ACAEBIN Quarterly Meeting

Auditing Big Data and Digital
Ben Afudego

December 2019

Presenting today

Ben Afudego

Partner, Advisory Leader West Africa

EY

Agenda

- **What is Big Data**

- **Case Studies**

- **Benefits of Big Data**

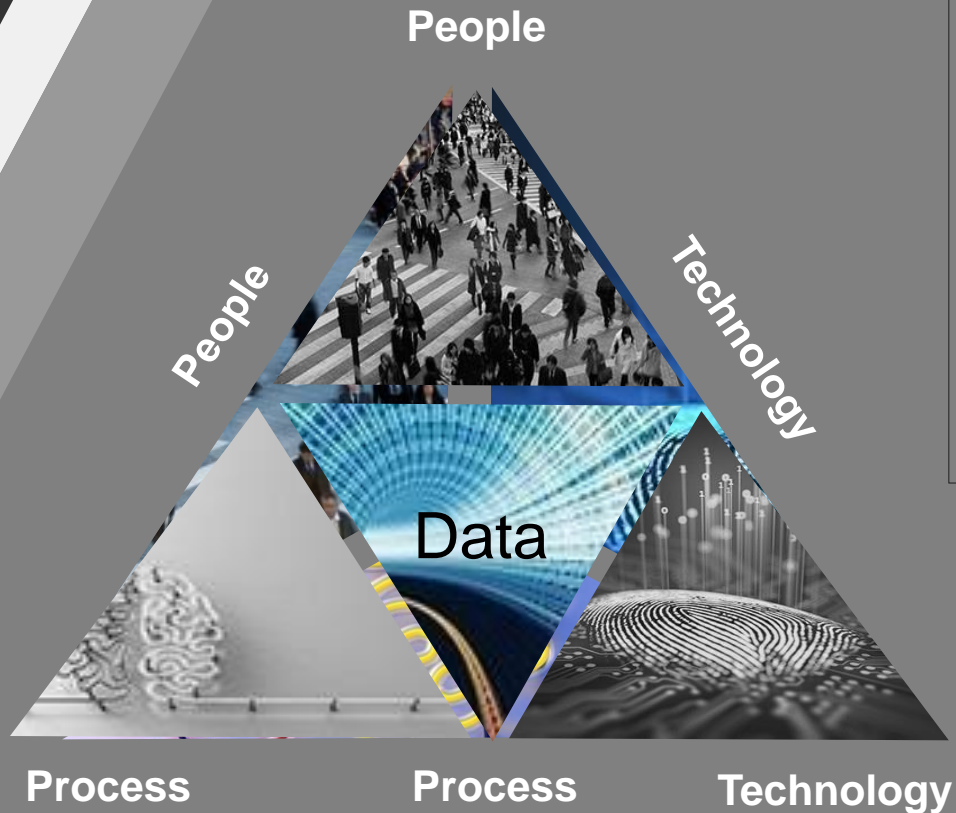- **The Role of IA in Big Data Usage**

EY

Industry disruption

34% of organizations reported seeing a significant **change in source of competitive** differentiation in their industry over the past 3 years based on data. This is expected to accelerate in the next 3 years.

———————— McKinsey

Data

By 2022, **90%** of corporate strategies will explicitly mention information as a critical enterprise asset and analytics as an essential competency.

———————— Gartner

People

People

Technology

Data

Process

Process

Technology

EY

What is Big data for the banking industry?

...Imagine watching all your "ad hoc" reports in real time, iteratively with unlimited granularity and aggregation power…

The key idea is **Be wiser and more agile than ever**

Banks have to be **wiser** to know how to act and react while **agile** to cope with this actions adapting to this new reality that the Big Data tsunami brings

Big data is not only a noun but a verb that encapsulates concepts like:

Data architecture and design

Granularity and aggregation of data

Analysis & machine learning

Visualization

Dissemination

Real time responses

Flexibility in ad hoc reporting

Business intelligence as a paradigm

EY

The global big data market is expected to grow at a CAGR of 10% during 2018–2027 and reach US$103b by 2027

Global Big Data Market, 2011 – 2027 (US$ billion)
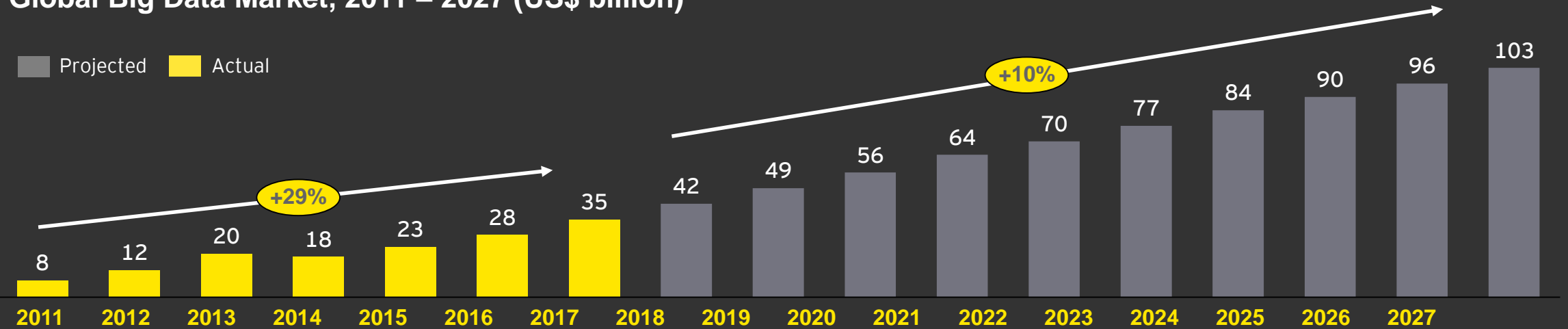
The big data market grew from $8b in 2011 to $35b in 2017, at a CAGR of 29%. The rapid stride in big data analytics is facilitated by declining technological costs, more accessible machine learning capabilities and an improvement in the analytics talent pool. Further growth in the big data industry will be led by the software market, which is expected to grow at a CAGR of 16% during 2017-27.

Source: Wikibon, Statista
Global WAM Trends: Big Data and Analytics

EY

Why Big Data – Client quotes?

