

# Regional Workshop on Statistical Business Registers and Industrial Classifications

---

Session 4b: Seven Dimensions of the SBR Maturity Model

Julian Chow, UNSD

Amman, Jordan, 25 – 27 June 2024



Shared Prosperity Dignified Life



# Maturity Model of Statistical Business Registers

## Seven dimensions

(D1) Legal and institutional framework

(D2) Data sources for SBR

(D3) Maintenance and update of SBRs

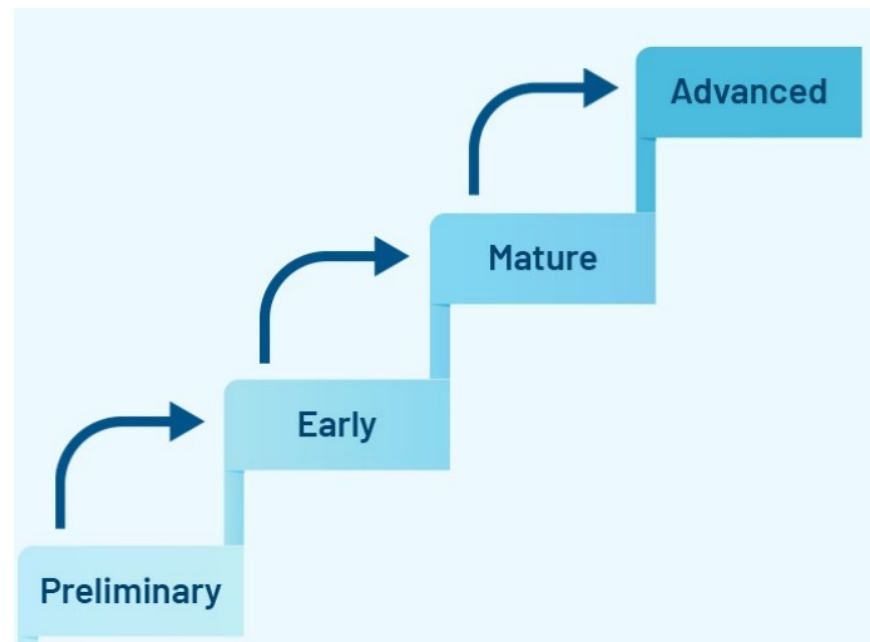
(D4) Coverage of an SBR

(D5) Use of an SBR

(D6) IT environment

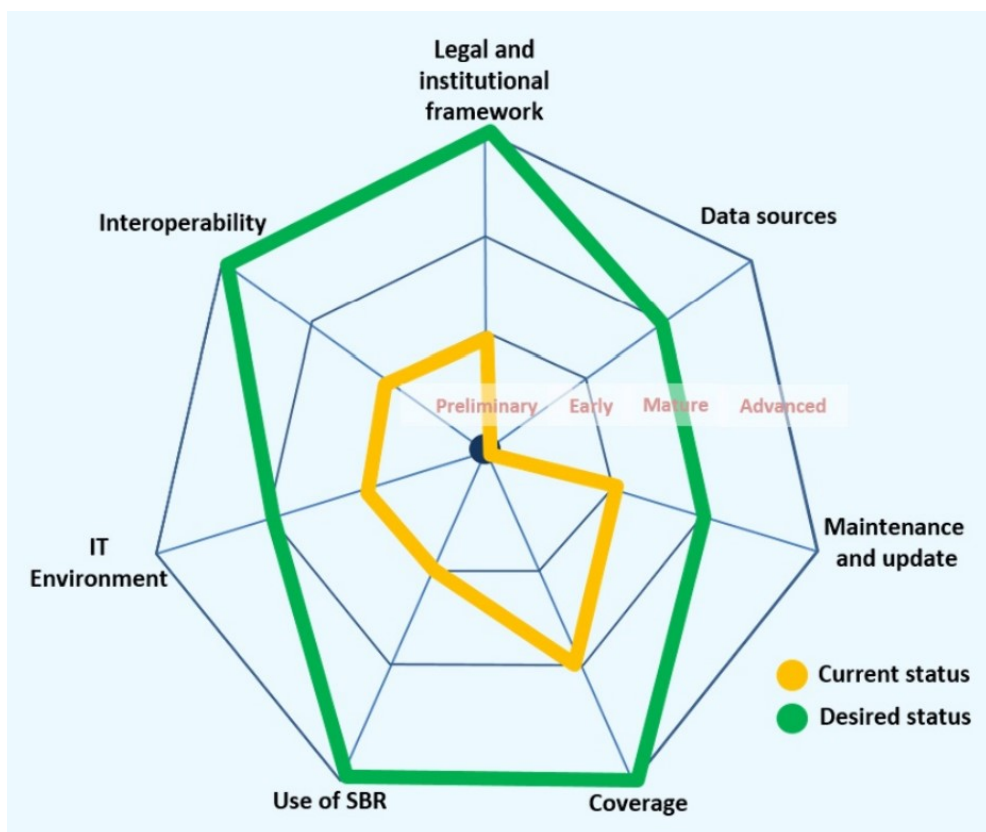
(D7) Interoperability

## Four maturity stages



## 2) Toolkit on SBRs (under construction)

### A questionnaire for self-assessment



**A list of training materials**



**Knowledge and practice sharing**



**Regular global assessment**

# Seven Dimensions of the SBR Maturity Model

---

Dimension 1: Legal and institutional framework

- Dimensions 2 & 3: Data sources, maintenance and updates

# Dimension 1: Legal and Institutional Framework



Shared Prosperity **Dignified Life**



# Dimension 1: Legal and Institutional Framework

## Legal Framework

- Statistical law
- Memorandum of Understanding or Service Level Agreement

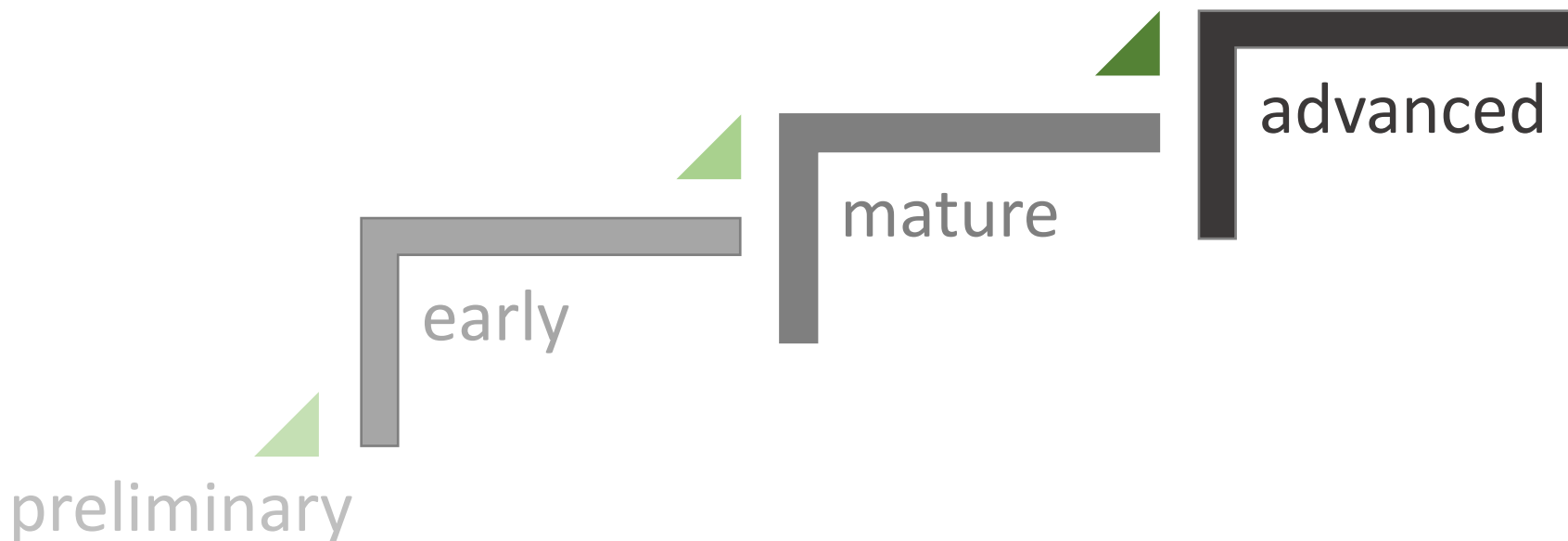
## Institution Framework

- Arrangements among the relevant institutions in the country for the maintenance and update of SBRs.



