

SHINING A LIGHT ON DARK DATA

A new approach to data extraction,
ingestion and analysis for the insurance industry

A perspective from Financial Services Technology Advisory



Automation is here to stay. Businesses are no longer just talking about automation or experimenting with automation; they are embedding automation as part of their long-term strategic objectives.

It is estimated that spending on automation (artificial intelligence and cognitive computing) will more than double from today's levels to \$97.9 billion in 2023.¹ For insurers, looking ahead the big question is how to automate at scale.

One of the great benefits of automation is in helping companies realize Straight-Through Processing (STP). This is key to streamlining end-to-end processes, such as claims processing, by reducing human intervention. Historically, STP has been considered an unrealistic objective, with businesses opting for human-in-the-loop and partial process automation. This is due primarily to the limiting factor of "Dark Data": data that the business possesses but cannot access, like photographs and handwritten documents. With new approaches to data extraction, and by leveraging a suite of powerful technologies like artificial intelligence, we are now in a position to unlock Dark Data and allow businesses to pursue STP, realizing the power of automation.

Once unlocked, Dark Data not only permits greater benefits from automation, it also feeds and enriches the models and analytics that create sales opportunities, measure risk, and drive business decisions. This creates a secondary value stream that can differentiate a business in the data-driven financial services market—particularly in the information-intensive insurance sector. It also creates the opportunity to reassign people resources to higher value activities and initiatives that support the insurer's growth agenda.

In this paper, we will explain the challenges posed by Dark Data and present insurance-industry-specific solutions. Subsequent papers will look at solutions applicable to the banking and capital markets sectors.

STRAIGHT-THROUGH PROCESSING IN INSURANCE

The benefits of intelligent automation have become particularly clear in the insurance industry. The availability of technology that can cut costs, reduce error rates, improve customer experience and speed of service, support compliance, and reduce monotonous tasks for human operators—while avoiding the prohibitive cost of legacy system changes or migrations—has made intelligent automation a “must-have” in the toolbox of any insurer.

Traditionally, insurers have automated individual, simple, stable processes with predictable inputs and predictable outputs to avoid the complexity of automating across the lifecycle of an account, policy, or claim. However, this approach has necessitated the inclusion of partial process automation, with handoffs between human operators and robots through various mechanisms.

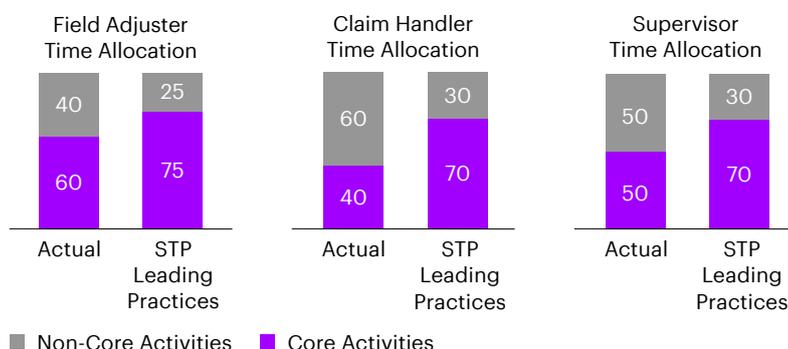
For example, leading segment software provider UI Path, Inc., suggests targeting only 70 percent of a process for automation, stating that it is usually necessary to “leave some steps of a process to human intervention.”²

While partial process automation can simplify deployment, the manual interventions and handoffs required introduce waste into the process (in terms of time, cost, and accuracy), eroding the business case for automation technology. Only when a process is automated straight-through are the full benefits of the technology realized. By automating entire processes end-to-end and providing a no-touch or low-touch approach to operations, handoffs are eradicated, further diminishing the risk of error, reducing handling time, and, as there will no longer be time spent ingesting or creating manual handoffs, creating efficiencies in processing capacity and significant improvements in the customer experience.

Figure 1. Straight-Through Processing Benefits in Claims

Claims Operations and Efficiency

Based upon Accenture’s experience, intelligent automation can reduce non-core activities for Claim Handlers and Field Adjusters (~40%-60% of their time), allowing them to focus on higher value core activities and reducing the number of Claims Supervisors needed for oversight.



Source: Accenture experience and work in this area

Projected Automation Benefits - Example

6000

Full-time Equivalent (FTE)

40 – 60%

Typical time spent on non-core activities

10 – 15%

Non-core document processing activities

160 – 540

FTE capacity freed by intelligent automation + enablers and can be reinvested in higher value activity

6 – 20M

USD savings assuming 40K blended cost per FTE

THE DARK DATA CHALLENGE

In insurance, automation clearly provides great efficiency gains when it is used to realize STP. Yet many insurers (and vendors) shy from targeting no-touch automations. One of the primary roadblocks is the issue of “Dark Data.” Up to 75% of a corporation’s data may be Dark Data.³

Dark Data is data that an organization has but cannot make use of. This includes both data that is not currently accessible to human or automated operators, as well as data that is accessible, is used by humans, but cannot be directly interrogated by systems and automation technologies. All forms of Dark Data can present roadblocks to STP.

Consider the example of a photograph that shows the scene of a car accident, as well as capturing the license plate, color, and severity of damage to the vehicle. There may be a street sign in the image indicating the location of the accident, and the image may show whether it is day or night, and whether the road is wet or dry. This information on its own could be enough for an insurance company to open a claim and produce an initial estimate of damages, determine whether to dispute the claim, and assess the probability of fraud. However, because that data is locked inside a photograph and is not presented in a structured format, it is not captured by the business; the image is indexed for future reference and the data is left on the shelf forever.