1 ...previous investment in IT systems - ERP, CRM, SCM, EPM, HCM etc. - has generated large volume of data yet to be used
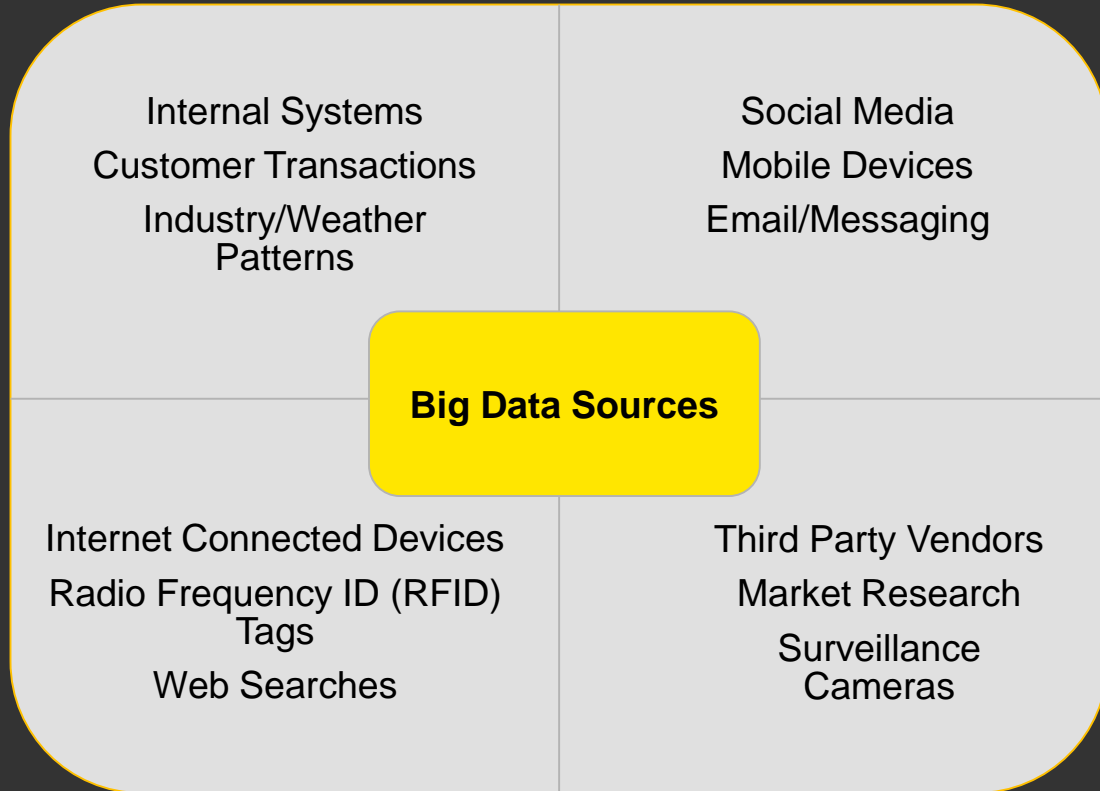
3 ...we have started to invest in basic Data and Analytics solutions - foundation / building blocks are in place for a quick start, however the same is not delivering perceived value

2 ... traditional data warehouse - are being replaced with Big Data to provide value add information to businesses

4 Big Data enables you to Access, Process and Consume Data from within and outside the organization, both structured and unstructured data...delivering valuable information for effective decision making

Auditing Big Data

EY

What is Big Data?

Examples of Big Data Sources

Internal Systems
Customer Transactions
Industry/Weather Patterns

Social Media
Mobile Devices
Email/Messaging

Big Data Sources

Internet Connected Devices
Radio Frequency ID (RFID) Tags
Web Searches

Third Party Vendors
Market Research
Surveillance Cameras

Characteristics of Big Data

3 Vs of Big Data

Volume: The amount of data being created is vast compared to traditional data sources

Variety: Data comes from all types of formats. This can include data generated within an organization as well as data created from external sources, including publicly available data.

Velocity: Data is being generated extremely quickly and continuously.

Additional Vs

Veracity: Data must be able to be verified based on both accuracy and context.

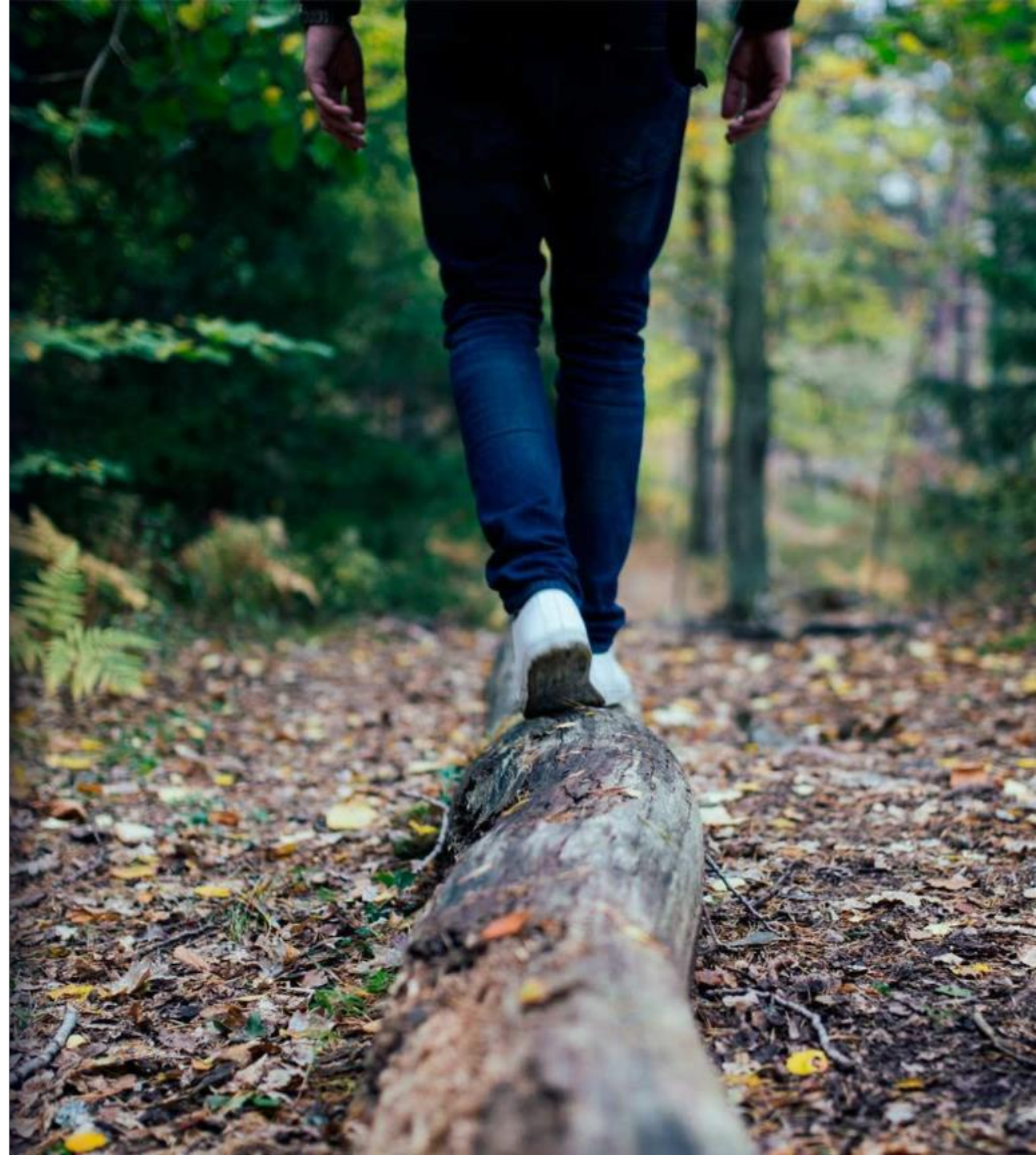**Variability:** Big data is extremely variable and always changing.