# (D1) Legal and institutional framework

This dimension refers to the legal and institutional framework relevant for establishing and maintaining a SBR.



# (D1) Legal and institutional framework

- A sound legal and institutional framework that regulates official statistics and, in particular, **the rights and obligations related to data collection and acquisition, including access to administrative records and other relevant sources**, provides an essential basis for the development and maintenance by the NSO of an effective and efficient of SBR.
- The legal framework refers primarily to the **Statistical Act** or **Statistical Law** regulating official statistics in a country but also other legal instruments and provisions regulating privacy and data protection, access, sharing and archiving.
- Some common elements are identified and elaborated in the Generic Law on Official Statistics (GLOS).

[https://unece.org/DAM/stats/publications/2016/ECECESSTAT20163\\_E.pdf](https://unece.org/DAM/stats/publications/2016/ECECESSTAT20163_E.pdf)



# (D1) Legal and institutional framework

- Aspect of Generic Law on Official Statistics (GLOS) relevant to SBR
  - Mandate for data collection (Article 15 of GLOS)
  - Access to administrative data (Article 17 of GLOS)
  - Legislative mandate for NSO on the establishment and maintenance of SBR (Article 19 of GLOS)
  - Data subject to statistical confidentiality (Article 20 of GLOS)
  - Access to individual data of the National Statistical System (Article 23 of GLOS).
  - Provision for data dissemination (Article 29 of GLOS).

# (D1) Legal and institutional framework

- Beyond the Statistics Act or Law, other provisions in administrative rules and regulatory frameworks can support the development and maintenance of SBRs. These include.
  - The establishment and use of a unique identifier for businesses in the country
  - Registration requirements for countries that have a business register, which is a data source for the SBR. Recommendations on business registers can be found in the Legislative Guide on Key Principles of a Business Registry, prepared by the United Nations Commission on International Trade Law (UNCITRAL).
  - Exemptions for the NSO from specific provisions in data privacy, protection, sharing and archiving acts to access and process microdata and to combine microdata from various sources for statistical purposes, including the development and maintenance of statistical registers.

## Business registers vs Statistical Business Registers

### Business Registers



It contains legal entities

Typically, businesses have to register in order to carry out the economic activity

Typically maintained by the business registrar authority

### Statistical Business Register



It contains statistical units for the collection and compilation of economic statistics

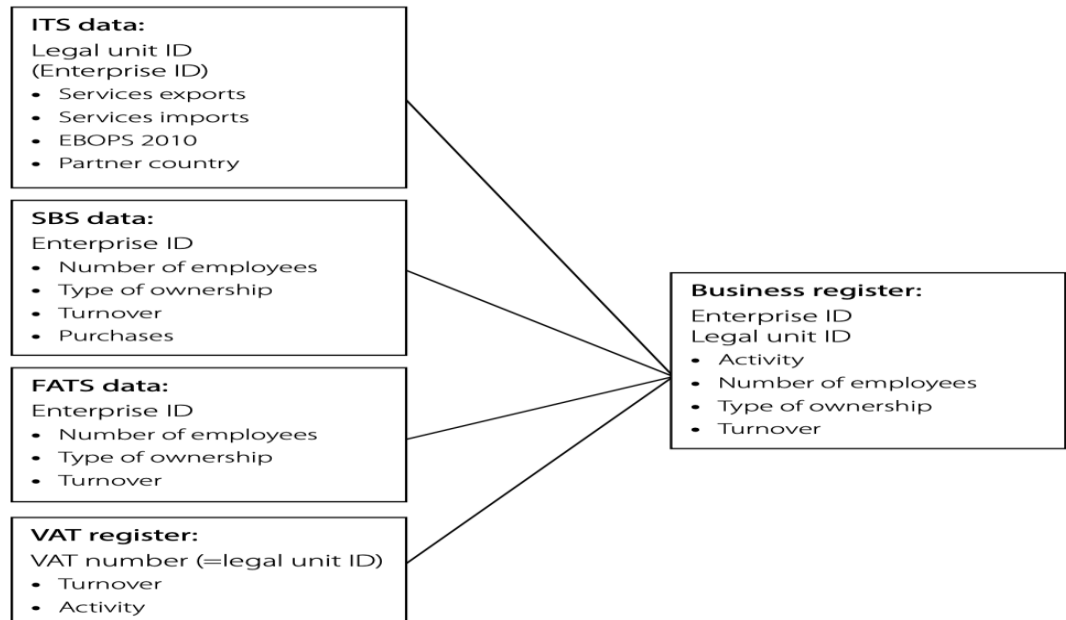
Maintained by the National Statistical Office

# (D1) Legal and institutional framework

- The availability of a national institution that is responsible for the registration of businesses is an important element for the SBR.
- A business register, administrative business register, or business registry are different from the SBR as they represent the country's mechanism for receiving, storing and making accessible to the public certain information about businesses, as required by domestic law.
- Most countries have a business registry supported by the legislation that require businesses to register in order to participate in the formal economy.
- When available, they represent an important source of information for SBRs.

## Use of unique legal identifiers

- Information on the same company has to be linked when combining information from different sources
- The availability of a unique legal identifier (among administrative data) is key for an efficient and effective micro data linking



# (D1) Legal and institutional framework

- A **memorandum of understanding (MOU)**, a **service level agreement (SLA)**, or similar arrangement can be used to formalize, in a less binding arrangement, the cooperation between institutions to access administrative data.
- These MOUs generally cover data flows, metadata, communications, protection of the confidentiality, and often include a clause ensuring that the NSO will be informed in advance of any changes made to administrative process that will affect the resulting data.

# (D1) Legal and institutional framework

- The Guidelines on the Legislative Framework for Civil Registration, Vital Statistics and Identity Management (United Nations 2019) , notes that the legal framework can cover the following aspects:
  - business registration;
  - business registers;
  - unique identifiers, data protection and privacy;
  - compliance, enforcement, rights and remedies; and transitional provisions.

# (D1) Legal and institutional framework

- Institutional arrangements refer to the arrangements among the relevant institutions in the country for the maintenance and update of SBRs.
- A clear understanding of the roles of the different institutions is important to establish an institutional arrangement. The NSO is typically the institution responsible for the SBR, and it is vital for the NSO to establish and maintain good relations with other institutions, especially the owners who are responsible for data sources.
- Depending on the nature of the national statistical system; i.e., centralized vs decentralized, there may be special institutional arrangements that guarantee the sharing of relevant information. Institutional arrangements can also be put in place with relevant agencies within the country, such as the institution in charge of businesses registration.



