Data Driven Credit

Panel Discussion

10,000+ data points from 100+ data sources ...

Other data
Lender interview notes, field verification, photos, call records

Social data
Google, Facebook, LinkedIn, Instagram, website data, news

Telco data
SMS, geo tagging, call records, recharge, M-wallet, mobile apps

Utility data
Electricity, phone, gas, internet

Internet transactions
Sales history, returns, customer rating, customer comments

Business systems data
ERP, payments, inventory mgmt.

Bank statement data
Current account, loans, transactions, bounces

Credit bureau data
Entity credit, proprietor credit

Tax data
GST, ITR, TDS, service tax, TIN, CRILC

Proprietor data
Aadhaar, KYC, PAN, voter Id, professional bodies

Entity data
MCA, Udyog Aadhaar, LEI, Shop Act, Exim code, ESIC, EPF

Legal
CIBIL suits, court cases, EPFO defaulters

Digital data types

○ Structured data   ○ Unstructured data

1

… with architecture enabling real time access
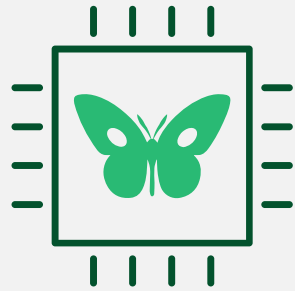
Digital lending enablers

Multiple enabler models to support end-to-end digital lending journeys
- Credit scoring & verification platforms
- API aggregators
- Surrogate data providers
- Digital process enablers (e.g., eSign, eKYC)

Digitization in banks & FIs

Banks setting up tech infrastructure to enable data consumption through
- API gateways
- Middleware
- Lending platforms integrated with data sources

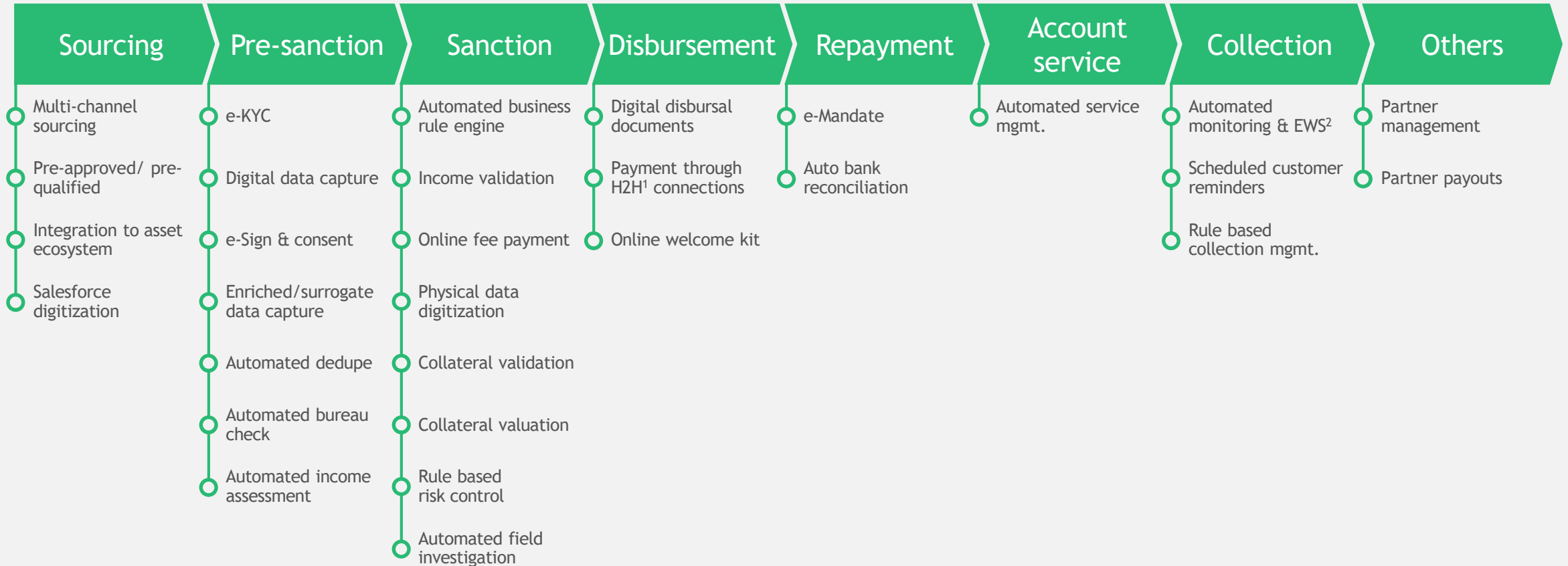Data storage, extraction & analytics platforms being set up

