



Barnabás Virág
Deputy Governor

CREDIT DATA - THE DRIVER OF LENDING IN THE 21ST CENTURY

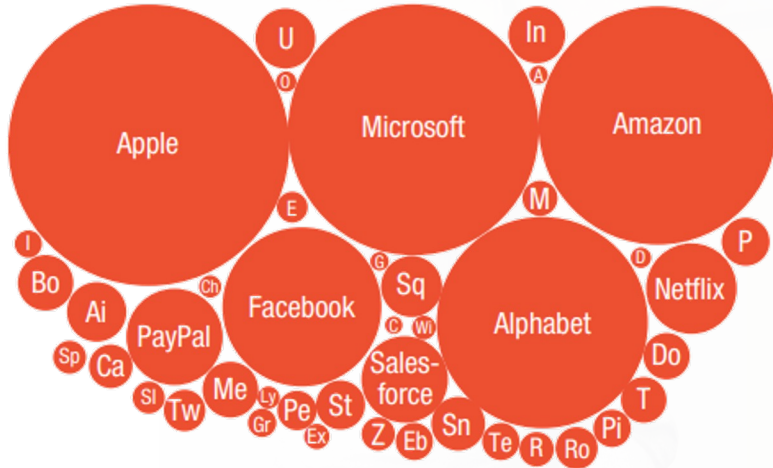


**(CREDIT) DATA IS THE NEW
FUEL OF THE ECONOMY AND
THE BANKING SECTOR**

EUROPE IS CURRENTLY LAGGING BEHIND IN THE DIGITAL RACE – THE UTILIZATION OF DATA WOULD BE KEY FOR A TURNAROUND



America



- Airbnb
- Alteryx
- Booking
- Carvana
- Chegg
- Doordash
- Dropbox
- Ebay
- Etsy
- Expedia
- Grainger
- Grubhub
- Instacart
- Intuit
- Lyft
- Match
- MercadoLibre
- Opendoor
- Palantir
- Peloton
- Pinterest
- Roblox
- Roku
- Slack
- Snap
- Splunk
- Square
- Stripe
- Teladoc
- Twilio
- Twitter
- Uber
- Wish
- Zillow

Europe

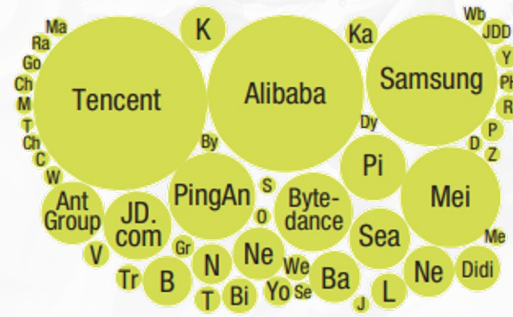


- Adyen
- Auto1
- Checkout
- Delivery Hero
- Edenred
- Hellofresh
- Farfetch
- Klarna
- Spotify
- Just Eat T.
- Yandex

Africa



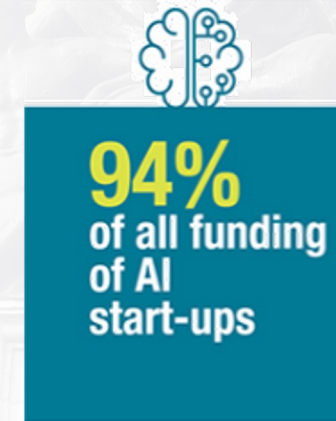
Asia and Pacific



- Baidu
- Beike
- Bilibili
- BYJU
- Chehaoduo
- Coupang
- Dada Nexus
- Didi Chuxing
- Go-Jek
- Grab
- JD Digits
- Kakao
- Kuaishou
- Lufax
- Manbang
- Meicai
- Meituan
- Mercari
- Naver
- Netease
- Ola
- OYO
- Paytm
- Pinduoduo
- PindAn Health
- Rakuten
- Rea
- Sea Group
- Seek
- Sensetime
- Tokopedia
- Trip.com
- VipShop
- WeBank
- WeDoctor
- Weibo
- YonYou
- Yuanfudao