# Summary of stages of maturity for Dimension 1

## Preliminary

- **No** statistical law or institutional arrangement and **no** MoU are in place in support of the SBR

## Early

- There is a **provision** in the statistical law and/or MoU that mandates the **NSO to collect data** that are relevant for SBRs
- Access to administrative data is **not guaranteed** by legislation or MOUs

## Mature

- A statistical law allows the NSO to **establish and maintain the SBR**
- A statistical law gives the NSO **access to relevant administrative data** for the SBR

## Advanced

- A statistical law allows the NSO to **release the public anonymized microdata**, while complying with the principle of confidentiality

# (D1) Legal and institutional framework

early

preliminary

## Preliminary stage:

- **No** statistical law or institutional arrangement and **no** MoU are in place in support of the SBR

## Early stage:

There is a legal framework or MoU that the NSO to collect data that are relevant for SBRs

### Example: Economic census

- a large network of enumerators to collect information
- data entry staff and managers to guide this process

→ the legislation needs to be supported by appropriate funding.



# (D1) Legal and institutional framework

early

mature

## Early stage:

An administrative business register may or may not be available.

\* Access to administrative data is **not guaranteed** by legislation or MOUs

## Mature stage:

A statistical law allows the NSO to **establish and maintain the SBR**

A statistical law gives the NSO **access to relevant administrative data** for the SBR

### Registration or inclusion of businesses

- regular access to and use of all necessary census, administrative business register or other administrative data.
- legal agreements and MoUs with several source owners

# Question - (D1) Legal and institutional framework

*Can you think of other characteristics at the mature stage?*

# Question - (D1) Legal and institutional framework

*Can you think of other characteristics at the mature stage?*

1. The use of administrative data in statistical production is mandatory where this is available.
2. Using a survey as an instrument to collect data with respondents is only permitted where information to produce reliable indicators is not available in administrative sources.
3. The legal framework covers aspects of business registration (enforcements of registration of new business (business births) and the cessation of businesses (business deaths); unique identifiers of business in registers; data protection and privacy; compliance, enforcement, rights and remedies; and transitional provisions.

# (D1) Legal and institutional framework

mature

advanced

## Mature stage:

Access to microdata services of the NSO are granted for public organizations for statistical purposes only.

The access is not allowed for commercial users outside of NSO.

## Advanced stage:

- A statistical law allows the NSO to **release the public anonymized microdata**, while complying with the principle of confidentiality
- There is a legislation on data protection, privacy and confidentiality that allows:
  - NSO to disseminate selected information in the SBRs;
  - NSO to share statistical information/micro data with outsiders (commercial users).

Businesses are granted:

- access to their information stored in the SBR
- allowed to update relevant information about their own organization.

# Question - (D1) Legal and institutional framework

*Can you think of other characteristics at the advanced stage?*

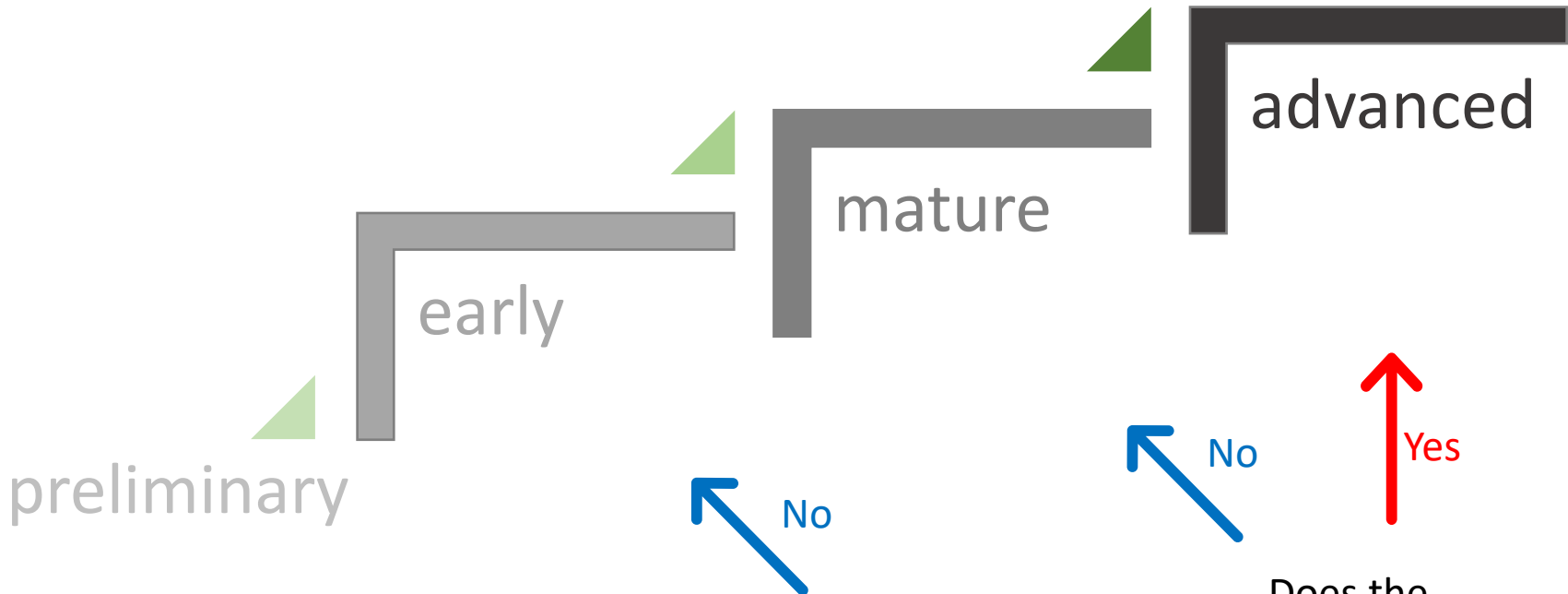
# Question - (D1) Legal and institutional framework

*Can you think of other characteristics at the advanced stage*

1. In an international context, regulations are in place to ensure the exchange, use and integration of national data, to create an international SBR.
2. Web scraping can be used for statistical production when it is not in conflict with other legislation.



# (D1) Legal and institutional framework



Do you have legal framework or MoU in support of SBRs?

**No** (blue arrow pointing up)

**Yes** (red arrow pointing right)

Does it enable NSOs to collect data that are relevant to SBR?

**No** (blue arrow pointing up-left)

**Yes** (red arrow pointing right)

Does it ensure regular access to and use of all necessary census, administrative business register or other administrative data?

**No** (blue arrow pointing up-left)

**Yes** (red arrow pointing right)

Does the legislation on data protection and privacy allow the NSO to disseminate selected information in the SBRs?

**No** (blue arrow pointing up-left)

**Yes** (red arrow pointing up)



Shared Prosperity Dignified Life



# Discussions

5. Give a brief description on the legal framework on official statistics in your country, with a focus on the aspects that are relevant in the production of business statistics.

- Is there a provision in the statistical law or a MOU (Memorandum of understanding) that mandates or allows the NSO (National Statistical Organization) to collect data that are relevant for SBR?
- Is there any provision on how your office should collect data on businesses?
- Is there any provision that grants NSO access to administrative data sources (e.g. records of businesses from the tax office)? If not, is there any institutional arrangement on data sharing in general among the NSO and other government agencies at various levels (national, regional or local governments)?
- Is there any mandate on the use of unique IDs for businesses in your country? If not, is there any institutional arrangements that helps the sharing of micro-data on businesses?
- Is there any provision on how data on businesses should be compiled and disseminated? If not, can you show some published reports or datasets on business statistics?

