- How do we move to adopting pseudo-anonymized data?

Technology

- How to build a scalable platform to handle the huge volumes of data?
- How do we ensure new capability can co-exist with current investments?
- How will the platform continue to embrace future innovation?
- Large volumes of sensitive data: how to ensure a granular security model?
- Will the approach be compatible with cloud strategies?

Common challenges that can be solved by adopting the right approach.



3.2 Contextual Decision Intelligence

In an ideal world, the model should be integrated in real time into the decision cycle, utilizing all the input data to bridge the data-decision gap. *Figure 6* provides an overview of how the cycle should operate and, in particular, the cycle that embeds continuous improvement and deployment of new models.

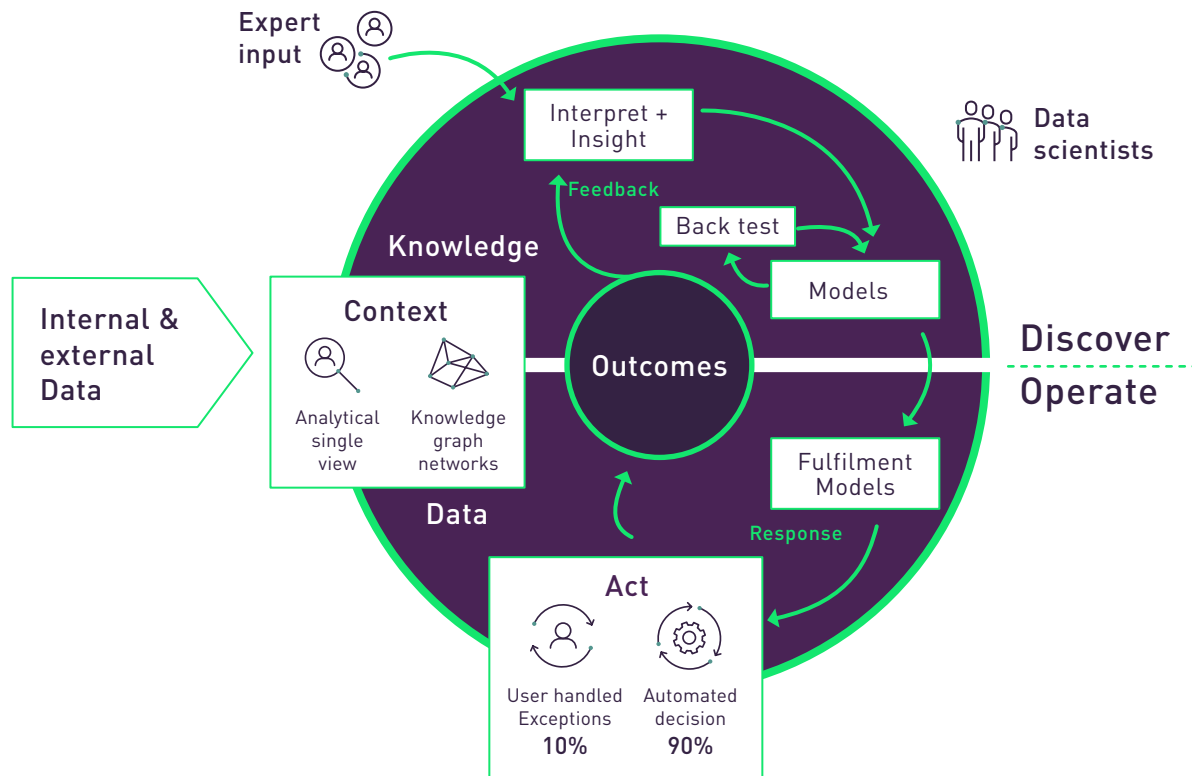


Figure 6. The Contextual Decision Intelligence cycle

There are two phases to the creation and management of ongoing models used for making decisions:

- **Discover** – the process of developing the model
- **Operate** – the ongoing running of the model to produce the desired outcomes

Regardless of the phase, a critical factor in the operational success and accuracy of the model is the availability of **context**.

Context has two elements, the first being the analytical single view of an entity, whether this is a citizen or a business. The analytical single view is generated by resolving the entity across many different records – some of which are internal to the organization and others may be external, such as third-party enrichment data.

The second element is the network that surrounds an entity. For example, when looking at an organization – who are the directors, shareholders and employees? Who does the company trade with? This context needs to be available to the data scientists as an input to the models they create (typically in a batch environment) as well as once the models have been deployed to operational production (where they may need to be generated in real time).