Visualization: Analytic results from big data are often hard to interpret; therefore, translating vast amounts of data into readily presentable graphics and charts that are easy to understand is critical to end-user satisfaction and may highlight additional insights.

Value: Organizations, societies and consumers can all benefit from big data. Value is generated when new insights are translated into actions that create positive outcomes.

Auditing Big Data

EY

Case Study

What this means in real terms

Use of Alternate Data Sources for New to Credit SMEs – Big Data

Big Data for Credit Scoring

- Social Networks
- Sensor-enabled ecosystems
- Mobile Data
- Maps
- Browser Data
- Digital footprint
- Email
- Form Filling Analytics
- Psychometric Data
- Financial Transaction Data

If financial data is scarce or traditional credit scores are non-existent such as in emerging markets, Big Data Scoring's solution is able to form the core of the credit scoring process

Auditing Big Data

EY

Along with traditional data, credit decisioning alternate data sources that can be considered for Consumer and SME financing

Case Study

1. Telco. Data

► Paytype: Prepaid / Post-paid
► Cumulative days of activity
► How many times been blocked / barred
► total airtime usage
► No. of deposits to in last 6 months

2. Utility Bills

→ Month on month gas bills
→ Month on month water bills
→ Month on month electricity bills
→ No. of Missed payments
→ Amount of missed payments

3. Tax data

→ Tax Amount paid
→ Missed tax payments
→ PF Deposits
→ Primary and Secondary source of income
→ Personal Investments

4. Bank statements

→ Debit and Credit Amounts (month on month)
→ Average Balance end of each month
→ High Value credits (Top 5)
→ No. of Credit cards
→ Missed salary record

5. SMS scrapping

→ No. of Bank Accounts
→ Salary Bank Account
→ Total CC transactions
→ Shopping bills
→ Online subscription bills
→ Online wallet details

6. Insurance data

→ No. of insurance policies
→ Type of insurance policies
→ Premium amounts
→ Maturity Date
→ Insurance Usage history

7. Trade Credits

→ Data on repayment track record
→ Defaults / delays
→ Trade credits history.

8. Bill of Entry

→ Pending Bill of Entries
→ GR return details

9. E-commerce

→ quantum of transactions
→ product quality / product returns
→ repayment behaviour
→ customer feedback / complaints

10. Legal Suits

→ Latest Legal Status,
→ Customer quality related cases
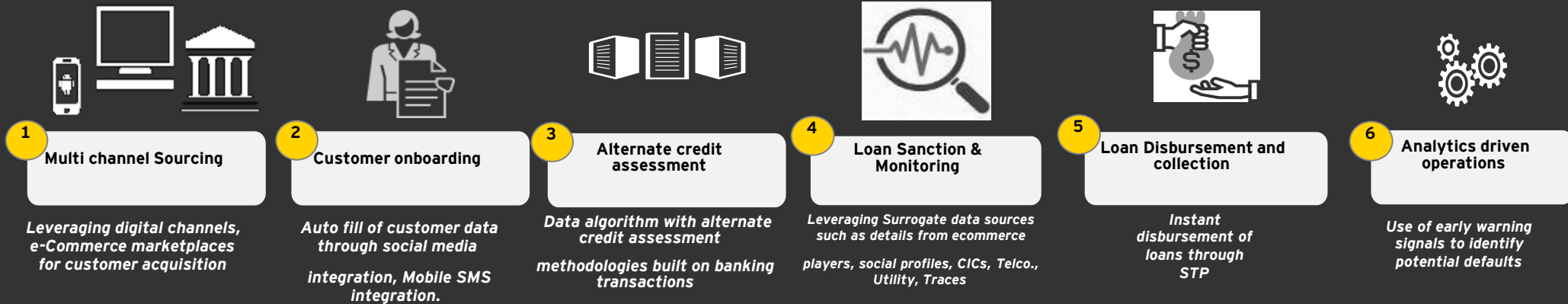→ Amount of litigation

11. Social Media

→ Customer's opinion,
→ Legal cases,
→ No of reviews
→ No of likes
→ Daily update analysis

12. Collateral Details

→ Details on the property / collateral type, collateral value and the encumbrance details
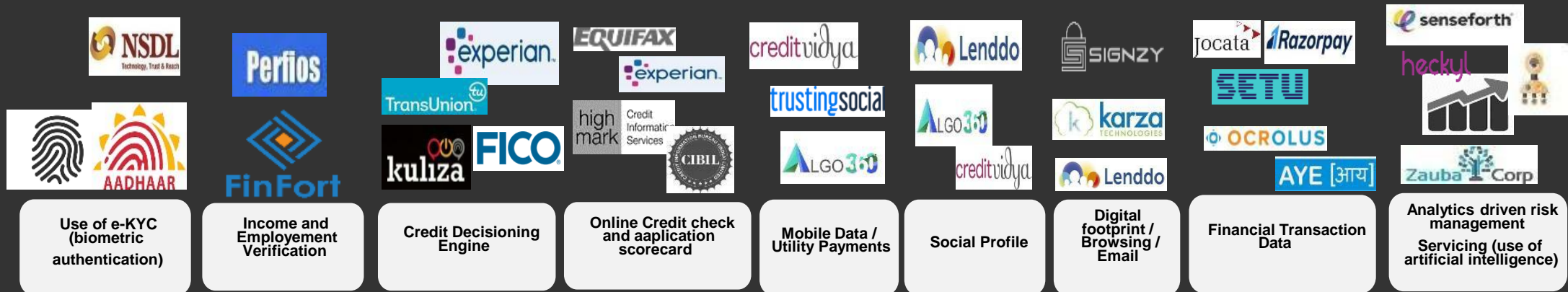
EY

For Banks & FIs, an optimal solution would be to move to adopting an end-to-end digital revamp of the entire credit eco-system

Value Chain

Case Study

1 Multi channel Sourcing

Leveraging digital channels, e-Commerce marketplaces for customer acquisition

2 Customer onboarding

Auto fill of customer data through social media

integration, Mobile SMS integration.