# Dimension 2: Data sources for the SBR

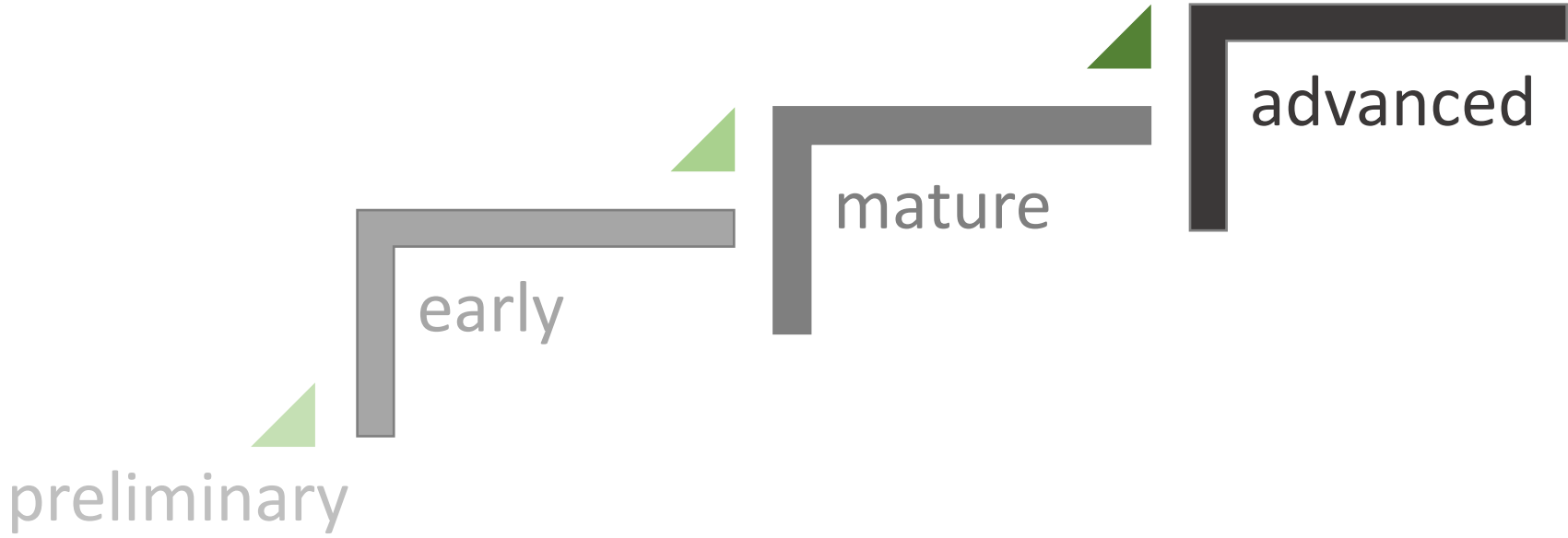


Shared Prosperity Dignified Life



# (D2) Data sources for the SBR

This dimension refers to the the characteristics of the SBR in terms of the data sources used for establishing and maintaining the SBR



## (D2) Data sources for the SBR

Data sources may be grouped into three categories:

- Administrative sources, such as business registration/license register, tax register, company/trade associations and chambers of commerce registers, social security registers, etc.
- Statistical sources: economic census and agricultural census, feedback from economic surveys, profiling and SBR improvement/maintenance surveys.
- Other sources: for example data from private data suppliers, telephone directories and the Internet, big data

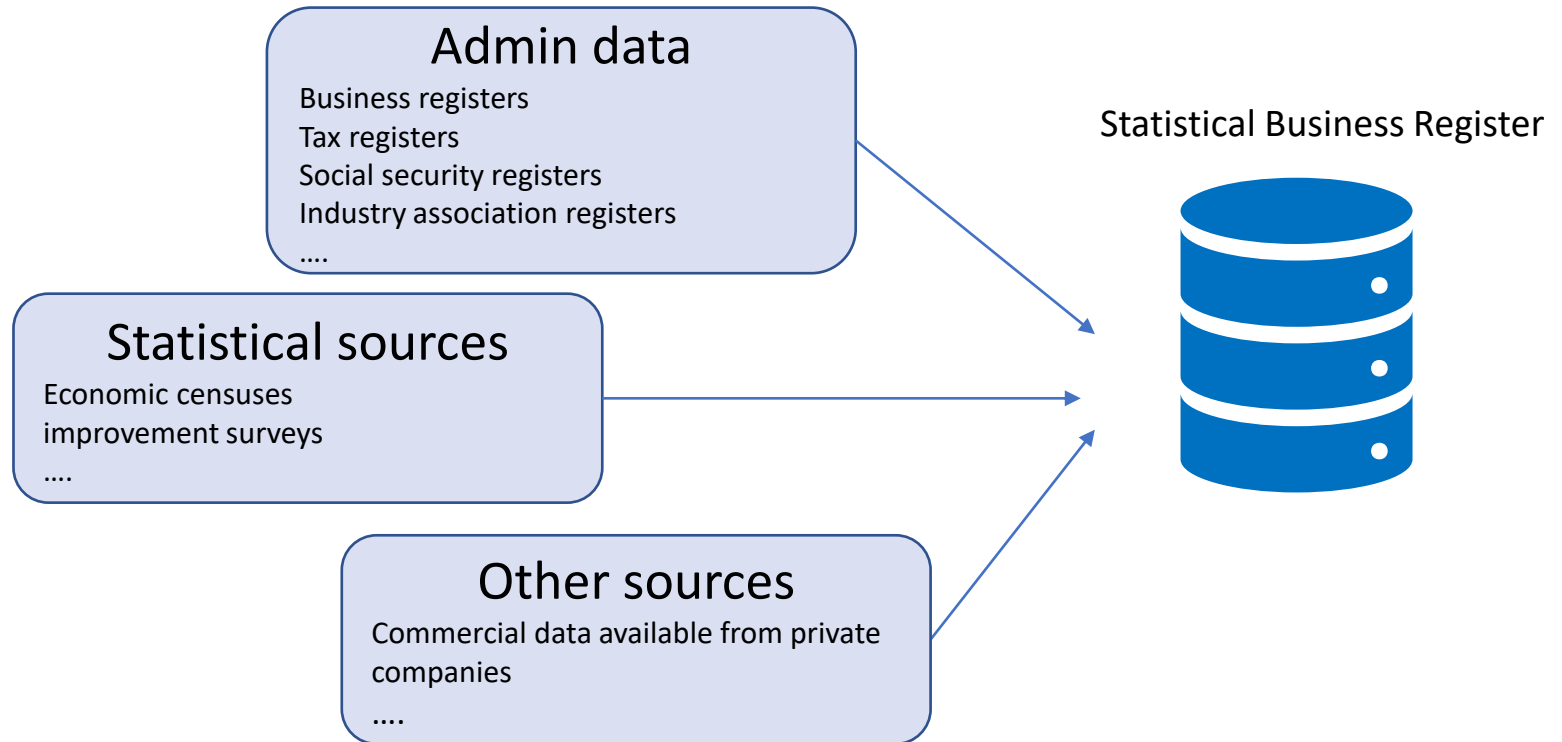
# Dimension 2: Data sources for the SBR

## Data Sources

- **Administrative sources**  
(*licenses register, tax register, social security register, labor and employment registers*)
- **Statistical sources**  
(*Economic Census, Enterprise/establishment surveys, SBR improvement surveys*)
- **Other sources**  
(*commercial data providers, utilities, internet companies*)



## Data sources for SBRs



## (D2) Data sources for the SBR

- The data sources depend greatly on the country's situation and the availability of relevant information from other institutions.
- Therefore, the choice of the best data source depends very much on the specific context of a country.
- However, in line with the UN Guidelines on Statistical Business Registers (para 6.4 UN 2020), the recommendation is for SBRs to be created and maintained primarily using administrative sources.
- They enable good coverage and stability. This approach is in line with Principle 5 of the United Nations Fundamental Principles of Official Statistics, which states “Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records.
- Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents



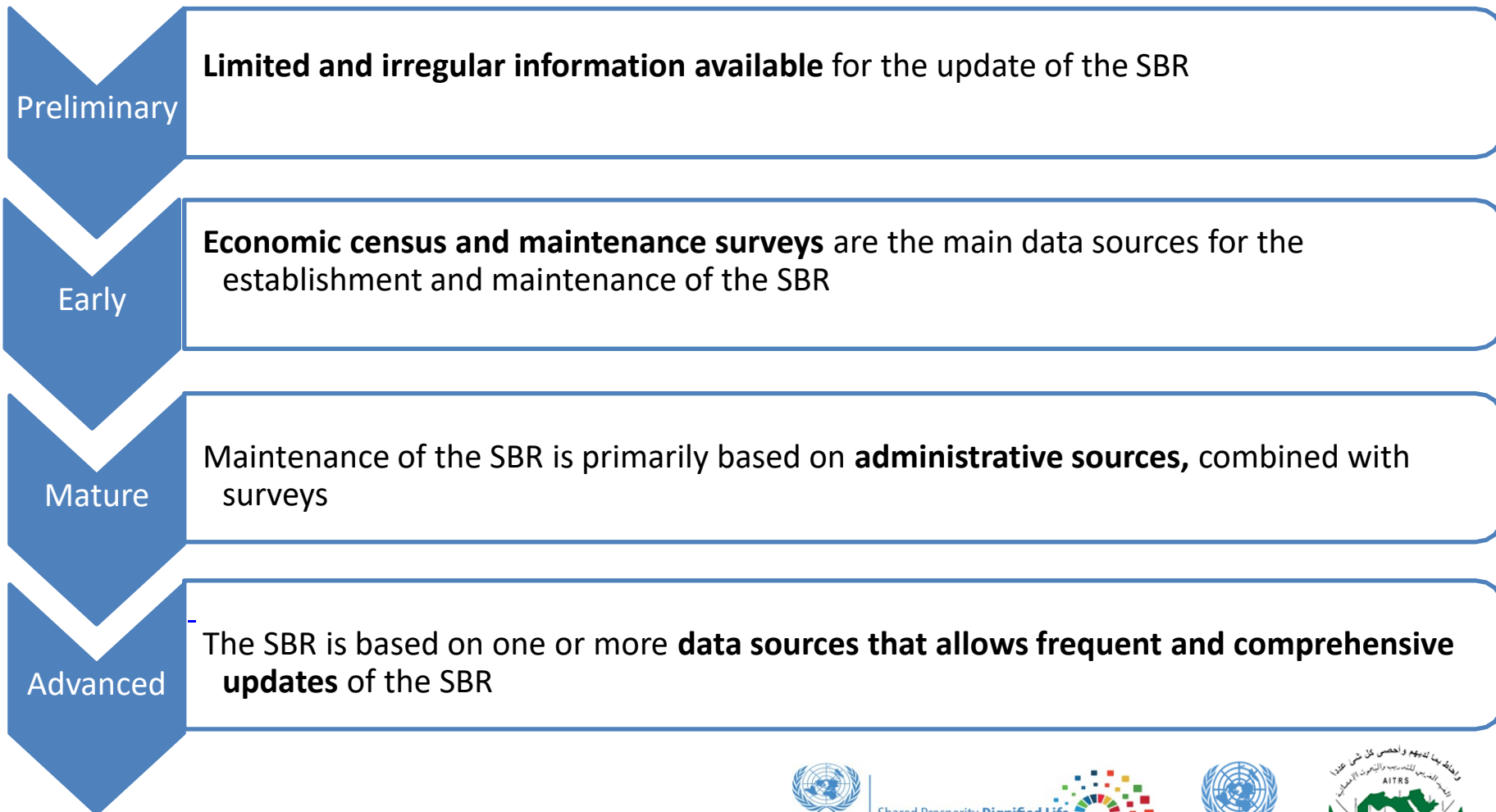
# Limitation on the use of economic censuses as a main source for the maintenance an SBR

- Economic censuses are very expensive even if the scope is restricted to businesses with identifiable premises. They are generally conducted at best every five or ten years. Thus, NSOs should consider a suite of annual surveys based on a register-based list to be much more cost-effective.
- SBRs need to be updated between economic censuses. The small businesses located during area enumeration are volatile in the sense that they may go rapidly in and out of production, or ownership, or change their activities or addresses. Thus, to be truly effective as a source of survey frames in the years between censuses, the SBR must be constantly updated, which itself can be costly.
- In addition to its high cost, the enumeration approach has the disadvantage of not being able to identify and document non-recognizable places of business, or enterprises without a fixed location, for example web-based businesses or individual entrepreneurs such as electricians and plumbers providing services at locations other than their homes.

# Further observations

- It should be recognized that there are countries where the NSO does not have access to administrative data, comprehensive administrative data do not exist, or the quality of the administrative data is not sufficient for the update of the SBRs.
- In these countries, the traditional economic censuses or other kinds of business surveys, as well as household surveys, are therefore an indispensable source for the SBRs.
- Economic censuses, mostly based on the establishment unit, deliver relevant and core information for the SBR on each establishment, such as address, economic activity, legal unit, number of employees. Even in countries that are able to use administrative data for their SBRs, data from censuses may be used as a complementary source for checking or updating the SBR.
- The preferred data sources depend on the specific situation in any given country, including the availability of administrative data and the scope and complexity of the national statistical system itself.
- However, efforts should be undertaken by countries in exploring potential administrative sources for use in the SBR, understanding their coverage and concepts, and assessing their quality. The improvement of the quality and timeliness of the SBR goes hand-in-hand with the improvement of administrative data. In addition, the NSO should aim to actively partner with administrative data providers and influence datasets for use in the SBR.

# Summary of stages of maturity for Dimension 2



# (D2) Data sources for the SBR

early

preliminary

## Preliminary stage:

### No SBR / Economic Census

(i.e. no up-to-date comprehensive data sets on establishments/enterprise/  
business units for statistical purposes/uses)

- \* Scoping exercises may be underway.

## Early stage:

The construction of the SBR depends mainly on the economic census  
together with maintenance surveys.

- \* There may be a periodic economic census used to update the SBR

## (D2) Data sources for the SBR



### Early stage:

No profiling or substantial data confrontation with other sources is undertaken.

- \* In general, the maintenance procedures run on a reference period with a significant lag (This hampers the survey data collection)

An assessment of the availability and quality of administrative data for SBRs is being undertaken.

### Mature stage:

The SBR is updated on a sub-annual basis using mainly administrative sources, such as business registrations, taxation data, etc.

- \* Data collection is more efficient because the maintenance procedures applied are close to the actual reference period.

# Question - (D2) Data sources for the SBR

*Can you think of other characteristics at the mature stage?*

# Question - (D2) Data sources for the SBR

*Can you think of other characteristics at the mature stage?*

1. Operational rules enable the integration and transformation of administrative data into statistical units.
2. There is a nationally consistent unique identifier used to maintain and identify economic units from a whole of government perspective.
3. More attention is given to national and possibly international profiling.

## (D2) Data sources for the SBR

mature

advanced

### Mature stage:

The SBRs is updated on a continuous and frequent basis and reflects almost in real time the 'events' that are recorded in the administrative data.

### Advanced stage:

There are multiple sources used to update the SBR on a sub annual basis. Commercial data sources are also used in the maintenance strategy of the SBR.

Businesses are granted:

- access to their information stored in the SBR
- rights to update relevant information about their own organization.