Government and regulator positive intent

Ensuring easier availability of data for consumption
- Open architecture platforms such as Adhaar, GSTN, BBPS
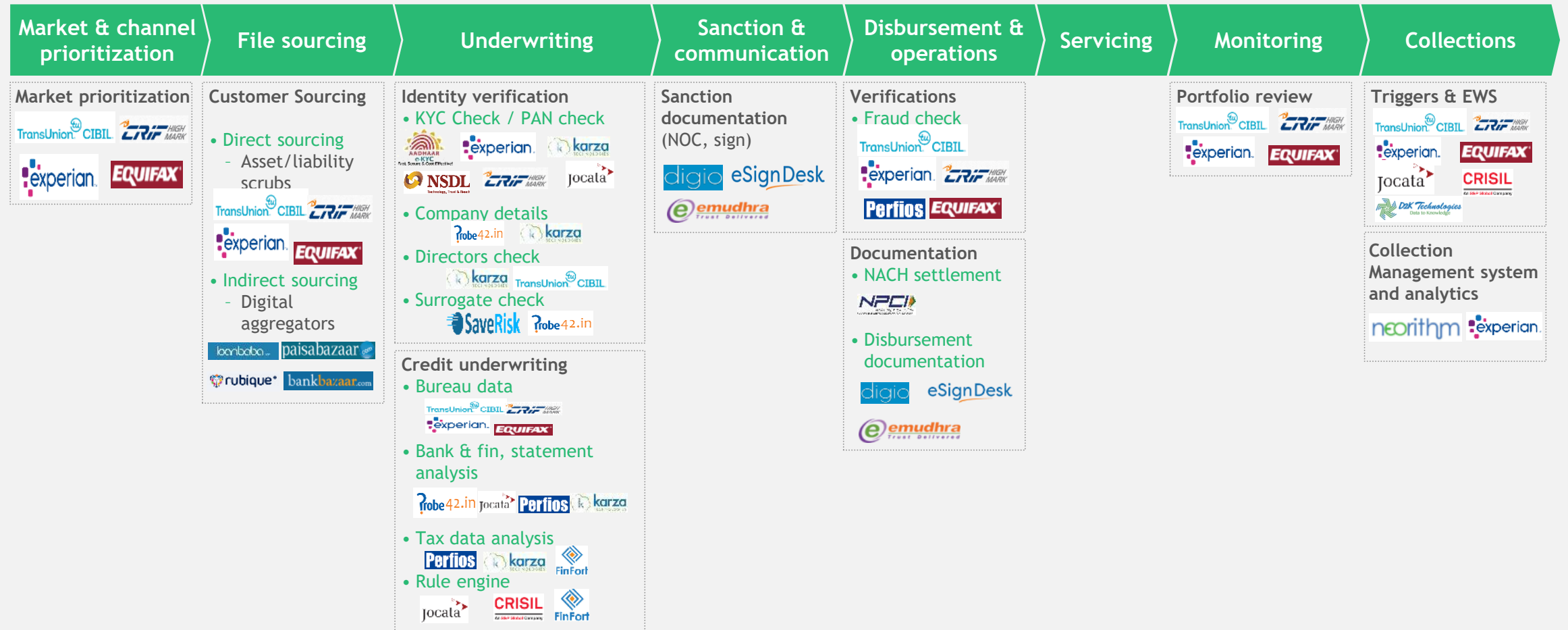- Setting up consolidated platforms such as account aggregators, cKYC