3 Alternate credit assessment

Data algorithm with alternate credit assessment

methodologies built on banking transactions

4 Loan Sanction & Monitoring

Leveraging Surrogate data sources such as details from ecommerce

players, social profiles, CICs, Telco., Utility, Traces

5 Loan Disbursement and collection

Instant disbursement of loans through STP

6 Analytics driven operations

Use of early warning signals to identify potential defaults

Key enablers

Use of e-KYC (biometric authentication)

Income and Employement Verification

Credit Decisioning Engine

Online Credit check and aaplication scorecard

Mobile Data / Utility Payments

Social Profile

Digital footprint / Browsing / Email

Financial Transaction Data

Analytics driven risk management

Servicing (use of artificial intelligence)

EY

Credit decisioning alternate data sources that can be considered for Consumer Lending and SME financing

Case Study
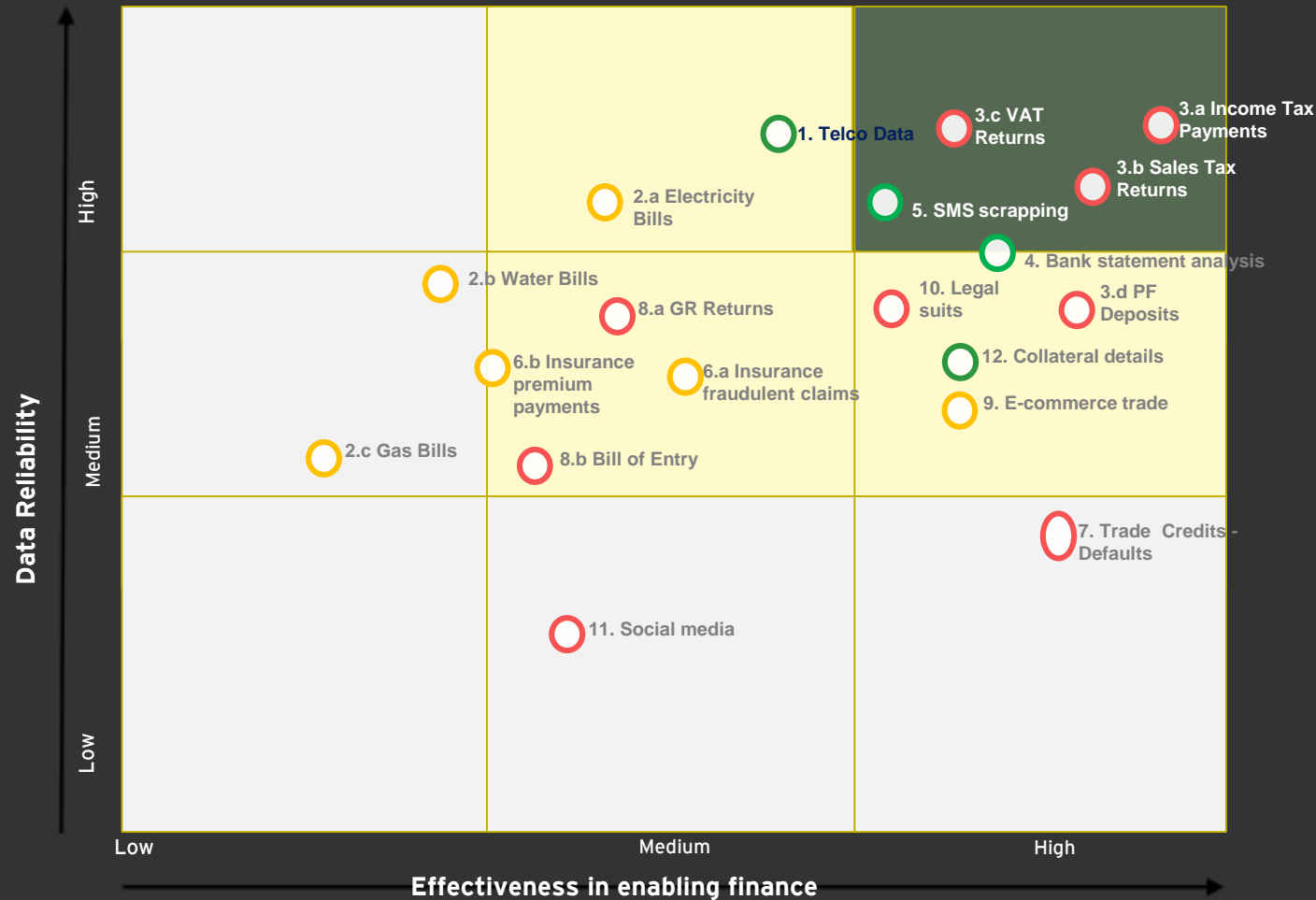
Key factors that can be used as a proxy that enables Banks to assess Self-Employe Small Businessmen:

➤ Use of surrogates for proving repayment conduct - repayment track record on utilities payments (telco, gas, water, electricity), statutory payments (however, in the unbanked segment data on statutory payments may not be readily available / reliable)

➤ Alternate data sources for estimating business volume/ cash flows: business volumes through information from E-commerce companies, trade credit information

Data Reliability (vertical axis: High, Medium, Low)

Effectiveness in enabling finance (horizontal axis: Low, Medium, High)

- 1. Telco Data
- 3.c VAT Returns
- 3.a Income Tax Payments
- 2.a Electricity Bills
- 3.b Sales Tax Returns
- 5. SMS scrapping
- 4. Bank statement analysis
- 2.b Water Bills
- 8.a GR Returns
- 10. Legal suits
- 3.d PF Deposits
- 6.b Insurance premium payments
- 6.a Insurance fraudulent claims
- 12. Collateral details
- 9. E-commerce trade
- 2.c Gas Bills
- 8.b Bill of Entry
- 7. Trade Credits - Defaults
- 11. Social media

Difficulty to capture
- Low (green)
- Medium (orange)
- High (red)