# Question - (D2) Data sources for the SBR

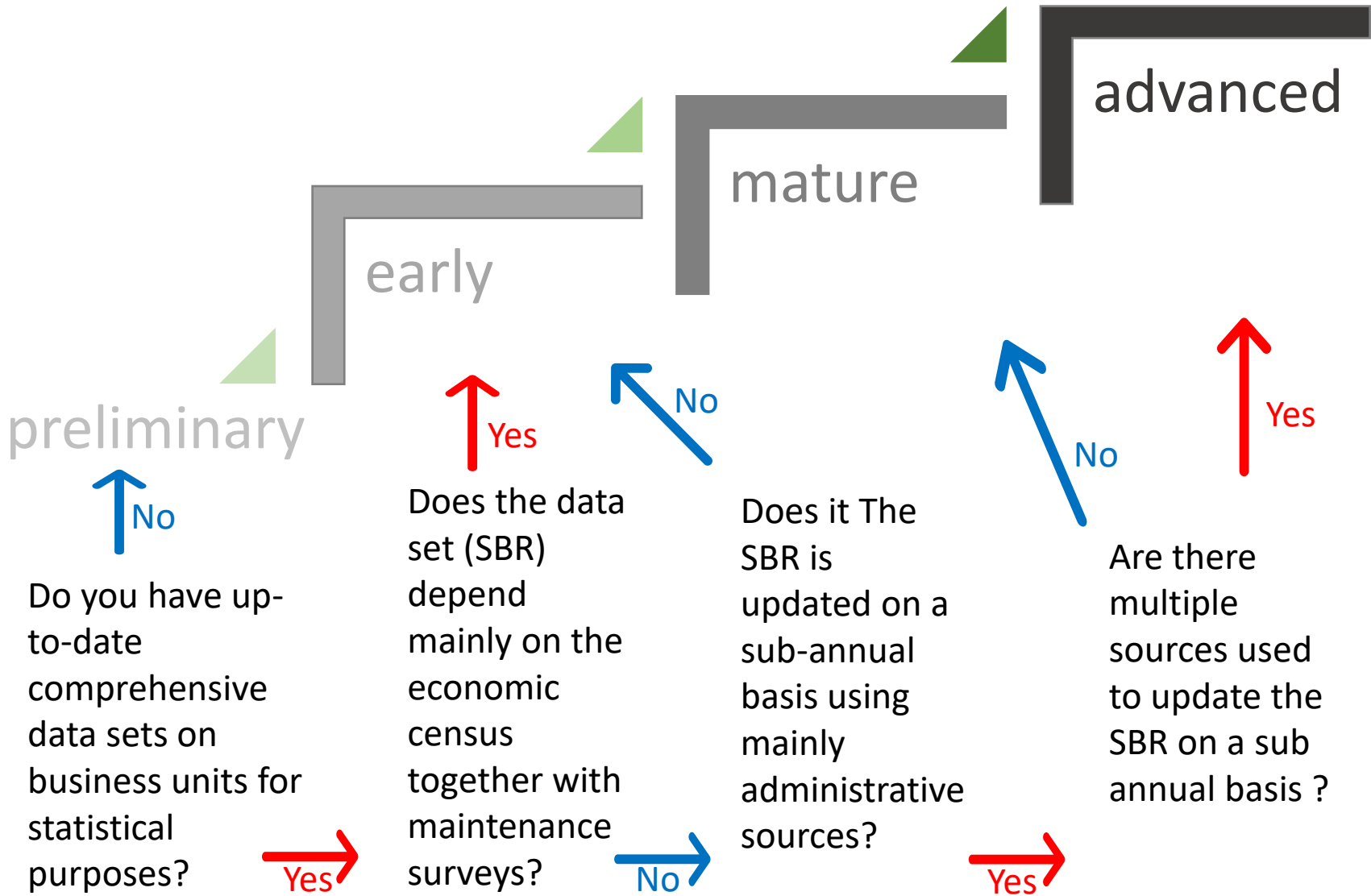
*Can you think of other characteristics at the advanced stage?*

# Question - (D2) Data sources for the SBR

*Can you think of other characteristics at the advanced stage?*

1. Where there is enabling regional legislation, data from other countries can be used to update the SBR.
2. Web scraping and data mining techniques are also used to improve or to verify the content of the SBR.
3. International coordination is optimized resulting in cooperation and data sharing between countries.

# (D2) Data sources for the SBR



# Discussions

1. Discuss the feasibility of using the following data sources to update the SBR in your country:
  - a. Economic census (frequency of the census, estimated numbers of units, coverage of the units, and if there is any update on the variables in the years between censuses)
  - b. Surveys (how will the results of the surveys be used to improve the quality of SBR)
  - c. Administrative data and tax registers (where to find newly established enterprises/legal units, and the participants are welcome to share their experience on data sharing on businesses between NSOs and other government agencies in general)
  - d. [Optional] Internet and big data (does your office have any experience in using big data or other non-traditional data sources?)

# Dimension 3: Maintenance and update for the SBR

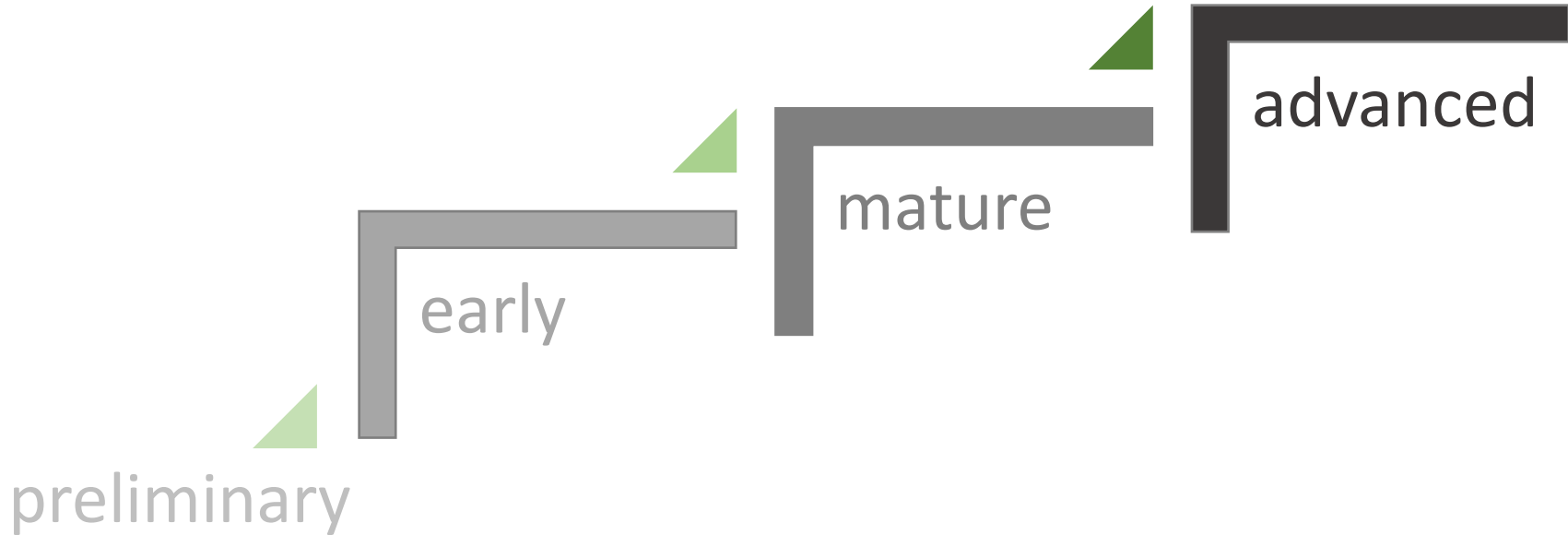


Shared Prosperity **Dignified Life**



# (D3) Maintenance and update of the SBR

This dimension includes the maintenance and update procedures for the SBR, and in particular the operational requirements of maintaining effective and reliable systems.