Stages of the Contextual Decision Intelligence cycle

Interpret + Insight – Data scientists work with experts in the field, such as tax inspectors, customs experts or intelligence specialists, as well as working with the data to build up an understanding of how a model could work.

Models and back test – The models are developed using any of the data scientists' favorite tools, such as Python, R, SAS or Scala, and are then back tested. This is normally performed in a batch environment, such as Spark backed by Hadoop. Critically, the data is already prepared with context to build analytical single views and networked knowledge graphs in the data.

Fulfilment models – These are the data scientists' models once they are promoted into the operational environment where IT takes control and ensures all the necessary governance and compliance to GDPR and similar regulations. Here, the models are executed either in batch to create alerts or in an event-driven manner where they score in real time. They also require the same context as an input to ensure the highest quality of decision is made as a response.

Act – This is where action is taken in response to the model. Ideally, in the majority of cases, it is desirable to have a real-time answer, such as a straight through process of a transaction, approval or rejection. However, in some cases, it may not be clear cut. The response is referred to a user who can review before recommending an outcome. This is where it is important to present to the end user with the bigger picture of the context, providing all the supporting intelligence highlighted to enable the user to make a rapid but informed decision.

It is this **contextual decision intelligence** approach that ensures high quality, explainable models can be operationalized to great effect and address the challenges and pitfalls discussed in this document.

3.3 The right technical architecture

Critical to the initial ongoing industrialization of a decision intelligence approach is to adopt the correct flexible architecture approach. As discussed, the volumes of data and compute capacity within the discover and operate phases of the cycle can be considerable when operating at national scale. The best practice answer is the adoption of cloud-based approaches, such as Microsoft Azure and the Azure HDInsight capability. These offer a multitude of benefits:

- They allow compute power to be rapidly mobilized, decreasing time to value
- They can be scaled horizontally, allowing limitless futureproofing for scaling the solution
- The containerized approach offers best practice for DevSecOps development and deployment cycles ensuring minimal risk of cyber-attack and data loss
- There are secure templates defined, removing the significant burden on your organization's IT staffing needs
- Using platform managed capabilities, such as Azure HDInsight and Azure SQL, you can be ready to go with architectures suited to analytics as well as decreasing the ongoing support costs
- This type of controlled environment offers support for separation between production and development activities and is GDPR-secure by design, with support for pseudo-anonymized data. Data does not have to be transferred to the data scientists own machines
- By having a central capability that resolves the data quality issues and generates the context, data scientists do not have to spend all their time preparing the data
- The overall total cost of ownership is significantly reduced
- It is possible to go from Proof of Concept to live operating in a few months
- The platform and approach are future proofed as open architecture and open source approaches support a multitude of free and commercial best practice analytical tools