End-to-end customer journeys are being digitized …

Digital building blocks in end-to-end customer journey

Sourcing	Pre-sanction	Sanction	Disbursement	Repayment	Account service	Collection	Others
Multi-channel sourcing	e-KYC	Automated business rule engine	Digital disbursal documents	e-Mandate	Automated service mgmt.	Automated monitoring & EWS[2]	Partner management
Pre-approved/ pre-qualified	Digital data capture	Income validation	Payment through H2H[1] connections	Auto bank reconciliation		Scheduled customer reminders	Partner payouts
Integration to asset ecosystem	e-Sign & consent	Online fee payment	Online welcome kit			Rule based collection mgmt.	
Salesforce digitization	Enriched/surrogate data capture	Physical data digitization					
	Automated dedupe	Collateral validation					
	Automated bureau check	Collateral valuation					
	Automated income assessment	Rule based risk control					
		Automated field investigation					

1. H2H stands for host to host 2. EWS stands for Early Warning Systems
Source: BCG case experience

... with help from ecosystem that exists today in India

Sourcing ——— Sanction ——— Disbursement ——— Monitoring ——— Collect.

Market & channel prioritization	File sourcing	Underwriting	Sanction & communication	Disbursement & operations	Servicing	Monitoring	Collections

Market prioritization
TransUnion CIBIL | CRIF HIGH MARK
experian | EQUIFAX

Customer Sourcing
• Direct sourcing
  – Asset/liability scrubs
  TransUnion CIBIL | CRIF HIGH MARK
  experian | EQUIFAX
• Indirect sourcing
  – Digital aggregators
  loanbaba | paisabazaar.com
  rubique | bankbazaar.com

Identity verification
• KYC Check / PAN check
  AADHAAR e-KYC | experian | karza
  NSDL | CRIF HIGH MARK | Jocata
• Company details
  Probe42.in | karza
• Directors check
  karza | TransUnion CIBIL
• Surrogate check
  SaveRisk | Probe42.in

Credit underwriting
• Bureau data
  TransUnion CIBIL | CRIF HIGH MARK
  experian | EQUIFAX
• Bank & fin, statement analysis
  Probe42.in | Jocata | Perfios | karza
• Tax data analysis
  Perfios | karza | FinFort
• Rule engine
  Jocata | CRISIL | FinFort

Sanction documentation (NOC, sign)
digio | eSignDesk
emudhra Trust Delivered

Verifications
• Fraud check
  TransUnion CIBIL
  experian | CRIF HIGH MARK
  Perfios | EQUIFAX

Documentation
• NACH settlement
  NPCI
• Disbursement documentation
  digio | eSignDesk
  emudhra Trust Delivered

Portfolio review
TransUnion CIBIL | CRIF HIGH MARK
experian | EQUIFAX

Triggers & EWS
TransUnion CIBIL | CRIF HIGH MARK
experian | EQUIFAX
Jocata | CRISIL
D2K Technologies

Collection Management system and analytics
neorithm | experian

Helping break through traditional challenges in lending

"Me too" play:
Inability to differentiate

1

Standardized product play
- Limited scope for innovation given operating model constraints
- Mostly driven by relationship play and hence making it difficult for new entrants