In Operations documents



20%
handwritten



45%
poor image
quality



66%
unstructured

In Front Office documents



25%
handwritten



69%
poor image
quality



41%
unstructured

Source: Based upon Accenture’s experience and work in this area

Dark Data in insurance can take many forms, including:

Structured Documents

Structured forms are used in insurance to gather information in a way that is easy for people to digest and identify key information. This information is often presented in tables or boxes with single-word descriptors as the only reference for what is contained in that field.

Unstructured Data

Unstructured data is everywhere in the insurance industry. Photographs, witness statements, legal documents, and medical records presented as part of an insurance claim, file notes, or even transcripts of call center conversations all constitute unstructured data. This data generally presents itself as written blocks of text with no clear indication of where certain key information is held.

- **Handwritten Documents.** Of the structured forms received, businesses can expect that between 20 and 25 percent will be handwritten. Handwriting has posed one of the greatest challenges to optical character recognition (OCR) technology for many years, with complexity driven by the fact that every person has their own style of handwriting. As a result, writing rules to differentiate one person's "o" from another person's "0" has been impossible to date.
- **Emails.** Emails are one of the highest volume types of communication in every business, used both internally for day-to-day process management, and externally for communications with customers. The time spent manually reading, classifying, routing, recording, and extracting information from emails is extensive.
- **Voice.** The conversations that agents and call center operators have with customers are important. Not only do they help build a relationship, but they also involve the oral transmission of a significant amount of data. Commonly, call center operators, claims handlers, or other employees make short-hand notes in customer files denoting what has been discussed on a call, but this high-level information only skims the surface of the data available during these calls.
- **Images.** Images and videos represent the next generation of largely untapped data sources. Due to the challenges presented by Dark Data, insurers have largely shied away from pursuing the generation and receipt of image-based data. Of those images that are received, data is rarely captured in an automated fashion; images are often manually classified and archived with the data they contain rarely being released.

Because critical elements of information about people, property, and policies are locked inside data that cannot be readily accessed by bots or systems, automation primarily takes place downstream from where the data first touches the organization. Automation is implemented only once a human operator has had the chance to touch the data and transform it into a structured format. That time spent manipulating and interpreting data increases the risk of manual error and erodes any benefit that the downstream automation can produce.

Businesses have employed several methods to work around the blockage caused by Dark Data, including the implementation of bots triggered by manually produced, templated emails, portals or applications for customers to fill in as their first point-of-contact, or complex and convoluted business rules to extract and route information. Each method has substantial downsides; for example, manual triggers produce waste and run the risk of potential errors, portals require a change in customer behavior that may be unwelcome and unwanted, and business rules are often imperfect and inflexible, creating errors and more manual work to fix.

Businesses often turn to human workers to unlock their Dark Data because there is no single tool or vendor in the marketplace that can solve for all types of Dark Data. For example, many vendors offer OCR, but few can accurately extract text from low quality, unstructured, or handwritten documents. Similarly, many offer transcription services, but few can offer highly accurate, fully automated, speaker-specific transcription and associated call center analytics. And many offer image recognition, but few can offer the flexibility required to solve specific business problems. Finding a tool that works well for a given problem is hard; finding a tool that works well for every problem has, up to now, been impossible.

HOW BIG IS THE IMPACT OF UNLOCKING DARK DATA?

The impact and benefit of unlocking Dark Data to pursue STP is enormous: by automating end-to-end, insurers can increase efficiency gains, reduce risk, and realize the potential of intelligent automation technologies. However, in the era of data-driven businesses, the benefit from unlocking Dark Data stretches far beyond the empowering of STP.

Insurers are increasingly making use of big data analytics and artificial intelligence to drive business decisions, generate insights, and identify risks. Based on Accenture's experience, insurance companies that are leaders in the analytics space can see an increase in profit of up to 21 percentage points.⁴ By unlocking Dark Data, insurers can obtain additional analysis and insights previously inaccessible due to a lack of usable data, providing them with a competitive edge when it comes to sales, risk management, and claims fraud prevention.

Within existing processes and automations, downstream human operators are often tasked with extracting specific information from physical and unstructured documents in order to complete the process. However, because of the time required and cost associated with this manual effort, only the minimum information required to complete those processes is ever extracted, leaving a significant amount of information on the table.

By unlocking their Dark Data, insurers can gain access to a significantly greater range of data points against which they are able to measure risks, identify opportunities, and make business decisions. Where today, human operators may only extract roughly 5-10 fields on a page in order to complete a specific process, the remaining 50-100 fields can be extracted and made ready for analysis by applying new methodologies for unlocking Dark Data. Multiplied over millions of documents across an enterprise, this improvement amounts to a paradigm shift in insurers' ability to generate insights into their businesses and customers.

Insurance Use Case: Straight-Through Claims Handling

An insured was just in a car accident. They use their insurance carrier's mobile application which has the insured take pictures of the accident. After the insured uploads the pictures to the app, a pre-trained model is used to assess the damage to the vehicle, confirm that the insured is covered, and estimates the potential cost of the repairs. The model suggests the vehicle appears drivable, so the system will notify the customer of nearby auto body shops that the carrier has worked with in the past. Based on historical data, the system verifies whether the final cost of repairs match the expected cost to repair the damage sustained. The system automatically sends a payment to the insured and closes the claim.