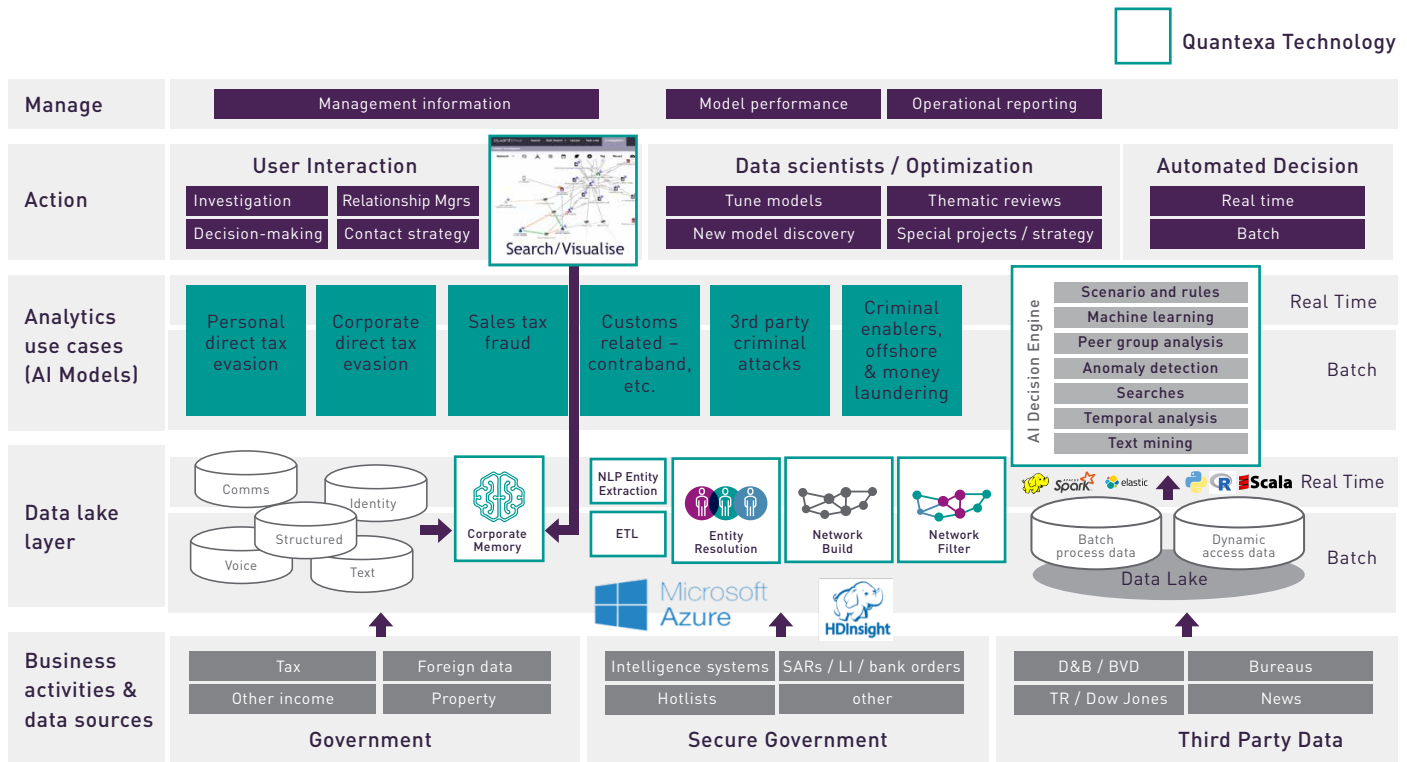
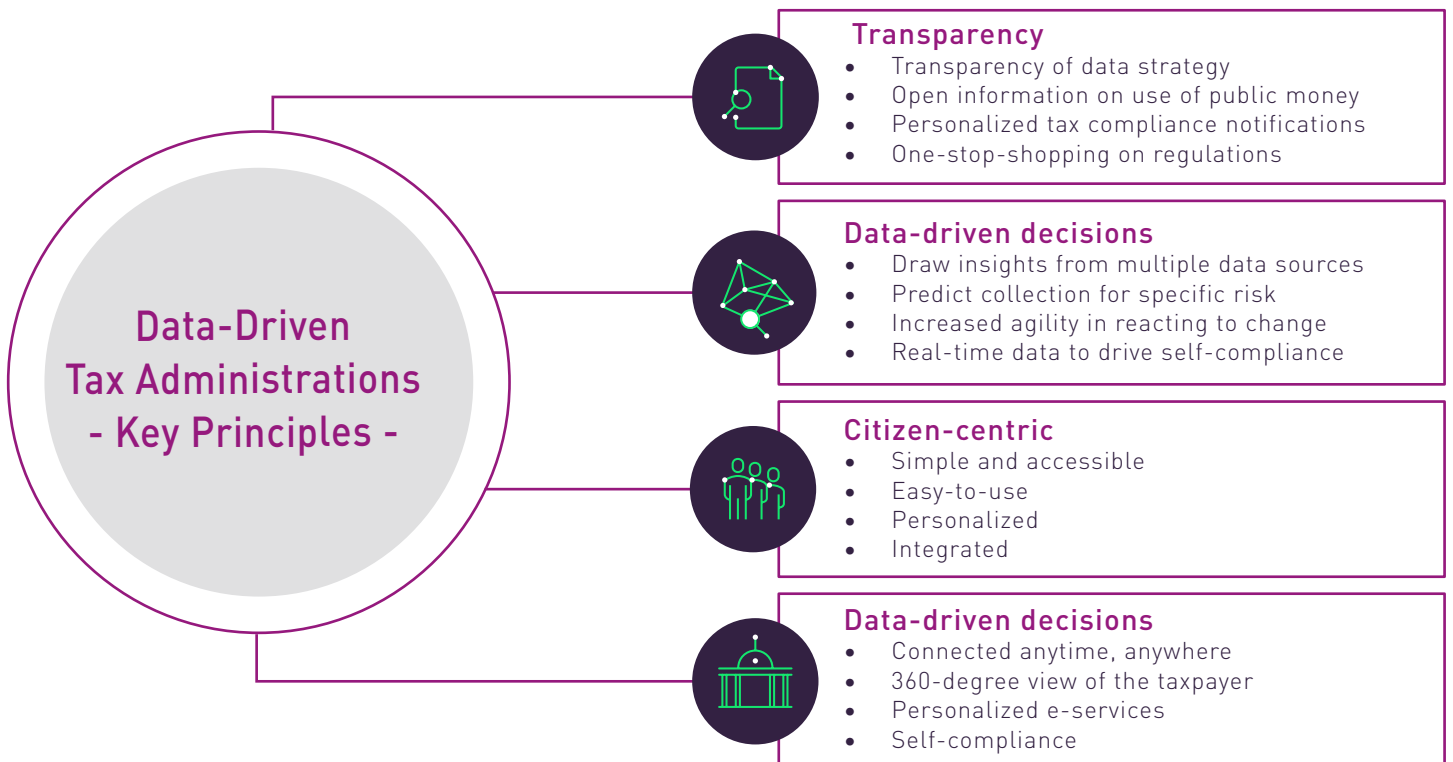