- High effectiveness and reliability of data
- Average effectiveness and reliability of data
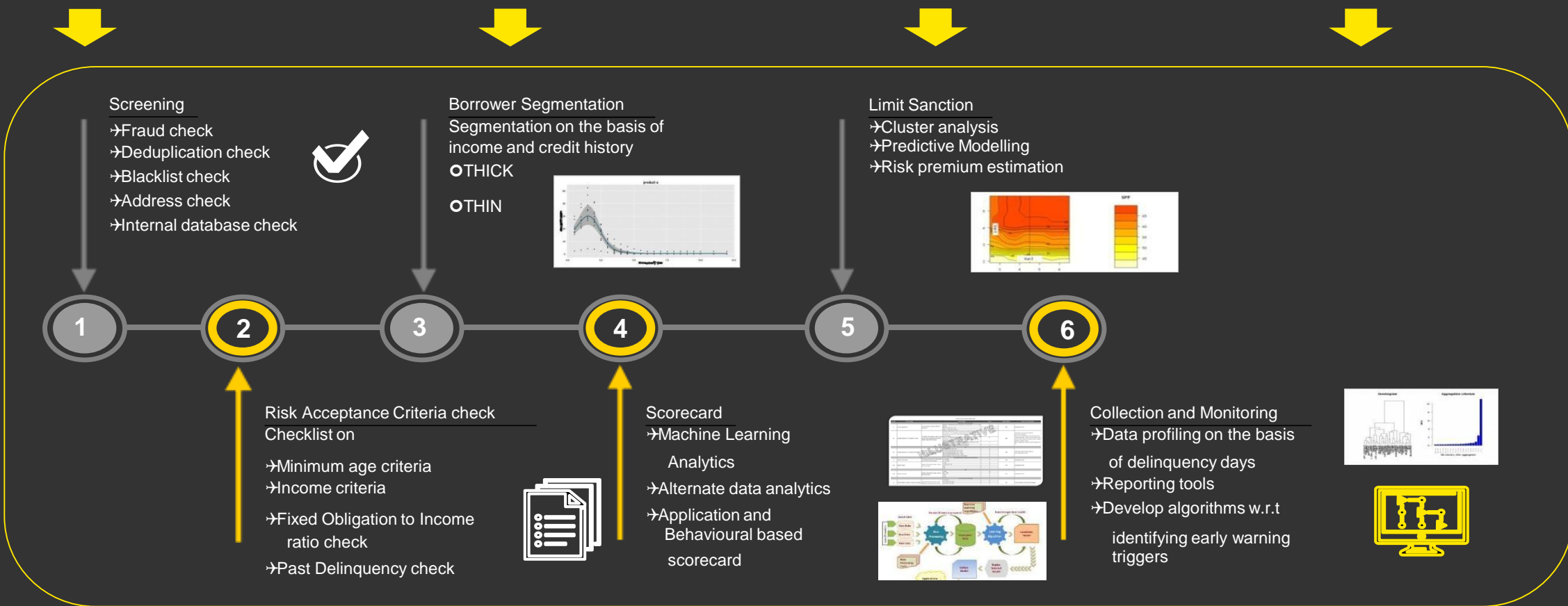- Low effectiveness and reliability of data

Auditing Big Data

EY

Based on the traditional and alternate data sources, Banks and FIs can implement credit decisioning models and algothrims at each step
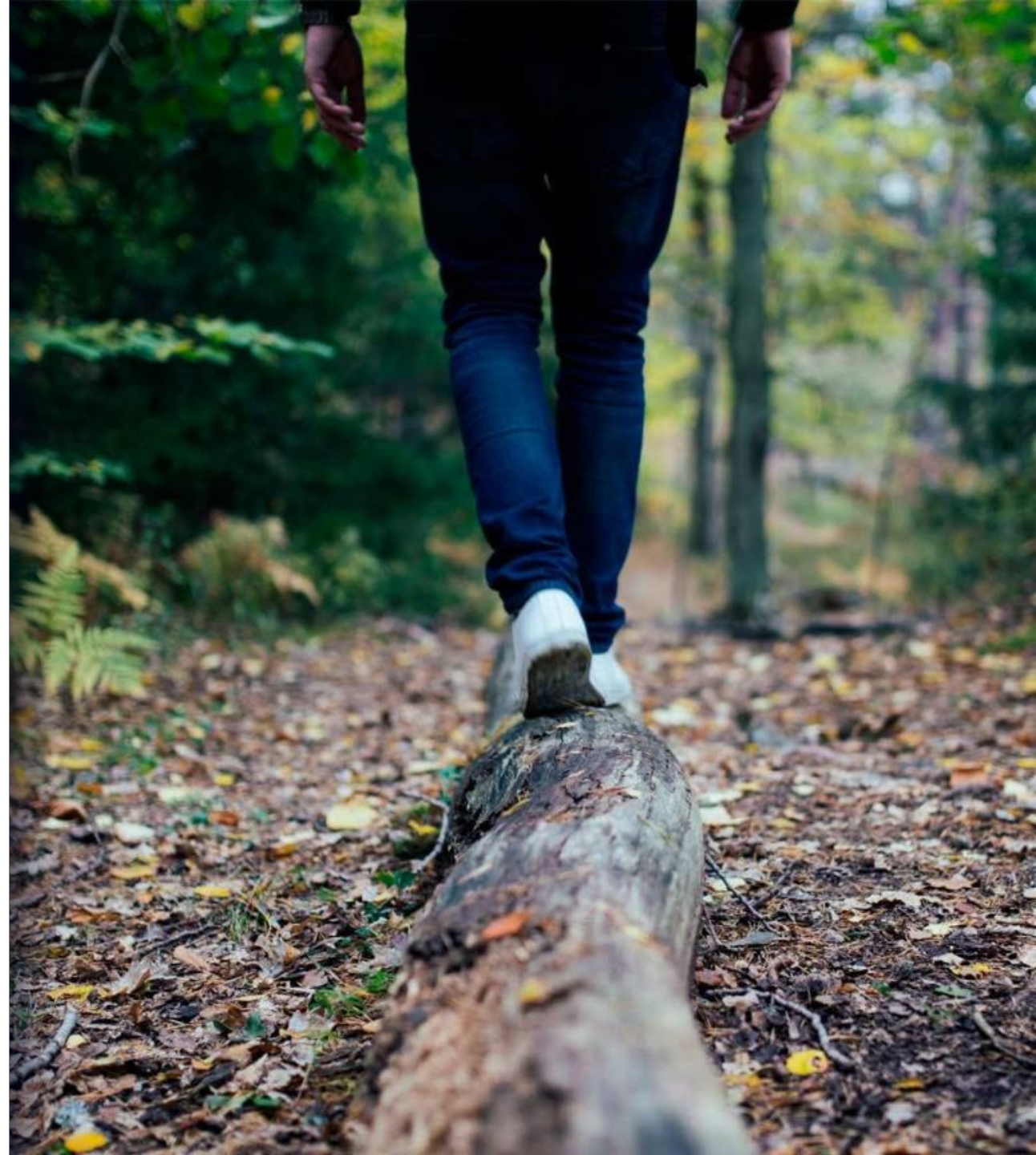
Case Study

Traditional Data sources + Alternate Data Sources

Screening
- ✈Fraud check
- ✈Deduplication check
- ✈Blacklist check
- ✈Address check
- ✈Internal database check

Borrower Segmentation
Segmentation on the basis of income and credit history
- ○THICK
- ○THIN

Limit Sanction
- ✈Cluster analysis
- ✈Predictive Modelling
- ✈Risk premium estimation

1 — 2 — 3 — 4 — 5 — 6

Risk Acceptance Criteria check
Checklist on
- ✈Minimum age criteria
- ✈Income criteria
- ✈Fixed Obligation to Income ratio check
- ✈Past Delinquency check

Scorecard
- ✈Machine Learning Analytics
- ✈Alternate data analytics
- ✈Application and Behavioural based scorecard

Collection and Monitoring
- ✈Data profiling on the basis of delinquency days
- ✈Reporting tools
- ✈Develop algorithms w.r.t identifying early warning triggers

Auditing Big Data

EY

Benefits and Risks of Big Data

Benefits of Big Data to the Organization

Big Data provides a number of important benefits for the Organization.