China and the United States

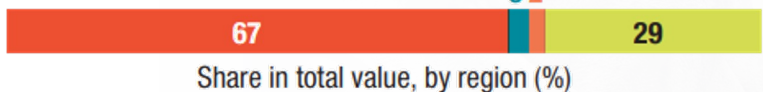
The front-runners in harnessing the value of data



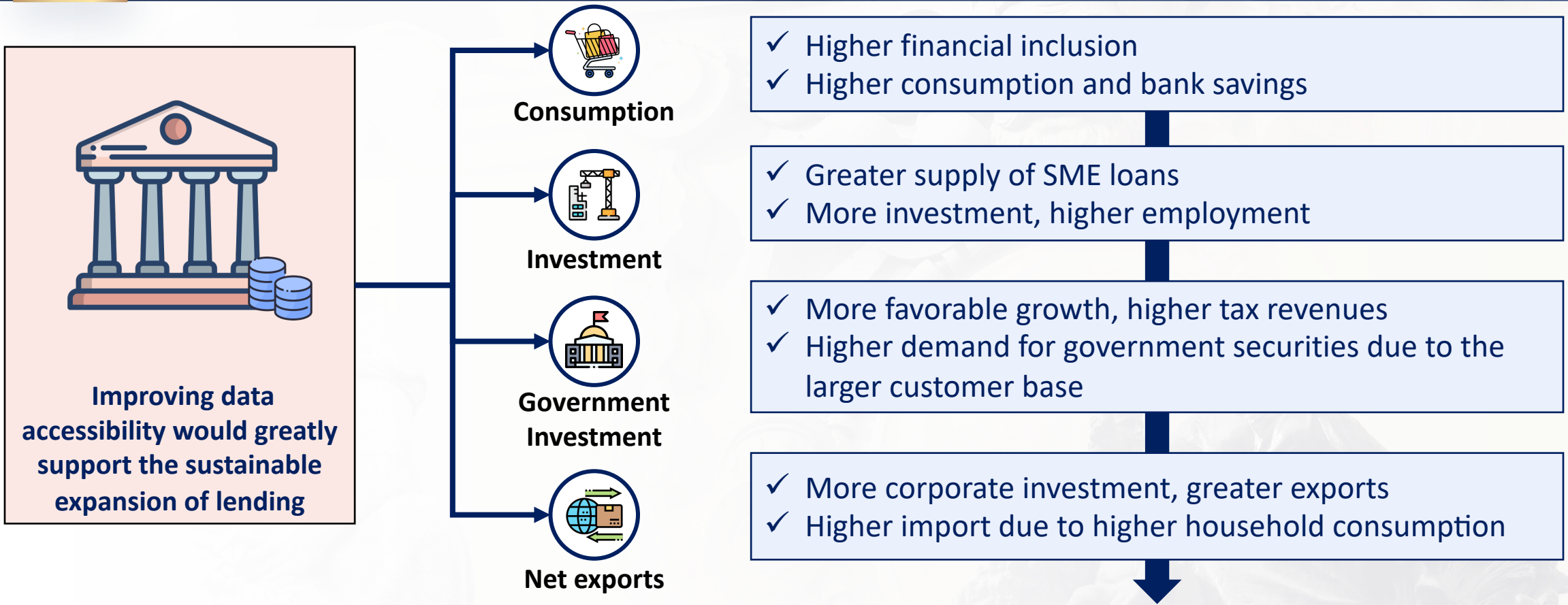
(Data for China and USA combined)

Source: [UNCTAD Digital Economy Report 2021](#)

Geographical distribution of the top 100 global digital platforms, by market capitalization 2021



OPEN ACCESS TO DATA IN THE FINANCIAL SECTOR COULD SIGNIFICANTLY SUPPORT ECONOMIC GROWTH



- Based on the international literature, an annual GDP of up to 1.5 percent higher can be achieved
- In Hungary, artificial intelligence based on data can result in an additional GDP increase of HUF 1 trillion by 2030

Note: The GDP impact for each economy in 2030 is estimated using MGI's Global Growth Model (GGM), a dynamic general equilibrium economic model spanning 100 countries; GGM estimates the economy-wide potential GDP impact by 2030 of 24 banking and payments use cases, each of which individually generates economic value. The attribution of potential economic value across market participants is estimated based on the potential impact of the 24 use cases on a standalone basis, without taking into account dynamic macroeconomic feedback loops that determine GDP impact. Source: McKinsey; European Data Portal; OECD; Magyarország Mesterséges Intelligencia Stratégiája - 2020-2030