Figure 7. An overview of a typical deployment architecture for a Contextual Decision Intelligence platform environment

3.4 Conclusion

AI changes the way we think about citizen-centric government. Data and AI adoption are governed by four key principles when applied to the transformation of tax administrations:

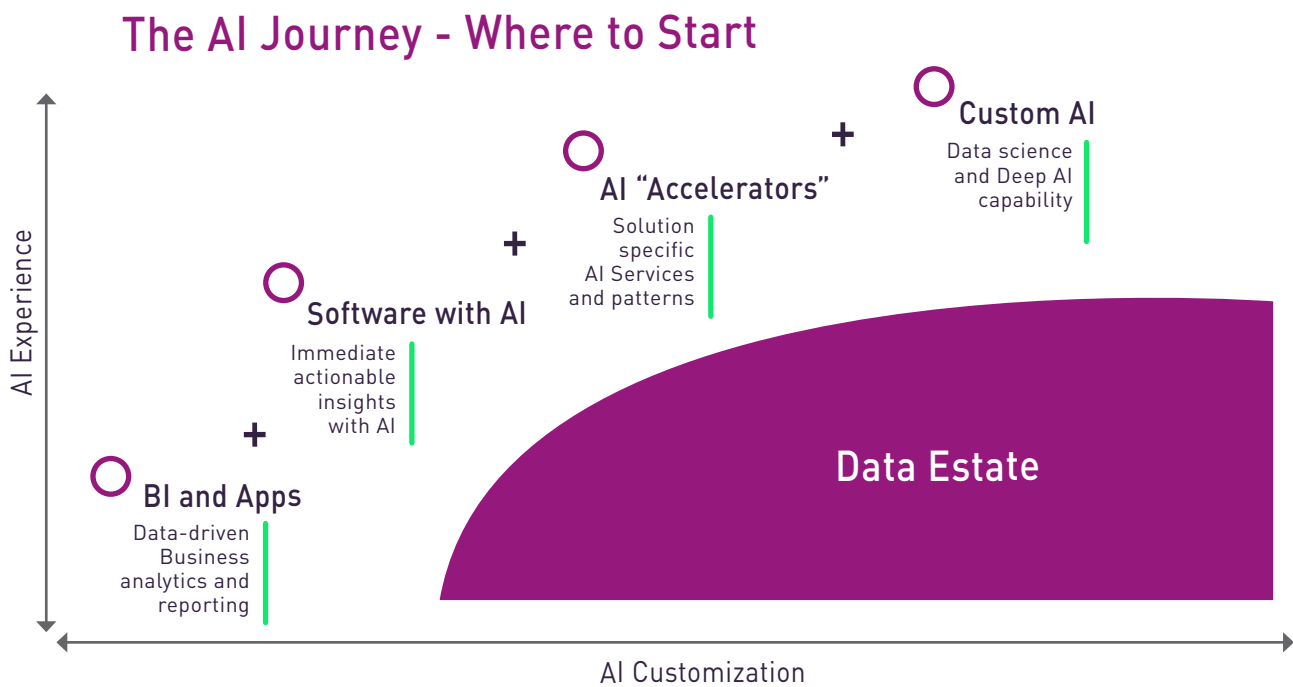


Many forward-thinking organizations are already deploying AI to support their citizen service efforts, raising the bar for the entire industry. Implementing AI helps to meet the rising taxpayer expectations, enabling customers to get answers at any time and in any channel.