Big Data Benefits for the Organization

- ► **Competitive advantage.**
- ► **Increased revenue.**
- ► **Innovation and faster product development.**
- ► **Market demands predictions.**
- ► **Well-informed business decisions.**
- ► **Operational efficiency.**
- ► **Enhance organization's transparency**

EY

Risks associated with Big Data

Is your organization ready for Big Data? There are risks associated with your organization's readiness to get on board with Big Data

Program governance

Lack of appropriate management support, funding, and/or governance over the big data program can expose the organization to undue risk or failure to meet strategic goals.

Data quality, management, and reporting

Gauging the quality of Big Data can be very difficult because of a lack of reference standards around Big Data concepts and properties

Additional Platform Requirements

Implementing Big Data requires additional layers of software and hardware at scale, which introduce new levels of architectural complexity and risk to your Organization.

Regulatory Compliance Risks

The nature of Big Data makes it difficult to define compliance frameworks around Big Data usage. This implies significant risks to regulatory compliance practice in this area.

Data Security and Privacy Risks

Ineffective information security standards and configurations may result in unauthorized access to — and theft of — data, inappropriate modifications of data, and regulatory compliance violations.
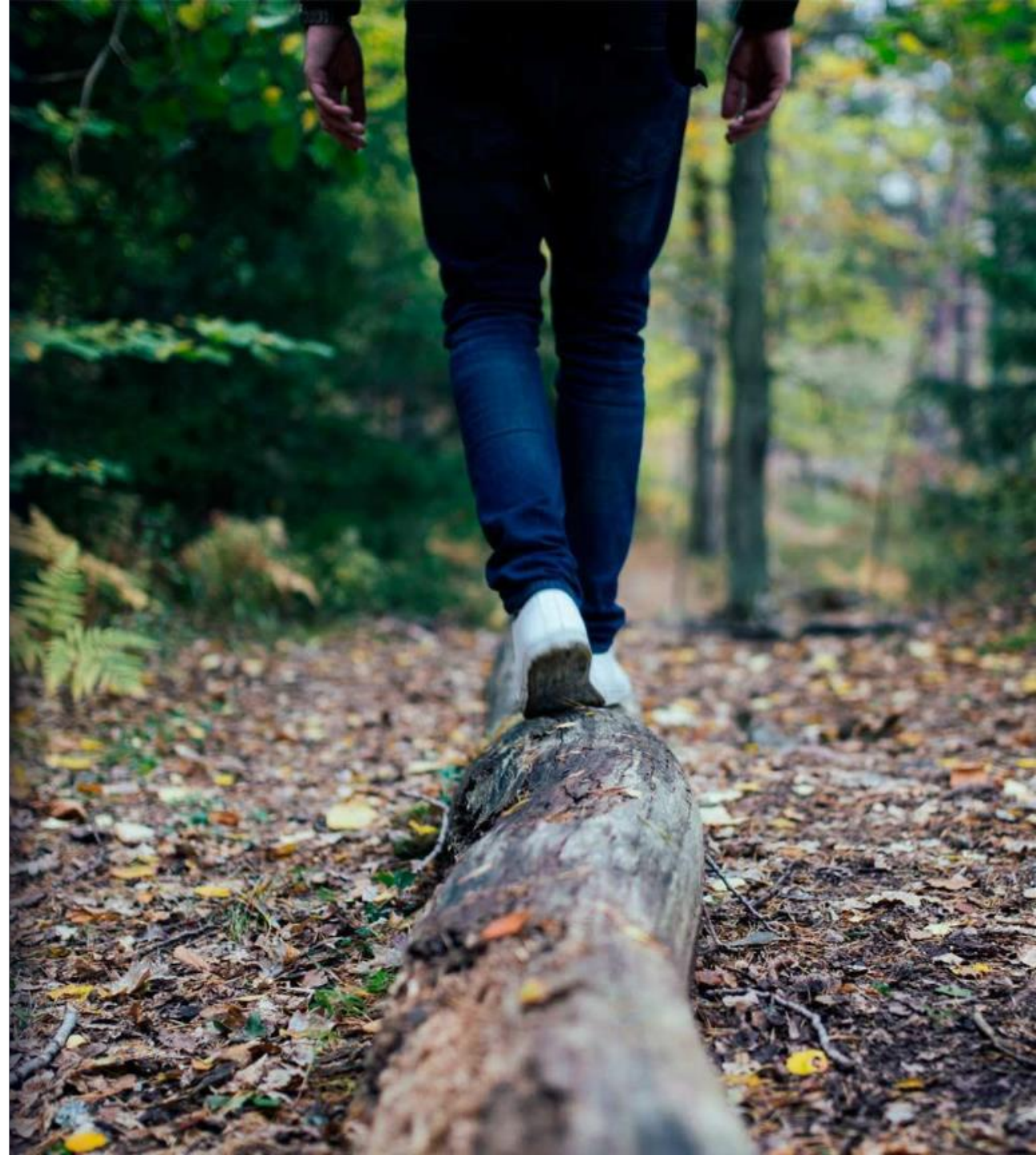
Knowledge Gaps & Limited Potentials

There are still significant gaps in global knowledge around Big Data which may limit the potential applications in your Organization

Auditing Big Data

EY

Role of IA in Big Data Usage

Role of IA in giving assurance on Big Data usage in the organization

Internal audit provides coverage of big data through multiple audits versus a single, stand-alone big data audit.

Internal audit should consider involvement through:

1 Formal and/or informal assessments - advisory projects, pre- or post-implementation reviews, and adequate participation in governance and steering committees.

2 Consider the role of big data within organizations as part of risk assessment and audit planning. If the risks are significant, internal audit can determine an appropriate plan to provide coverage of big data risks and controls.

3 Verify that the objectives of a big data program align with the enterprise-wide business strategy

4 Perform tests to ensure the big data program provides value and is fully supported by appropriate leadership in the organization.

5 Focus significantly on how the data is being consumed and the actions the organization is taking based on results obtained from big data analysis.

6 Leverage big data solutions to support data analytic efforts for audit projects

7 Ensure the confidentiality, integrity, availability, and performance of big data systems aligns with management's business requirements and needs.

8 Audit programs should include test steps to ensure the quality, security, and privacy of the data used for analysis, as well as analytic outputs

9 Educate the board on the organization's big data initiatives, the resulting risks and challenges, and the significant opportunities and benefits.

Auditing Big Data

EY

Challenges for IA functions in auditing Big Data

Risks to Data Quality Audits

Big Data tends to be largely unstructured. This makes it hard to evaluate Data Quality since there are no reference standards to evaluate against.