MORE INFORMATION CAN LEAD TO FASTER FINANCIAL DEEPENING, BUT DRAWBACKS NEED TO BE MANAGED



USA

- Standardized calculation on centralized data (FICO) - **since 1989** ✓

- **Based on income and wealth** 🧩

- It can exclude people with low incomes ⚠️

- Control of data by the private sector 🔑

- *Typical data used:* regular monthly savings, monthly payment of bills 📊



China

- State, centralized database with large data requirements - **mandatory after 2020** ✓

- **Based on social behavior** 🧩

- It raises ethical and data protection issues ⚠️

- Control of data by the Government 🔑

- *Typical data used:* spending habits, charity, violations of regulations 📊



Hungary

- For the time being, application scoring is the most common form of credit assessment
- The stored data is not always up-to-date, the scoring methods are not innovative
- Control of data is with the individual borrowers
- Access to public or market databases is limited



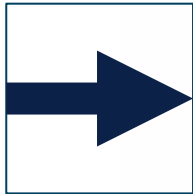
CENTRAL BANKS AS DRIVERS OF CREDIT DATA USAGE



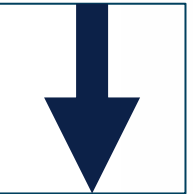
CENTRAL BANKS AS DRIVERS OF CREDIT DATA USAGE



Financial crisis - accelerated the **HORIZONTAL** and **VERTICAL** data demand



Horizontal: covering more sectors, instruments



Vertical: granular, instrument-level, transactional data instead of aggregate



+3rd dimension: data quality – demand for good quality data, standardization

Standardization of national-level Central Credit Registers lead to a common requirement with harmonized methodology for **LOANS**



AnaCredit - a harmonized loan data collection of the European Central Bank since 2018

https://www.ecb.europa.eu/stats/money_credit_banking/anacredit/html/index.en.html

- Compulsory for euro area, Extra-EU countries may join – Hungary has not joined AnaCredit
- Scope: credit to **legal persons** – loans to individuals are excluded!
- Around 100 attributes, tables for instruments, counterparties and collaterals
- Reporters: credit institutions of euro area
- Threshold: 25,000 EUR

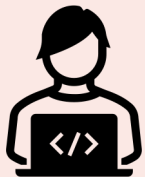
Although AnaCredit was a big step forward in European data collection, the current coverage of the database is not sufficient for the supervision and resolution activities of the central bank. The Magyar Nemzeti Bank therefore started building its own, broader database.

THE MNB HAS UTILISED CONSIDERABLE RESOURCES TO BUILD ITS OWN CREDIT REGISTER



HUNGARIAN CENTRAL BANK CREDIT REGISTER -

Biggest credit
register in
Europe



- ✓ Implemented in **2019** - more than 4 years experiences
- ✓ **Monthly frequency** – some attributes reported at quarterly frequency
- ✓ 59 data suppliers - all (36) credit institutions and 23 other financial institutions

- ✓ Normalized, flexible data model - extended data model of AnaCredit
- ✓ 25 tables
- ✓ 750 attributes, of which ~500 meaningful attributes above identifiers

- ✓ All partner sectors including individuals in an anonymised way

- ✓ **Comprehensive data quality framework**
- ✓ 2300 entry-point validation rules – blocking and plausibility checks
- ✓ Several other data quality checks (cross-checks, outlier detections, ML, etc.)

- ✓ ~ **6 millions credit instruments/months** (including off-balance sheet items)
- ✓ ~ **150 thousands corporations** as debtor
- ✓ ~ **4 millions individual debtors**
- ✓ **Number of records** stored from the beginning: **2.2 billion**

THE MNB'S CREDIT REGISTER IS USED FOR A WIDE RANGE OF CENTRAL BANK TASKS



UTILIZATION OF THE CENTRAL BANK CREDIT REGISTER

Multipurpose - serves **core** and **supervisory** central bank functions

Other utilizations

CORE FUNCTIONS

SUPERVISORY FUNCTIONS

- ✓ regular and ad-hoc analysis
- ✓ financial stability
- ✓ monetary policy forecasting
- ✓ stress tests
- ✓ minimum reserve strategy



offers new possibilities in data driven supervision

- ✓ closer cooperation with the supervised institutions
- ✓ risk modelling
- ✓ increasing the efficiency of the early warning reporting