# Dimension 3: Maintenance and update of the SBR

## Maintenance /Update

- Maintenance groups
- Changes in characteristics of statistical units
- SBR maintenance procedures
- Consideration in dealing errors



# Dimension 3:

## Maintenance and update of the SBR

- In general, the key objective of maintenance is to update the coverage and content of the SBR, taking into account continuity and stability rules, according to a well-defined calendar, and in as timely a fashion as the information sources allow. This is so the SBR is able to provide economic surveys with sampling frames that are accurate and as up to date as possible
- The operational requirements of a maintenance program include the modification of records and preservation of stored records. Modernizing and maintaining the operational requirements is a prerequisite of contemporary functions of NSOs, essential to improving services to the public



# Dimension 3:

## Maintenance and update of the SBR

Part of the maintenance of the SBR includes the digitalization and automatization of procedures. However, these aspects are elaborated under the IT dimension of the Maturity Model. In this dimension, the focus is on the following aspects:

- the sources of information described in Dimension 2 (e.g., administrative sources, economic censuses, feedback from surveys and SBR improvement surveys) and rules for dealing with conflicting information;
- continuity and stability rules determining whether a statistical unit is deemed to have continued despite significant changes (for example, of ownership, size, economic activity and/or location), or to have died and been replaced by another unit, as are stability rules (sometimes called resistance rules) that restrict the speed with which changes of characteristics are applied in order to inhibit unwanted oscillations in values;
- consistency in maintenance rules, procedures, methodologies, standards, definitions and concepts, and over time;
- quality evaluation of the SBR
- preservation of stored records (e.g., procedures for storing and preserving records);
- validation procedures;
- the frequency of updates

# Summary of stages of maturity for Dimension 3

Preliminary

The SBR is **not maintained or updated, or is updated irregularly**

Early

The maintenance and update procedures for the SBR are **being developed**

Mature

**Maintenance and update procedures and methods of making changes to records** are clearly defined and executed

Advanced

Maintenance and update of the SBR are conducted **on a continuous basis**

# (D3) Maintenance and update of the SBR

early  
preliminary

## Preliminary stage:

Effectively no maintenance activities undertaken as there is no SBR / Economic Census available.

\* Scoping exercises may be underway.

## Early stage:

Methods of making changes to the records, the ways of maintaining the integrity of the records, and keeping a log of all such changes are being developed.

# (D3) Maintenance and update of the SBR



## Early stage:

Procedures for storing and preserving records is not in place.

\*Internal review mechanisms for system functions may or may not be elaborated.

Validation procedures are not regularly done (ad-hoc and highly manual).

\* This is due to a lack of resources or a lack of suitable data sets for confrontation. (e.g. lack of unique identifier, databases or legal authority to collect information)

## Mature stage:

Procedures for storing and preserving records and internal review mechanisms for system functions are well established and functioning.

Validation procedures are done on a regular basis by confronting the multiple sources used to update the SBR.

# Question - (D3) Maintenance and update of the SBR

*Can you think of other characteristics at the early and mature stage?*

# Question - (D3) Maintenance and update of the SBR

*Can you think of other characteristics at the mature stage?*

Early stage

1. The economic census frame may be updated using surveys (not a dedicated one). In general, the maintenance procedures run on a reference period with a significant lag (This hampers the survey data collection)
2. The SBR team is relatively small compared to the population of the country and national statistical office.
3. The SBR is not updated on a regular basis.

# Question - (D3) Maintenance and update of the SBR

*Can you think of other characteristics at the mature stage?*

Mature stage

1. Methods of making changes to the records, the ways of maintaining the integrity of the records, and keeping a log of all such changes are well developed and documented.
2. There may also be a periodic census used to update the SBR as well.
3. The primary data sources are used to update the SBR on a regular basis at sub annual frequency.

# (D3) Maintenance and update of the SBR

mature

advanced

## Mature stage:

Maintenance procedures are in place to integrate and transform administrative data into statistical units through profiling.

\* Profiling team engages with the largest and most statistically significant organizations for a nation's economy.

## Advanced stage:

### Profiling

- covers a substantial proportion of a country economic activity in terms of industry value add;
- ensures that where there are significant contributors to economic activity;
- uses timely administrative data to identify units for investigation.

There is a nationally consistent unique identifier used to maintain and identify economic units from a whole of government perspective.



# Question - (D3) Maintenance and update of the SBR

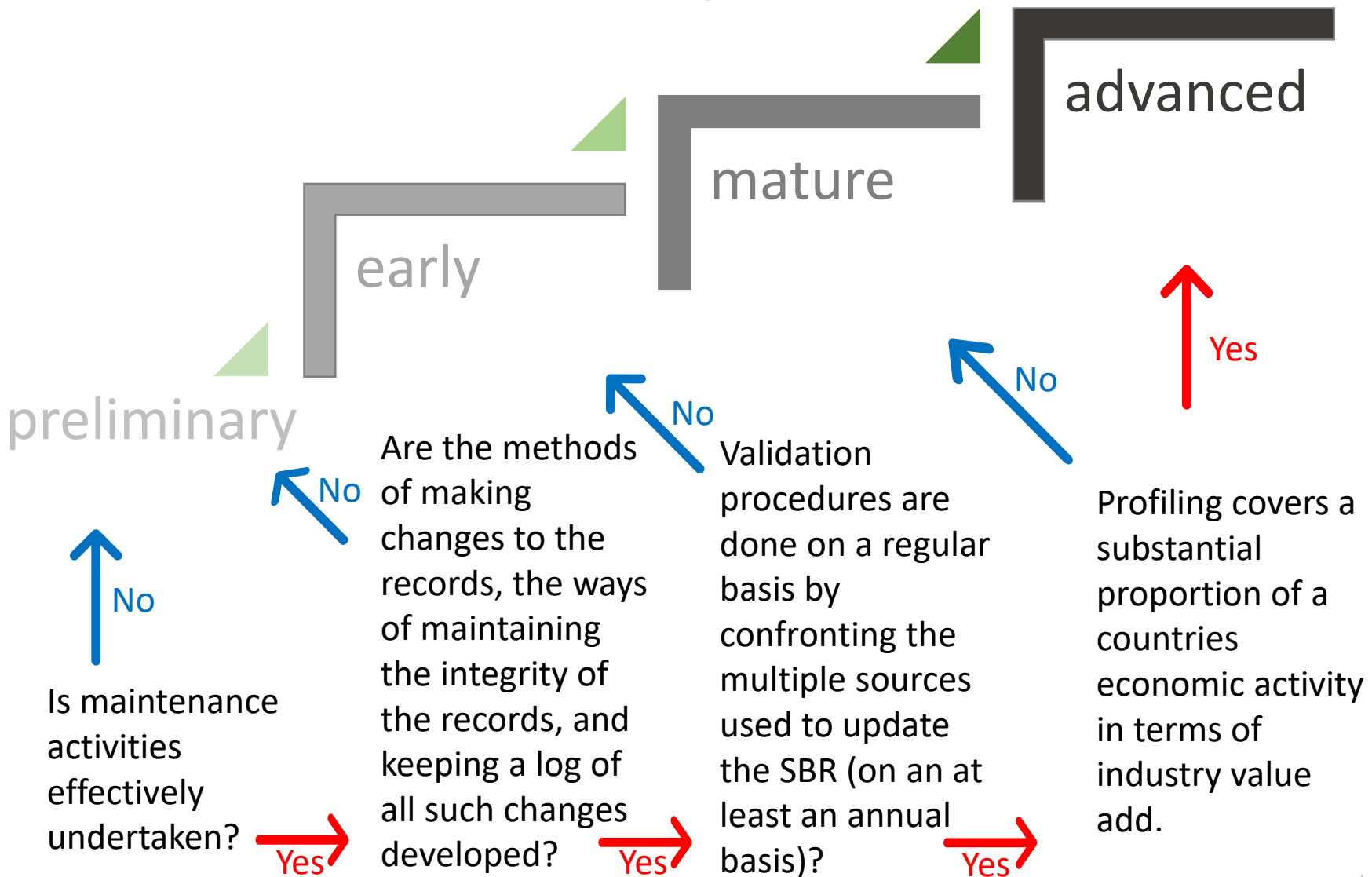
*Can you think of other characteristics at the advanced stage?*

# Question - (D3) Maintenance and update of the SBR

*Can you think of other characteristics at the advanced stage?*

1. The SBR is updated on a sub-annual basis in line with the updates from the data sources.
2. The government databases include all key data requirements for the SBR.
3. The SBR combines / integrates data from many sources so that it reflects the real world.
4. Such a system may have one data item derived from one administrative source but many other key data items derived from others; with the NSO determining the rules depending on reliability and timeliness.
5. The data items are all derived based on the economic unit's model.
6. Detailed unit record level reports are created which explain all key changes and are made available to all internal stakeholders.