Tax administrations are deriving some of the following benefits from adopting AI:

- Maintain and enhance knowledge and expertise within your teams
- Gain speed and improve on resolution time, collection and refund time
- Cost reduction from improved efficiency and automation of services
- Gain flexibility through elastic compute capacity that adapts to your seasonality and forms a hybrid interoperable platform
- Enable a real-time engagement with your taxpayers for improved compliance

The roadmap to AI adoption



There is no one-size-fits-all application of AI for customer service. Every organization is at a unique place in their digital transformation journey and has distinct needs, so naturally they will deploy AI with differing levels of sophistication. As organizations grow and mature, their needs will change, so it's critical that their AI solutions are flexible enough to scale and update accordingly. Regardless of where your organization is on its digital transformation journey, AI has the potential to deliver incredible value to you and your customers.

The most important part of AI development and adoption is the ethical frame.

What is **ethics** in AI?

- A responsible, principled approach for the delivery of AI technologies and solutions that governs our thinking and behavior
- The ability to consider the societal and human impact of the technology solutions we create and influence governments and the industry
- An approach that enables us to create a culture of purpose that empowers our people to produce solutions with positive outcomes

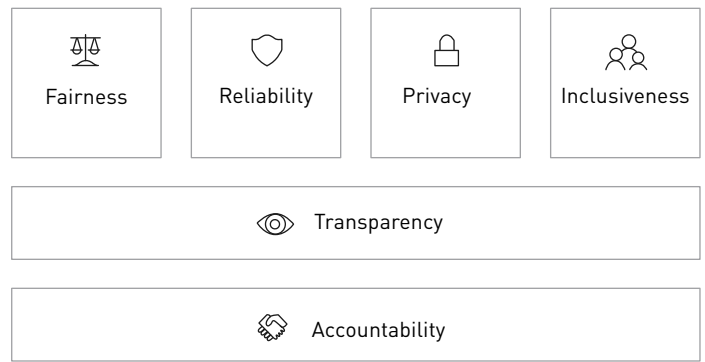


Figure 8. Microsoft's six principles on which all AI systems need to follow

3.5 Getting started

The first step is to identify a good candidate project to pilot the adoption of AI. *Figure 9* provides a guide to selecting the right pilot and making it a success.

Selecting your first pilot

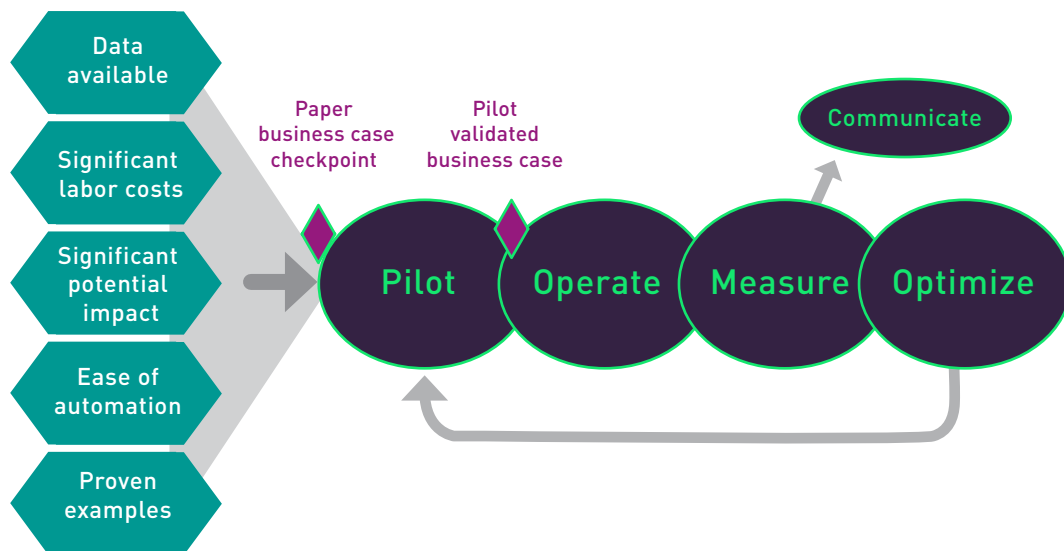


Figure 9. Selecting a pilot

Once you have demonstrated to the business the value that can be achieved using **contextual decision intelligence**, then consider your programme towards building a data-driven organization with parallel streams, as indicated in *Figure 10*.