- ✓ Dissemination
- ✓ Replacement of aggregated data collection
- ✓ Data transmission to the members of national official statistical service



THE CREDIT REGISTER CAN SERVE AS A BASIS FOR FURTHER DATA COLLECTIONS AND SUBSTITUTE AGGREGATE DATA REPORTINGS



EXPERIENCES



High efforts for ensuring data quality



New IT solution in data collection, data quality management and compilations of granular loan data



Intensive coordination among the stakeholders – reporting agents, users, IT partners, etc.

WAY FORWARD

Granular data collections for other data areas – e.g. deposits



More intensive use of administrative data sources, linkage of available databases

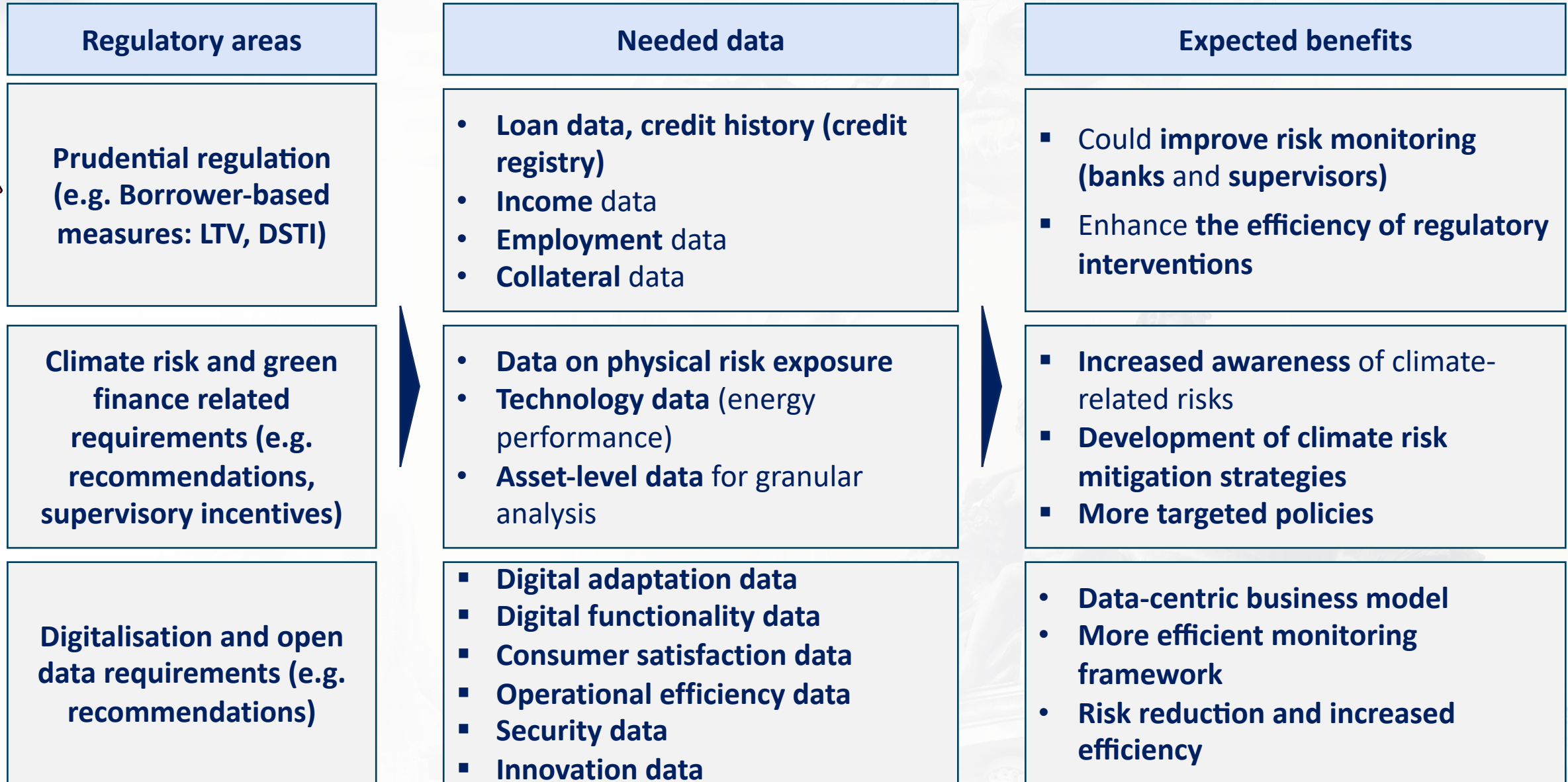
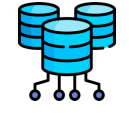