- Other risk factors such as the scale of Big Data makes it difficult to obtain representative sample sizes

- The scale of Big Data also reduces the meaningfulness of results, according to the **Central Limit Theorem**

- Knowledge gaps in Big Data concepts might also affect auditors' abilities to carry out effective Big Data audits.

Risks to Data Fit-for-Purpose Audits

Fit-for-Purpose audits require application-level testing that may be out of reach for most data auditors because of the knowledge gaps that exist around Big Data concepts as well as the lack of specialist Big Data teams.

- Many fit-for-purpose data audits require the use of scripts or "test programs", as well as decision trees or process flows that help evaluate whether a sample of data is fit for purpose.

- With Big Data, these test programs and process flows do not follow the usual paradigms: as an example, typical programs are usually a combination of Object-Oriented, Procedural, Functional or Imperative paradigms, while Big Data applications employ **MapReduce** and **Parallel Programming** paradigms

Key Risk Factors

Custom Big Data Formats

Big Data can take a variety of custom formats. Without a well-defined list of standard formats, it becomes hard to qualify data

No Audit Standards for Big Data

Data auditors have no agreed reference standard for auditing Big Data, making case-based methods and findings questionable

Wrangling Big Data is Hard Work

The sheer scale of big data makes data wrangling during audits very hard work requiring teams and very long hours. This can discourage auditors or affect the accuracy of findings

Huge Knowledge Gaps still exist

Big Data is still very much a technical area, and there might be significant knowledge gaps for most auditors, affecting their ability to carry out effective audits

"Because Big Data is complex, large and mostly unstructured, it poses significant risk to data auditors during Data Quality and Data Fit-for-Purpose Audits."

EY

How IA teams can mitigate the challenges of Big Data audits

Collaborate & Learn.

1

- ❑ Work with Software Testers in your Organization to define Big Data reference and implementation standards.

- ❑ Bridge the Knowledge Gap by conducting more research into this area.
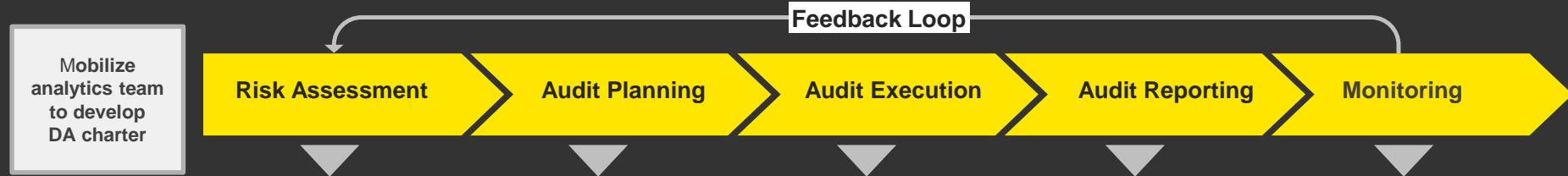
Build.

2

- ❑ Build tools to automate Big Data testing in your organization.

- ❑ This reduces the manpower requirements and scales up your ability to provide periodic or on-demand Big Data audits

Enforce.

3

- ❑ Once you have established a framework around Big Data auditing in your organization, you need to ensure compliance to minimize the risks of Big Data adoption for your organization.
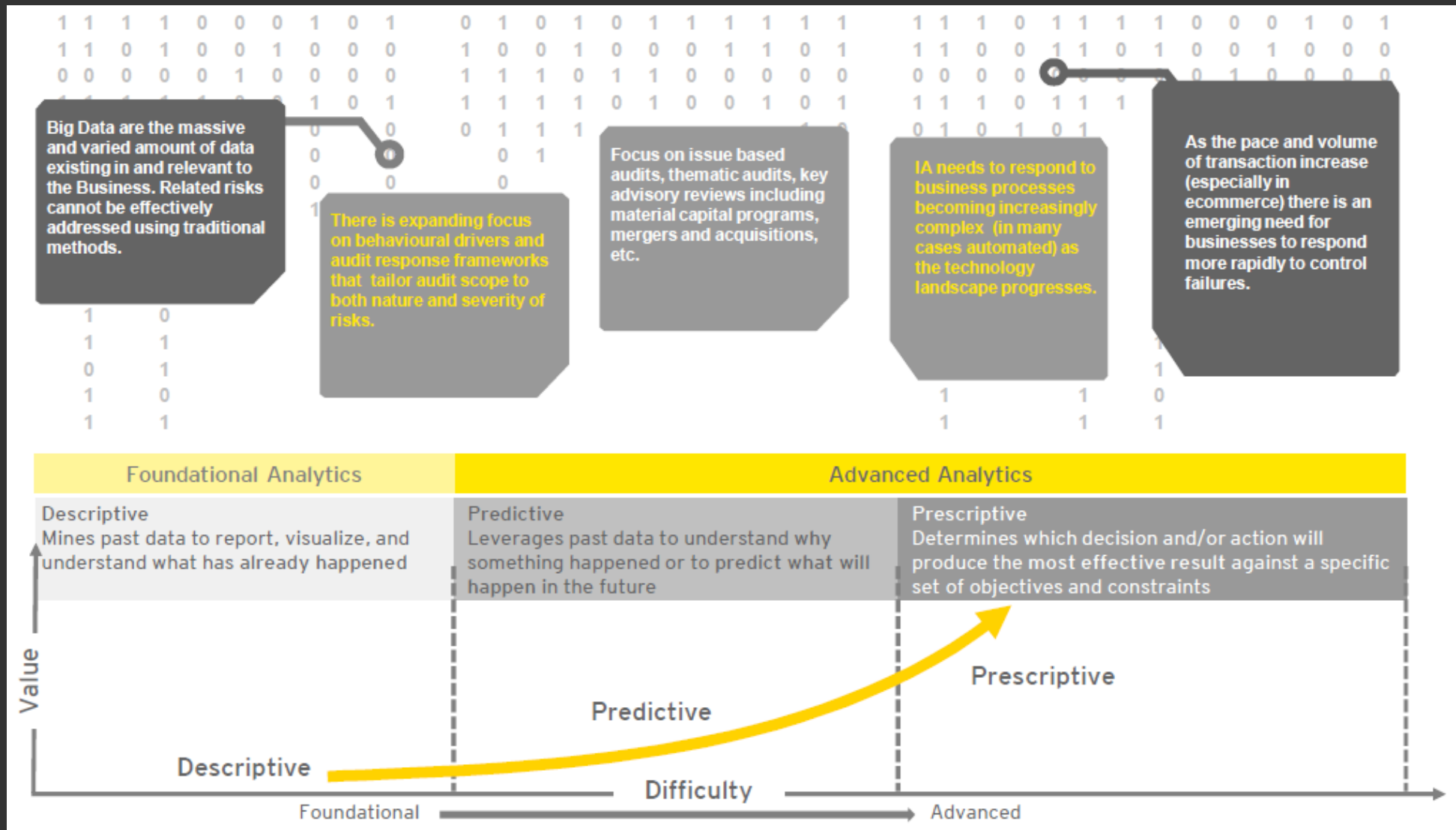
EY

How can IA deliver through analytics?