DARK DATA: THE ACCENTURE APPROACH

Accenture has developed a multi-tool capability to help insurers unlock Dark Data, making use of advanced technologies to extract valuable information from unstructured documents, documents with poor image quality, handwritten documents, voice-based interactions, and images.

Insurers applying Accenture’s framework not only permit STP and better leverage automation, they can also generate additional benefit streams (see Figure 2). Accenture’s approach to Dark Data combines these major elements:

1. Use Case Definition and Enhancement
2. Tool Selection
3. Concept Verification and Refinement
4. Business Process Transformation

1. Use Case Definition and Enhancement

Accenture has a market-tested framework for identifying, sizing, and understanding an insurer’s business and its Dark Data challenges.

By performing an in-depth analysis of documents received, channels, content-management systems, existing routing and extraction technology, and the associated processes and indexing requirements, Accenture can help insurers build an accurate picture of the Dark Data landscape within their enterprise, which can further inform the solution design and tool selection outlined below.

Figure 2. Accenture’s Dark Data Assessment Framework Output - Example

KEY STATISTICS

of mailroom FTEs: **230**
 Current Electronic Document Management (EDM) solution budget (annual): **\$1.2M**
 Avg # Documents per Envelope: **1.9**
 Avg # Documents per Fax: **1.4**
 Avg # Documents per Email: **3.2**
 Avg # Attachments per Email: **2.4**
 Avg # Pages in a Document: **5.1**
 Avg # of document variations: **23**
 Percentage Handwritten: **28%**
 Image Quality
 H: 50.2%, M: 33.8%, L: 16.0%

KEY FINDINGS

Seasonality impacts: 60% of volume received within first 5 business days of each month
 State mail forms updated on the first business day of the calendar year
 10% of all packages contain checks

DOCUMENT TRENDS

Top 10 document types

	# of versions
General Liability Application	85
Commercial Insurance Application	251
Loss Runs	91
Endorsement	15
Property Supplement	5
Statement of Values	9
Blank	62
Terrorism Risk Insurance Act	42
First Report of Injury	6
Wage Loss Information	3

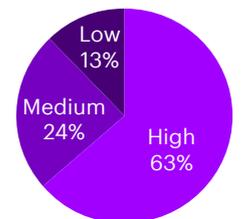
Data type

PDF	Searchable PDF	Other
75%	20%	5%

Handwriting

Fully	Partial	Signature	None	N/A
22%	6%	6%	61%	5%

Image quality



37% of documents have less than 300 DPI resolution

Source: Accenture, January 2020

2. Tool Selection

There are hundreds of vendors and tools that claim to be able to ingest Dark Data. Knowing which tools can yield the best results for a given Dark Data challenge is very difficult. Accenture has identified, tested, and benchmarked specialized Dark Data ingestion tools across the industry to be able to help insurers select the tools expected to be most effective in a given use case (see Figure 3).

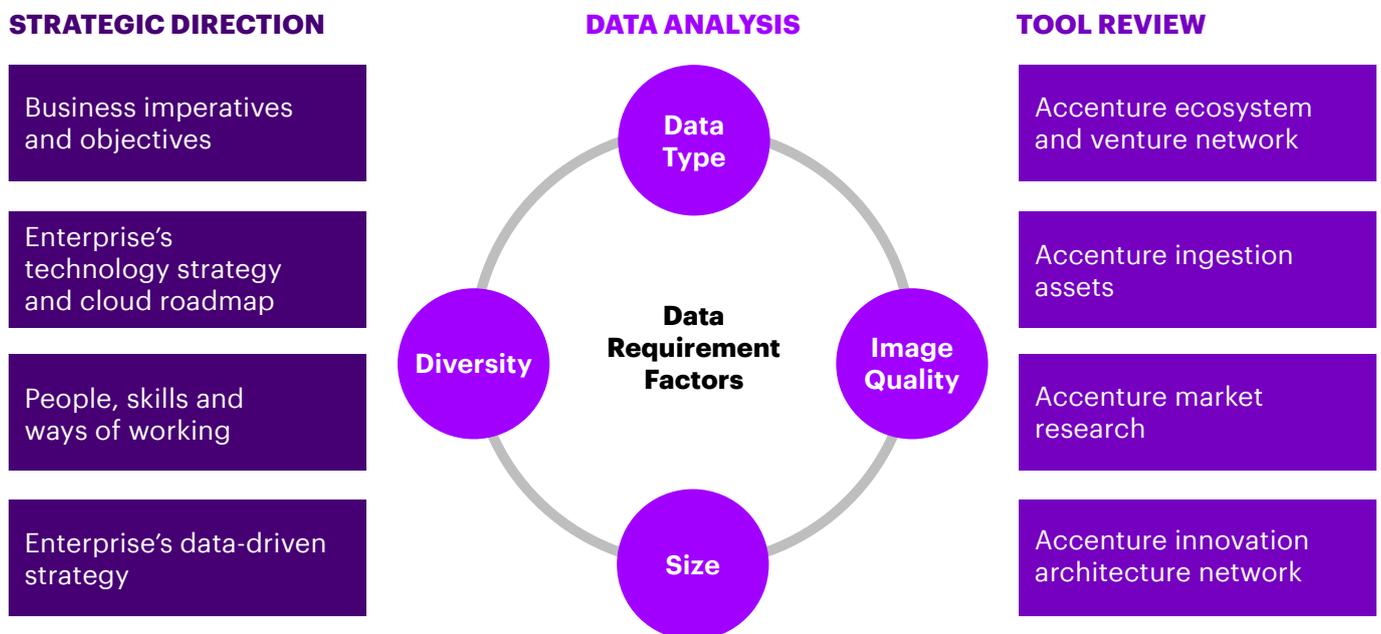
Tool selection should not only consider which tool yields the highest accuracy rates, although this is obviously a core component. Insurers should also evaluate:

- Their overall business objectives and technology strategy, including their underlying data architecture and available data.
- Their appetite for cloud technology over on-premise deployments.
- Whether they intend to implement strategic enterprise-wide ingestion or tactical process-based ingestion.

- The level of analytics that might be desired post-ingestion.
- Existing capabilities within their own Centers of Excellence and their ways of working, including any skills gaps that might be present.

Many companies attempt to solve their Dark Data challenge with a single-vendor solution, and as a result struggle to see transformational results across all types of Dark Data. Accenture links together the right combination of high-performance tools, not just to get the data, but to get it more accurately, more quickly, and to a greater level of depth than any single-tool solution. Additionally, relationships with market-leading ingestion technology providers uniquely position Accenture to leverage the power of these resources beyond just the strength of the tools themselves. This includes the co-creation of deployable industry-standard ingestion models, competitive price negotiation positioning, access to third-party-data, and first-class technical support.

Figure 3. Accenture’s Approach to Dark Data Tool Selection



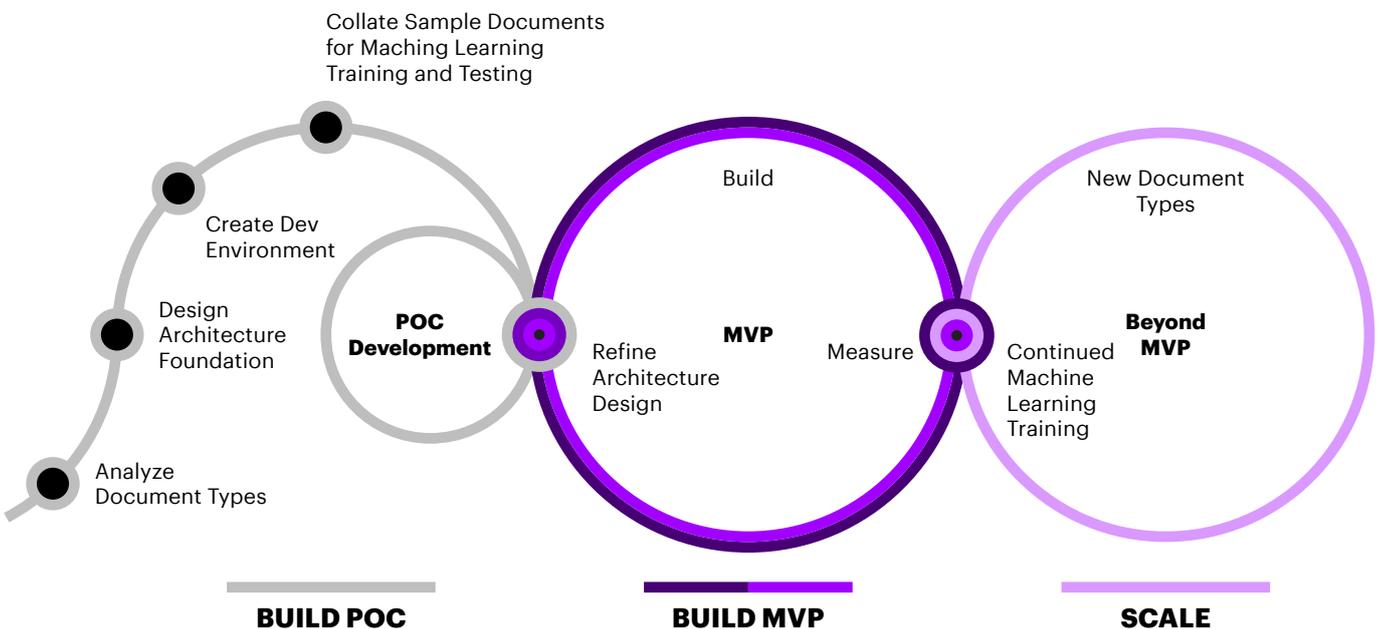
Source: Accenture, January 2020

3. Concept Verification and Refinement

In many cases of industry-standard Dark Data challenges, it is possible to leverage pre-trained models for demonstration and assessment. However, where more specific challenges have to be addressed, a sample of data provided by the business can help develop a custom ingestion solution and train new models to prove the applicability of specific technologies to actual business challenges.

Once a concept is proven, Accenture can combine the proof-of-concept (POC) results and Accenture's Assessment Framework to identify and prioritize ideal opportunities for a Minimum Viable Product (MVP) (see Figure 4). In many cases, Accenture can call on its library of production-ready models for specific document and data types, developed in conjunction with its ecosystem vendors, alliances and partners. However, where the MVP requires additional tuning, development, or training, Accenture can carry out the end-to-end delivery and deployment of a model customized to that specific challenge.