Connection of available databases – further utilization



Further decrease of reporting burden

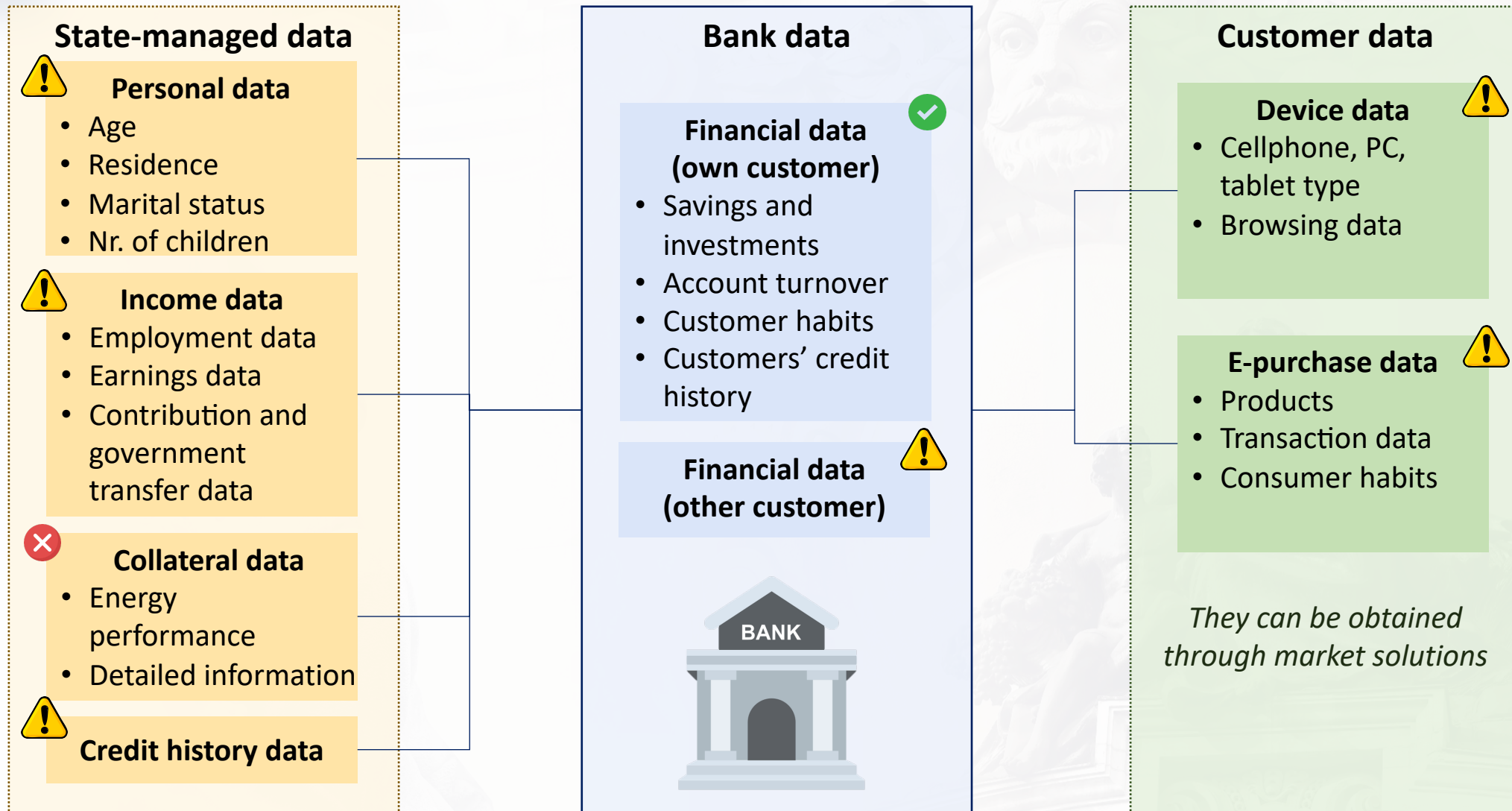
CENTRAL BANKS CAN DRIVE BANKS' DATA COLLECTION VIA PRUDENTIAL REGULATORY TOOLS AS WELL