High unit cost

2

High touch operating model leading to high opex
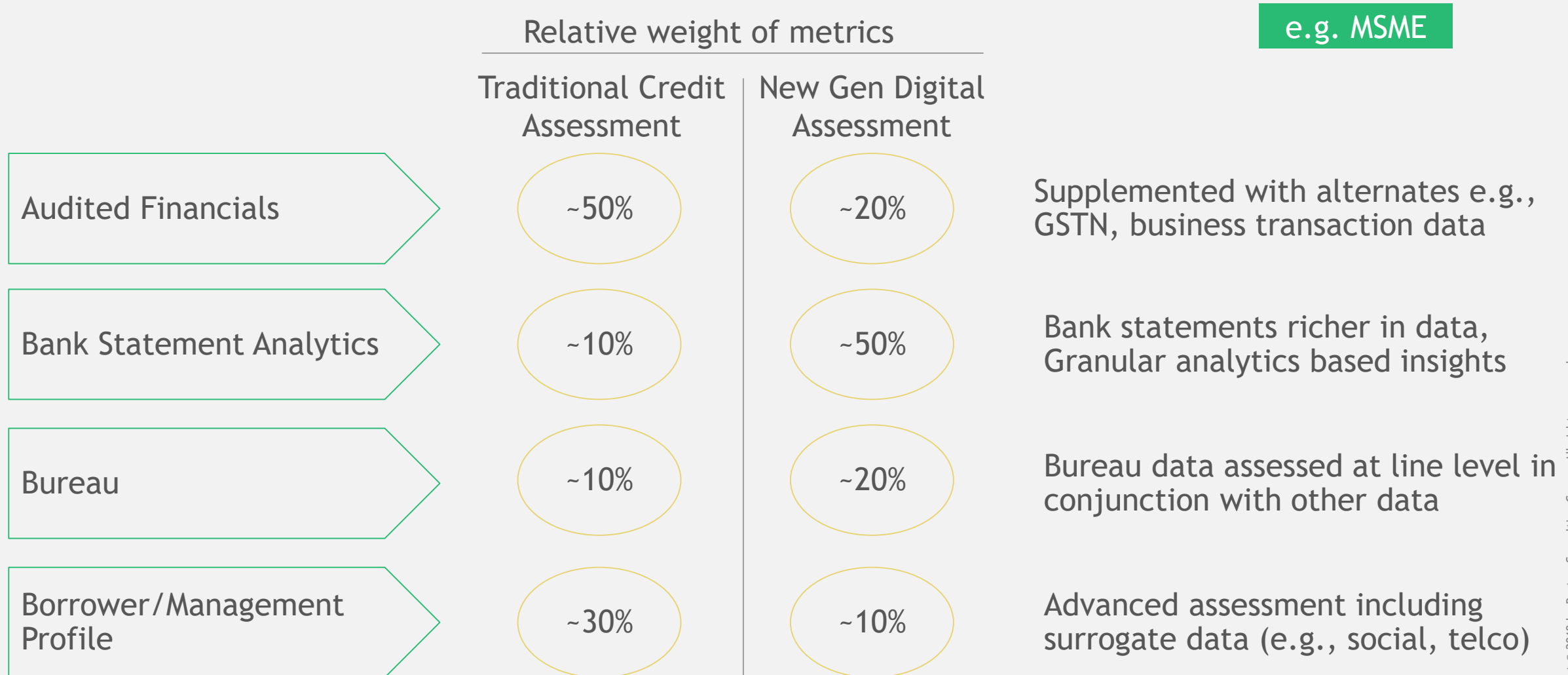- Driven by high number of sales touch points, low credit & operations productivity and high collections cost
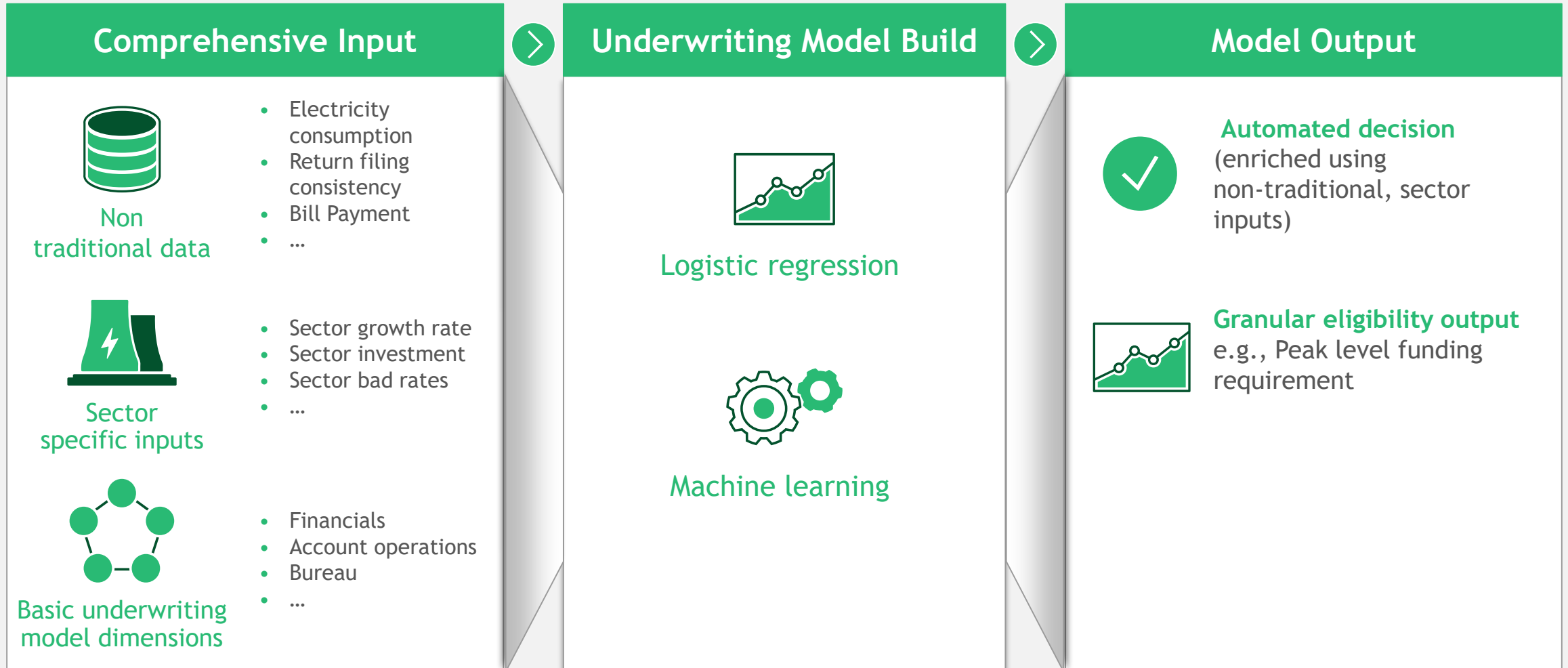
Information arbitrage