Figure 4. Accenture's Approach from POC to MVP to Scale



Source: Accenture, January 2020

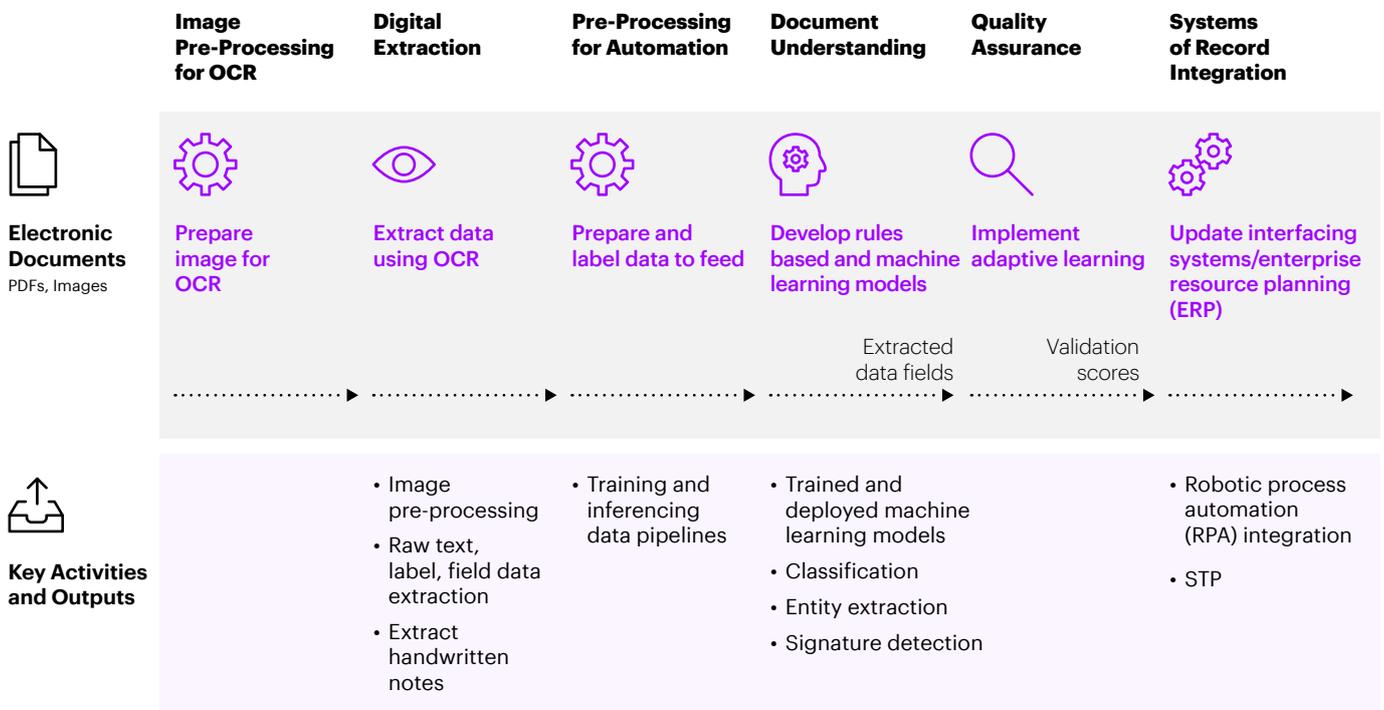
4. Business Process Transformation

Combining enterprise Dark Data solutions with analytics and predictive modeling, process modernization, operations, and automation technologies can help insurance companies access and leverage their Dark Data to transform the way they do business.

Transforming the Technology: This is where data science, model development and enhancement, data security, cloud, and systems integration meet to move beyond MVP and to deliver the “right”

technology solution for the enterprise at scale. Based on Accenture’s experience, enterprise-wide ingestion is effectively tackled by working closely with individual lines of business (LOBs) to apply the enterprise-wide ingestion technology stack to the specific data challenges faced by each line. For each LOB, extraction models can be customized and developed by Accenture following the process outlined in Figure 5.