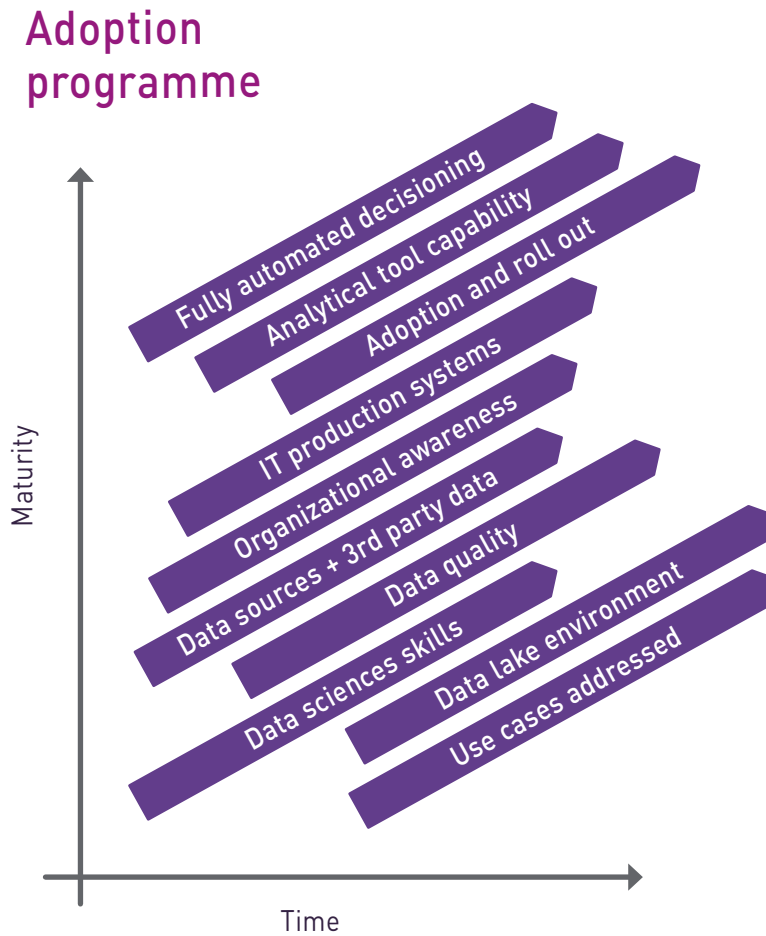


Figure 10. Adoption programme

The data, analytics and AI revolution is reshaping the way organizations make decisions. Governments must proactively innovate and adopt new technologies to transform business operations and leverage the huge quantities of data produced and collected by public sector bodies. Within the tax domain, a decision intelligence capability using AI and machine learning can improve taxpayer and employee experience, as well as enhance tax compliance. However, you must ensure you adopt the right approach to avoid common pitfalls that result in organization's failing to truly operationalize solutions. This approach is **contextual decision intelligence**.

To find out more on how to maximize the value of your data and effectively adopt AI using high quality and explainable models, speak with an expert by clicking the links below.

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

[Speak with Microsoft](#)

[Speak with Quantexa](#)



About Microsoft

Microsoft is the leading platform and productivity company for the mobile-first, cloud-first world, and its mission is to empower every person and every organization on the planet to achieve more. Our vision for government organizations is to enable them to better serve and protect citizens and build more secure, productive nations by connecting people, systems, and information resulting in more impactful outcomes.

[Learn more here](#)



About Quantexa

Quantexa's Contextual Decision Intelligence (CDI) platform is a new approach to data that gives organizations the ability to connect internal and external data sets to provide a single view, enriched with intelligence about the relationships between people, places and organizations. By generating context, Quantexa automates millions of operational decisions, at scale, to effectively and efficiently identify complex compliance issues and potential security threats.

[Learn more here](#)

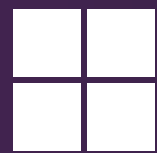
© 2020 Quantexa Limited. All rights reserved. This document is provided "as-is." To the maximum extent permitted by law, Quantexa excludes any warranties, express or implied or statutory including without limitation, merchantability, accuracy, fitness for a purpose, satisfactory quality and non-infringement. Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. This document does not provide any legal rights to any intellectual property in any Quantexa proprietary information or products. You may copy and use this document for your internal, reference purposes.

© 2020 Microsoft Corporation. All rights reserved. This document is provided "as-is." Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. You bear the risk of using it. This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.

A dark purple background with a light purple network pattern of interconnected nodes and lines.

quantexa

Connecting data | empowering decisions



Microsoft