3

Information arbitrage in the segment makes it difficult to price the risk
- Incomplete / manipulated financials
- Lack of verifiable digital trail

5

Primary sources of credit insight beginning to change

Relative weight of metrics

	Traditional Credit Assessment	New Gen Digital Assessment	
Audited Financials	~50%	~20%	Supplemented with alternates e.g., GSTN, business transaction data
Bank Statement Analytics	~10%	~50%	Bank statements richer in data, Granular analytics based insights
Bureau	~10%	~20%	Bureau data assessed at line level in conjunction with other data
Borrower/Management Profile	~30%	~10%	Advanced assessment including surrogate data (e.g., social, telco)

6

Source: BCG case experience, expert interviews

Analytics driven decision helping improve credit assessment quality (across underwriting and EWS)

Comprehensive Input

Non traditional data
- Electricity consumption
- Return filing consistency
- Bill Payment
- ...

Sector specific inputs
- Sector growth rate
- Sector investment
- Sector bad rates
- ...

Basic underwriting model dimensions
- Financials
- Account operations
- Bureau
- ...

Underwriting Model Build

Logistic regression

Machine learning

Model Output

Automated decision (enriched using non-traditional, sector inputs)

Granular eligibility output e.g., Peak level funding requirement

Hype vs reality ?

Bankers' perspective
- Data driven credit assessment/EWS : Where are we in the journey of evolution ?
- How does MSME evolution compare with retail ? How can we unlock the potential of digitization, data and analytics in MSME – similar to what we have seen in retail ?
- How do banks evolve organizationally to capitalize on the emerging opportunity ?

Ecosystem partners
- How does data architecture in India compare to global economies ? What can be done to further strengthen it ?
- With data democratization taking more space in government's agenda – how do ecosystem players evolve to provide value added offerings to clients (Banks / FIs) ?
- How do we address the issues around data authenticity / frauds ?