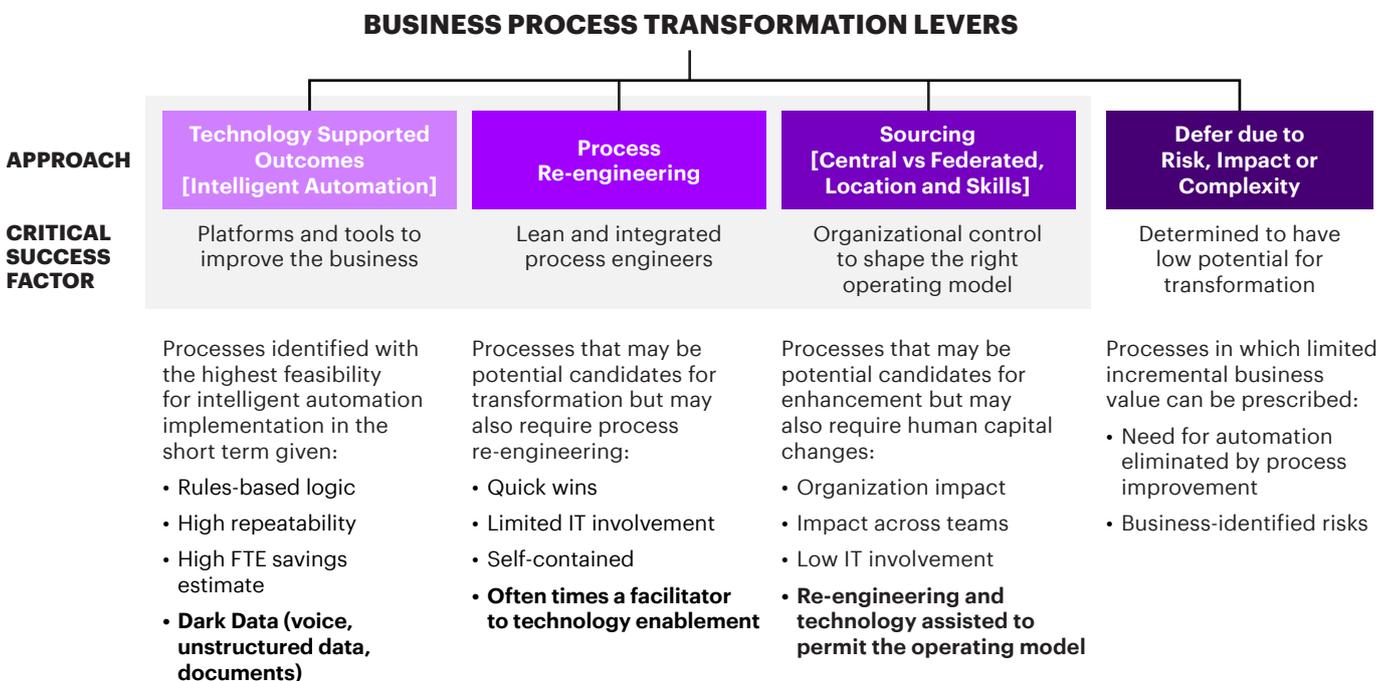
Figure 5. Illustrative Process for Developing Dark Data Extraction Models



Source: Accenture, January 2020

Transforming Processes: Businesses looking to leverage Dark Data to allow STP should understand not only the inefficiencies created by its Dark Data, but also the broader opportunities for process enhancement. Using a leading process-mining technology, Accenture can deliver to an insurer a comprehensive understanding of its process architecture, inefficiencies, and opportunities for improvement. This positions the insurer to leverage Dark Data, process re-engineering, sourcing and automation to permit STP and realize operational efficiencies. (see Figure 6).

Figure 6. Core Process Transformation Levers



Source: Accenture, January 2020

Transforming Operations: Where the scale of a Dark Data challenge warrants transformative document management solutions, Accenture has the operational footprint, technology know-how and capabilities, and system integration experience to provide a full end-to-end document ingestion or virtual mailroom managed service (see Figure 7). Accenture's managed service is orchestrated using market-leading workflow management solutions, and integrates seamlessly with existing business processes, automations, and systems of record.

Turning People into Precious Talent:

Unlocking Dark Data and realizing STP allows insurers to redeploy resources from lower value tasks to higher-value activities, functions and initiatives that create sustainable business value.

People can be assigned to support the drive to innovation, an area often under-resourced among insurers. They can be employed in completing the digital transformation of the business, or entrusted to turn information into actionable insights to support strategic decision making, new product and market development and extend an insurers value proposition to new client groups.

These people resources can also be retrained and upskilled to become the precious talent for addressing skill gaps that limit an organization's ability to grow and prosper.

Transforming Business Intelligence:

Having unlocked its Dark Data, an insurer should have a comprehensive understanding of how to leverage the unlocked data using analytics to create valuable insights.