Feedback Loop

Mobilize analytics team to develop DA charter	Risk Assessment	Audit Planning	Audit Execution	Audit Reporting	Monitoring

Key activity

Risk Assessment	Audit Planning	Audit Execution	Audit Reporting	Monitoring
► **Identify risk** assessment priorities ► **Determine scope** of audit plan activities	► **Preliminary "scan"** of relevant audit information to drive project scope, sampling and fieldwork procedures	► **Identify anomalies**, trends and potential fraud indicators ► Replace sample testing approaches with **full-coverage data analytics**	► Provide **quantifiable**, fact-based information for reportable issues and exceptions ► **Visualization** of audit findings	► Provide an automated **basis for continuous auditing** & controls monitoring. ► Provide analytical **input for follow-up Risk Assessment**.

Example analytics

Risk Ranking	Regional benchmarking	Red Flags / Observations	Report Visualizations	Controls Monitoring
Value at Risk Analysis	Key Risk Indicators	Robotic Process Automation	Risk Quantification	Risk / Action Monitoring

Big Data are the massive and varied amount of data existing in and relevant to the Business. Related risks cannot be effectively addressed using traditional methods.

There is expanding focus on behavioural drivers and audit response frameworks that tailor audit scope to both nature and severity of risks.

Focus on issue based audits, thematic audits, key advisory reviews including material capital programs, mergers and acquisitions, etc.

IA needs to respond to business processes becoming increasingly complex (in many cases automated) as the technology landscape progresses.

As the pace and volume of transaction increase (especially in ecommerce) there is an emerging need for businesses to respond more rapidly to control failures.

Foundational Analytics	Advanced Analytics	
Descriptive Mines past data to report, visualize, and understand what has already happened	Predictive Leverages past data to understand why something happened or to predict what will happen in the future	Prescriptive Determines which decision and/or action will produce the most effective result against a specific set of objectives and constraints

Value

Prescriptive

Predictive

Descriptive

Difficulty

Foundational ————————————→ Advanced

How do Analytics add value to IA?

	Issue	How has analytics helped us get better or become more relevant?	Resulting value-add	Big Picture Impact
Focus audit scope on risks that really matter	Audit scope not focused on high risk business processes or relevant business areas (e.g cost centers)	» Improved audit scope based on analytics driven risk assessment (i.e. selecting cost centers that show potential process circumvention) » De-scoping low risk areas based on analytics results (i.e. de-scoping VMF maintenance due to low or no duplicate vendors / blank or missing fields / changes to VMF)	» Audits are focused on relevant high risk areas which results in greater ability to improve the business and provide increased assurance on risks that matter	RELEVANCE
Time savings though in depth upfront knowledge (e.g. process, irregularities) and repeatable analytics	Auditor taking too much time from the business during on site visits Collection of data is too slow and often incomplete which requires back and forth communication Risk of setting up DA across the IA group is inefficient	» Upfront analytics results improve audit preparation and process understanding resulting in more efficient meetings with the business » IA access to system data allows for higher independence and faster collection of data. Transaction tests on entire population can reduce traditional (expanded) sample testing saving time without compromising assurance » Standardized analytics procedures allows for fast, repeatable and scalable setups that can be leveraged for new DA rollouts within the IA group	» Time is more productively used on investigating potential exceptions based on full population testing and conducting in-depth root cause analysis, supporting documentation gathering, and defining action plans for process/control improvement » Data analytics implementation throughout the IA group is standardized and scalable ensuring increased risk coverage, quality and speed	EFFECTIVENESS
Improved audit value through high impact findings and meaningful recommendations	Audit results and recommendations do not add much value to the business	» Analytics has enabled IA to quantify the extent of operational business risks based on full population testing or predictive data analytics (i.e. stating number of transactions affected and (future) financial impact to the business)	» Recommendations and action plans are based on quantitative results which better explains risk impact leading to greater acceptance by the business » Predictive analytics allows the business to stay ahead of the risks	RELEVANCE

Auditing Big Data

EY

How do Analytics add value to IA?

	Issue	How has analytics helped us get better or become more relevant?	Resulting value-add	Big Picture Impact

Sustainable Reporting	Audit reports are cumbersome or not easily understood by the business
	
	Risks or issues are not raised on time
		» Enhanced reporting through data visualization and impact representation
		» Take advantage from innovative reporting methods:
		  » Digital reporting
		  » Continuous Reporting Dashboards
		  » Video Reporting
			» Audit reports are simpler, concise and present data in an understandable way
			» Stakeholders have better access to on-time or real-time reporting, enabling them to address risks faster
				RELEVANCE

People Motivation and Auditor Acceptance	Employee dissatisfaction leading to higher turnover Inconsistent quality throughout the IA group	» Increased auditor skill set (i.e. IT, data analytics) resulting in better audits » Enhanced toolbox for the auditor (i.e. analytics visualization dashboards) leading to greater audit efficiency » Consistent quality on every audit as a result of audit teams receiving defined set of analytics	» Auditors adapt an analytics mindset which increases confidence and job satisfaction (reduces job fluctuation) » Better scoring on IA group quality assessments » Data driven auditing can reduce the need to travel (save travel costs / reduce traveling related fluctuation) increasing flexible work arrangements	*EFFECTIVENESS*
IA contribution towards realizing digital strategy of the organization	Organization is not taking advantage of benefits of digitalization and lacking innovation IA group is not recognized throughout the organization	» Enabling a data driven audit approach by using all available (raw) process data » Taking advantage of automation to reduce tedious manual tasks » Digitalizing the IA function with analytical procedures has spearheaded the business' use of analytics throughout (i.e. leading 1st and 2nd line functions to embed CCM/CRM)	» A better, faster, smarter audit, transforming the Internal Audit experience for all stakeholders » Changed perception of Internal Audit to that of a change agent for realizing organizational digital strategy and as a leading function within the business and peer group	*RELEVANCE*

Auditing Big Data

EY

EY | Assurance | Tax | Transactions | Advisory

About EY
EY is a global leader in assurance, tax, transaction and advisory services. The insights and quality services we deliver help build trust and confidence in the capital markets and in economies the world over. We develop outstanding leaders who team to deliver on our promises to all of our stakeholders. In so doing, we play a critical role in building a better working world for our people, for our clients and for our communities.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. For more information about our organization, please visit ey.com.

EYG no. XXXXXX

BMC Agency
GA 1007360

ED None

ey.com

Thank you