CREDIT DATA: POSSIBILITIES AND CHALLENGES IN HUNGARY



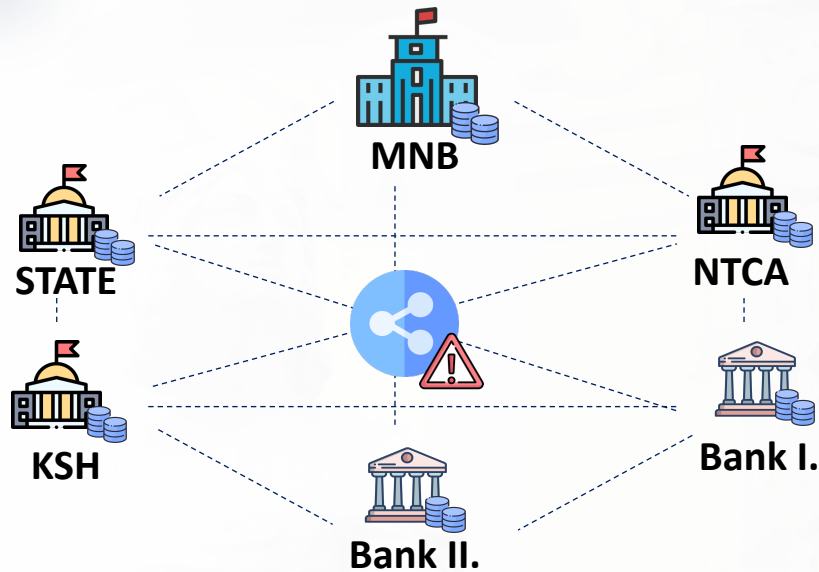
THERE IS A WIDE ARRAY OF DATA NOT CURRENTLY AVAILABLE TO BANKS FOR CREDIT SCORING PURPOSES



Note: Data used by banks according to the current Hungarian practice of electronic access to them.

✔ Readily accessible; ⚠ Consent needed; ✘ Not accessible

DOMESTIC DATABASES OPERATE IN ISOLATION: ACCESS AND USABILITY ARE LIMITED



- ISOLATED, SEPARATE FUNCTIONING OF AUTHORITIES, WITH LITTLE COORDINATION
- THE POSSIBILITIES OF OPEN BANKING (PSD2) ARE UNTAPPED

Disadvantages

Authority databases:

- Data sharing is difficult (contract, channel, no conceptual design)
- Methodological differences, data provider burdens, regulatory burdens
- Tracking changes, quality control
- Lack of information: what data is available in what form for which period and from whom

Bank databases (PSD II):

- PSD2 requires banks to create an API access option → however, the APIs providing data access do not yet work at many banks or only work on paper → state involvement is necessary to ensure data access

EFFECTIVE ACCESS TO THE DATA WOULD SIGNIFICANTLY SUPPORT THE SPREAD OF ONLINE LENDING PROCESSES



Central Credit Information System



More than **1 million** retail loan contracts annually

Credit information database (positive and negative data) → expansion, development, wide access required

Energy performance certificates



Energy characteristics of **1.5 million** homes

The most important energy characteristics of real estate → automated bank access required

Central statistical valuation database



Nearly **100,000** omissible valuations per year

Data on which statistical valuation is based → a central database is required to access data on a level playing field in mortgage lending

E-Land Register



Nearly **100,000** housing loans disbursed annually

Digitization of Hungarian land registry procedures and access to a wider range of data → introduction recommended as soon as possible

Online earnings statement



More than **1 million** income certificates per year

Loan application without personal administration, faster and easier evaluation → expansion and development required

Necessary regulation: reducing legal barriers, creating a supporting framework (Government, Parliament)

**THANK YOU FOR
YOUR ATTENTION!**