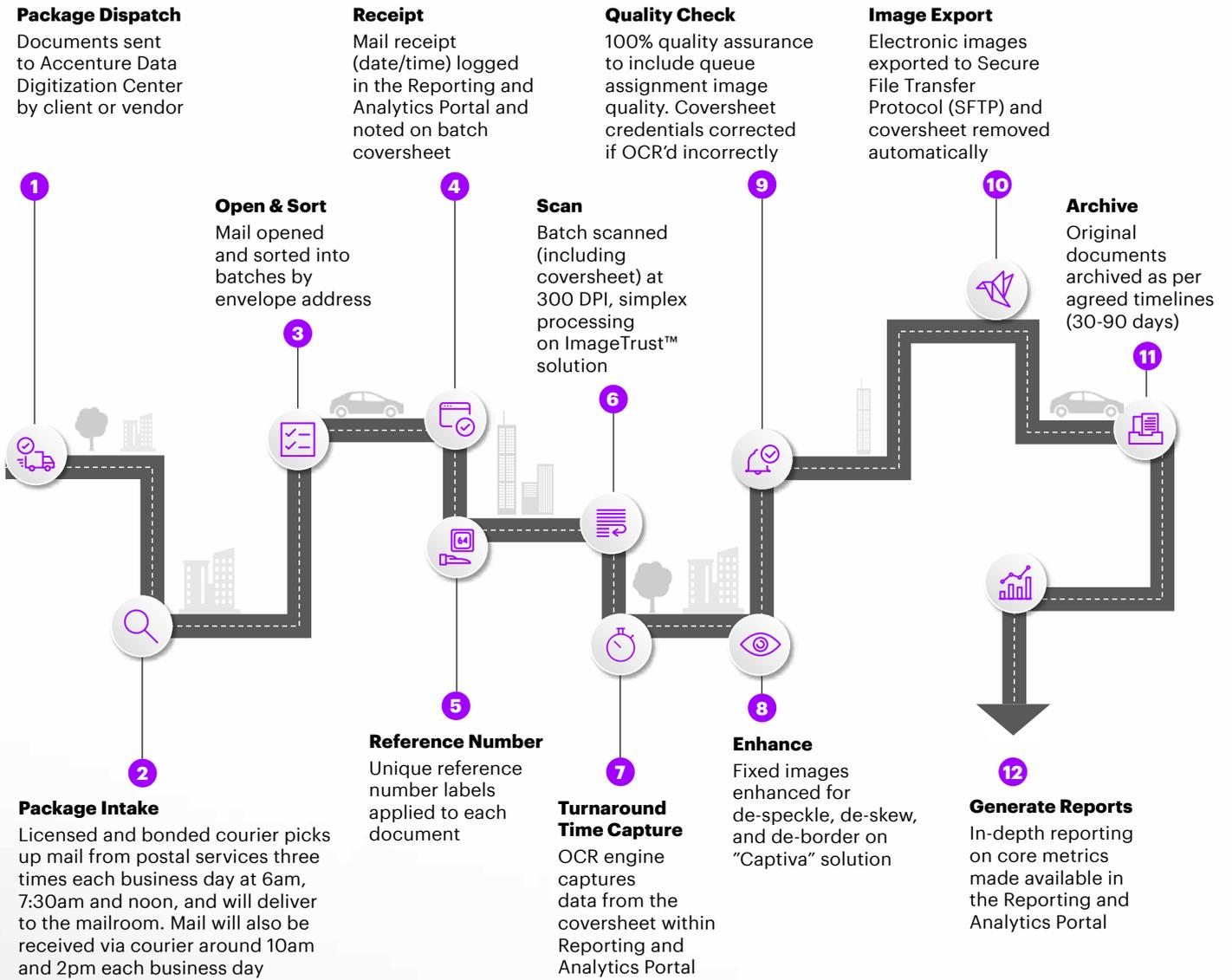
Accenture's approach to Dark Data ingestion means in many cases the data will be extracted by tools that already permit sophisticated analytics, reducing the incremental lift required to generate those insights. Regardless of the extraction tool, however, Accenture will apply a range of analytics and modeling techniques in combination with relevant third-party data sources to generate insights with the power to transform the way an insurance company understands its business.

This includes the application of unsupervised machine learning techniques to enhance customer and claim segmentation, as well as creating powerful predictive analytics to better triage new business, manage claims and reserves, identify opportunities for cross-selling, and enhance underwriter and claim-handlers' decision-making abilities. On top of this, Accenture applies market-leading data visualization software so that insights are accessible and impactful.

For more information on how insurance carriers can make the most of their data by applying advanced analytics, please see our publication "Harnessing the Insurance Data and Analytics Exhaust Stream."⁵



Figure 7. Accenture's Intelligence-based Virtual Mailroom Managed Service – Example



Source: Accenture, January 2020

CONCLUSION

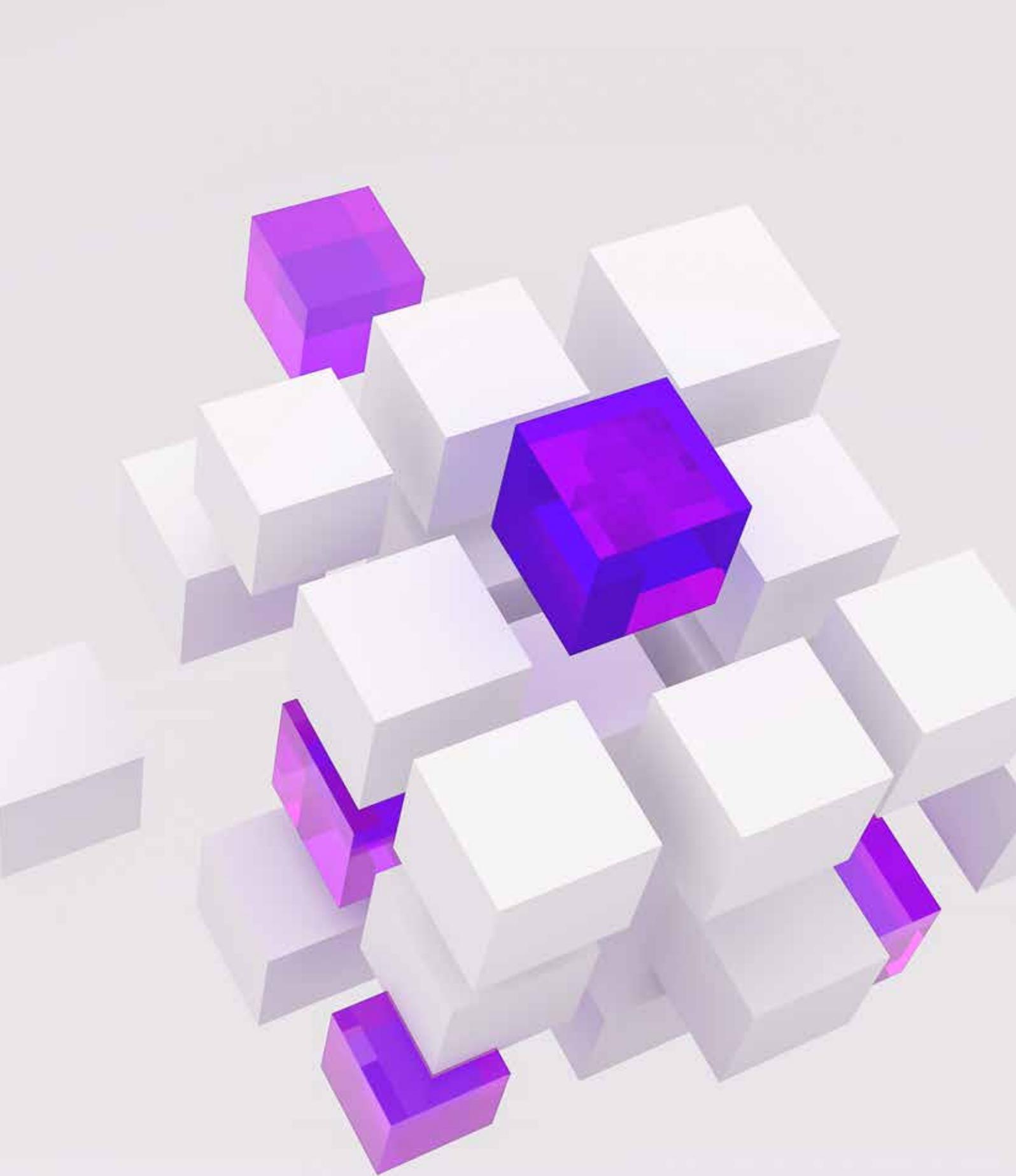
For insurers, Dark Data is a major obstacle on the path to Straight-Through Processing and the development of new revenue streams.

While barriers such as handwritten documents, call center transcripts, photographic records and unstructured medical forms abound, new approaches combining multiple tools can help insurers organize and ingest this data and extract the valuable information it contains.

Accenture's approach to unlocking Dark Data allows insurers to pursue both Straight-Through Processing and advanced analytics by:

- Understanding the scope and extent of their Dark Data challenge.
- Identifying the ideal technology approach for a given Dark Data challenge, existing technology landscape, and set of data and security restrictions.
- Capturing incremental value from extracted Dark Data through advanced analytics and insight-generation.
- Developing proof-of-concept ingestion models to improve their understanding of the Dark Data challenge and hone the business case for ingestion technology.
- Providing end-to-end development of ingestion technology, deployed either within an insurer's own ecosystem or provided as a managed service.

To find out more on how Accenture can help you shine a light on your dark data to strengthen your analytics capabilities and support your growth agenda, please contact one of the authors.



References

1. [“Worldwide Spending on Artificial Intelligence Systems Will be Nearly \\$98 Billion in 2023, According to New IDC Spending Guide,”](#) IDC, September 4, 2019.
2. [“From pilot to full scale RPA deployment – A comprehensive guide to the business transformation journey,”](#) UiPath.
3. [“Dark Data Research Reveals Widespread Complacency in Driving Business results and Career Growth,”](#) Splunk, April 30, 2019.
4. [“Harnessing the data exhaust stream: Changing the way the insurance game is played,”](#) Accenture 2016.
5. Ibid

Authors



Duane Block

Managing Director
Accenture Financial Services,
Technology Advisory
Intelligent Solutions Global and
North America Lead
duane.j.block@accenture.com



Geoffrey Hodgson

Manager
Accenture Financial Services,
Technology Advisory
geoffrey.hodgson@accenture.com

Contributor



Elodie de Fontenay

Director
Accenture Financial Services,
Technology Advisory
Offering & Innovation Development
Global Lead
elodie.b.de.fontenay@accenture.com



Carolyn Goldbeck

Managing Director
Accenture Financial Services,
Technology Advisory Intelligent Solutions
North America Insurance Lead
carolyn.r.goldbeck@accenture.com



Devin Loe

Consultant
Accenture Financial Services,
Technology Advisory
devin.loe@accenture.com

Additional Contacts

Tim Broome

Managing Director
Accenture Financial Services,
Technology Advisory
Intelligent Solutions Africa and Asia Pacific Lead
tim.broome@accenture.com

Philipp Gaida

Senior Manager
Accenture Financial Services,
Technology Advisory
Intelligent Solutions Europe Lead
philipp.gaida@accenture.com

About Accenture

Accenture is a leading global professional services company, providing a broad range of services and solutions in strategy, consulting, digital, technology and operations. Combining unmatched experience and specialized skills across more than 40 industries and all business functions—underpinned by the world’s largest delivery network—Accenture works at the intersection of business and technology to help clients improve their performance and create sustainable value for their stakeholders. With more than 505,000 people serving clients in more than 120 countries, Accenture drives innovation to improve the way the world works and lives. Its home page is www.accenture.com.

Stay Connected

Accenture Financial Services

Technology Advisory

www.accenture.com/us-en/services/financial-services/technology-advisory-index



Connect With Us

www.linkedin.com/showcase/16197660



Follow Us

twitter.com/TechAdvisoryFS

Disclaimer: This document is intended for general informational purposes only and does not take into account the reader’s specific circumstances, and may not reflect the most current developments. Accenture disclaims, to the fullest extent permitted by applicable law, any and all liability for the accuracy and completeness of the information in this document and for any acts or omissions made based on such information. Accenture does not provide legal, regulatory, audit, or tax advice. Readers are responsible for obtaining such advice from their own legal counsel or other licensed professionals. If you require advice or further details on the topic, please contact your Accenture representative.