# (D3) Maintenance and update of the SBR



# Discussions

Discuss the challenges of establishing an SBR in your country with regard to the following aspects:

- Current and new data sources of the SBR
- Maintenance and update of an SBR

# Quiz (1)

*'My country has a legal framework to use economic census but not administrative data; SBR is based mainly on economic census; update of the records is taking place systematically.'*

Q. Which of the following should be highly likely to be the assessment of their SBR?

- (a) D1. Preliminary, D2. Early, D3. Early
- (b) D1. Preliminary, D2. Preliminary, D3. Preliminary
- (c) D1. Early, D2. Early, D3. Early
- (d) D1. Early, D2. Preliminary, D3. Preliminary

# Quiz (1)

*'My country has a legal framework to use economic census but not administrative data; SBR is based mainly on economic census; update of the records is taking place systematically.'*

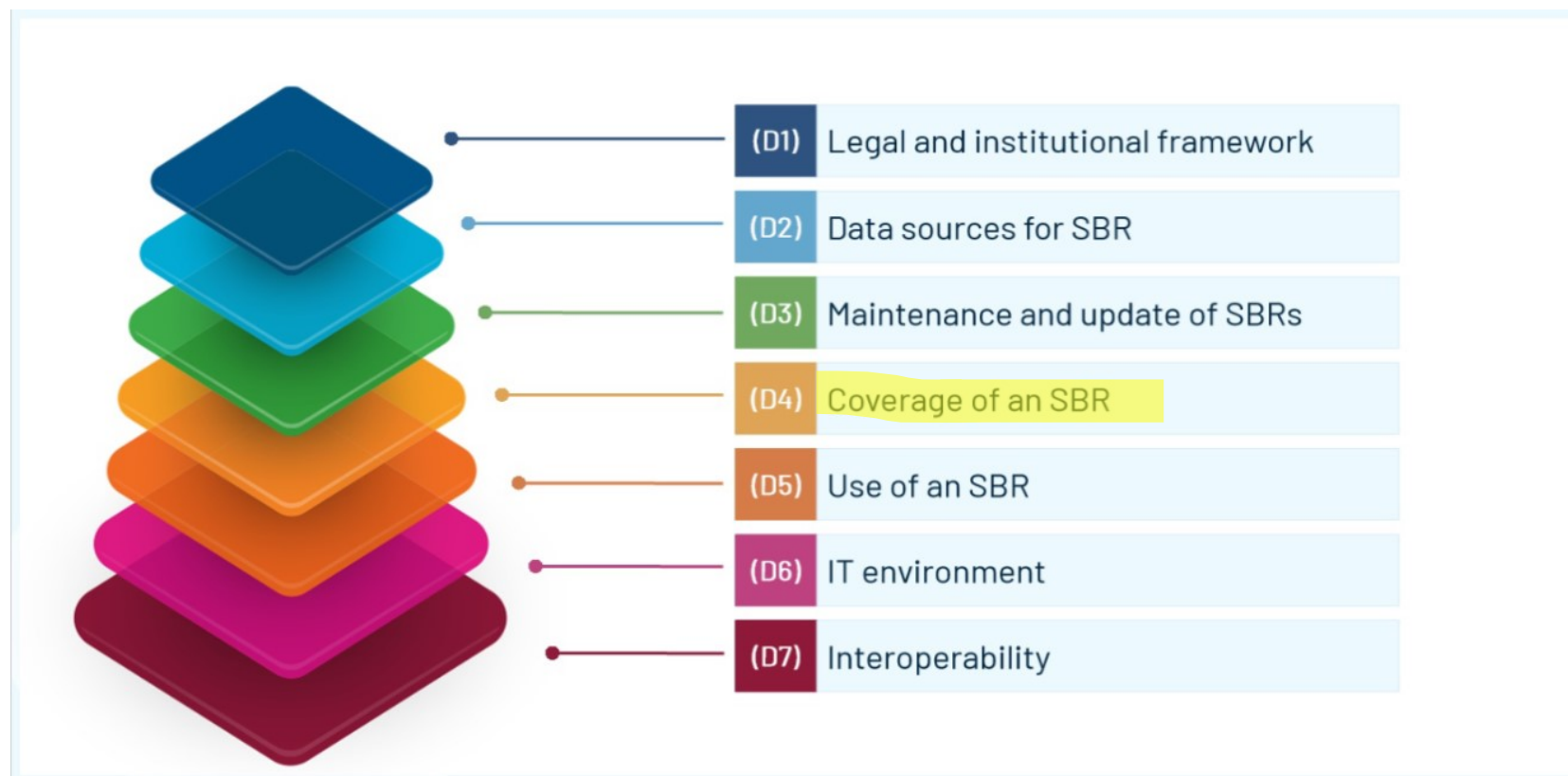
Q. Which of the following should be highly likely to be the assessment of their SBR?

- (a) D1. Preliminary, D2. Early, D3. Early
- (b) D1. Preliminary, D2. Preliminary, D3. Preliminary
- (c) D1. Early, D2. Early, D3. Early**
- (d) D1. Early, D2. Preliminary, D3. Preliminary

# Seven Dimensions of the SBR Maturity Model

- Dimension 4: Coverage of an SBR
  - Dimension 5: Use of an SBR

# Seven dimensions





## Dimension 4: Coverage of an SBR

There are three key aspects:

- **Completeness** is the extent to which the SBR includes all institutional units within the 2008 SNA production boundary
- **Coverage** is the proportion of total national economic production that the units represent.
- **Content** is the set of characteristics (e.g., types of units, institutional sector, size, location and registration status) of the units contained in the SBR.

# Coverage of an SBR

- In principle, an SBR should record all institutional units in the national economy that are engaged in productive economic activities; i.e., activities contributing to the gross domestic product (GDP). Thus, NSOs should aim to cover as much national production as possible in their SBRs by including all types of institutional units engaged in production.

***However, where complete coverage is not easily attainable in practice, and a more realistic aim for coverage can be taken.***

# Coverage of an SBR

- One particular issue is coverage of the informal economy, which is highly diversified and for which no administrative data exist.
- ***It is usually not recommended*** to focus on coverage of the informal economy in early stages of SBR development. However, in certain countries – particularly developing and emerging economies – the informal economy is important, and the collection of economic data about it is a priority.

# Coverage of an SBR

- In general, groups of units should not be omitted without an assessment of their economic importance. SBRs should aim to record, at a minimum, all active economic units in the formal economy that are engaged in economic activities contributing to the gross domestic product (GDP)

# Coverage of an SBR

- In general, **when first establishing an SBR, the corporations sector is the focus, as it is the most easily covered.**
- The inclusion of the government sector, non-profit institutions and the household's sectors provide more complete coverage.

# Coverage of an SBR

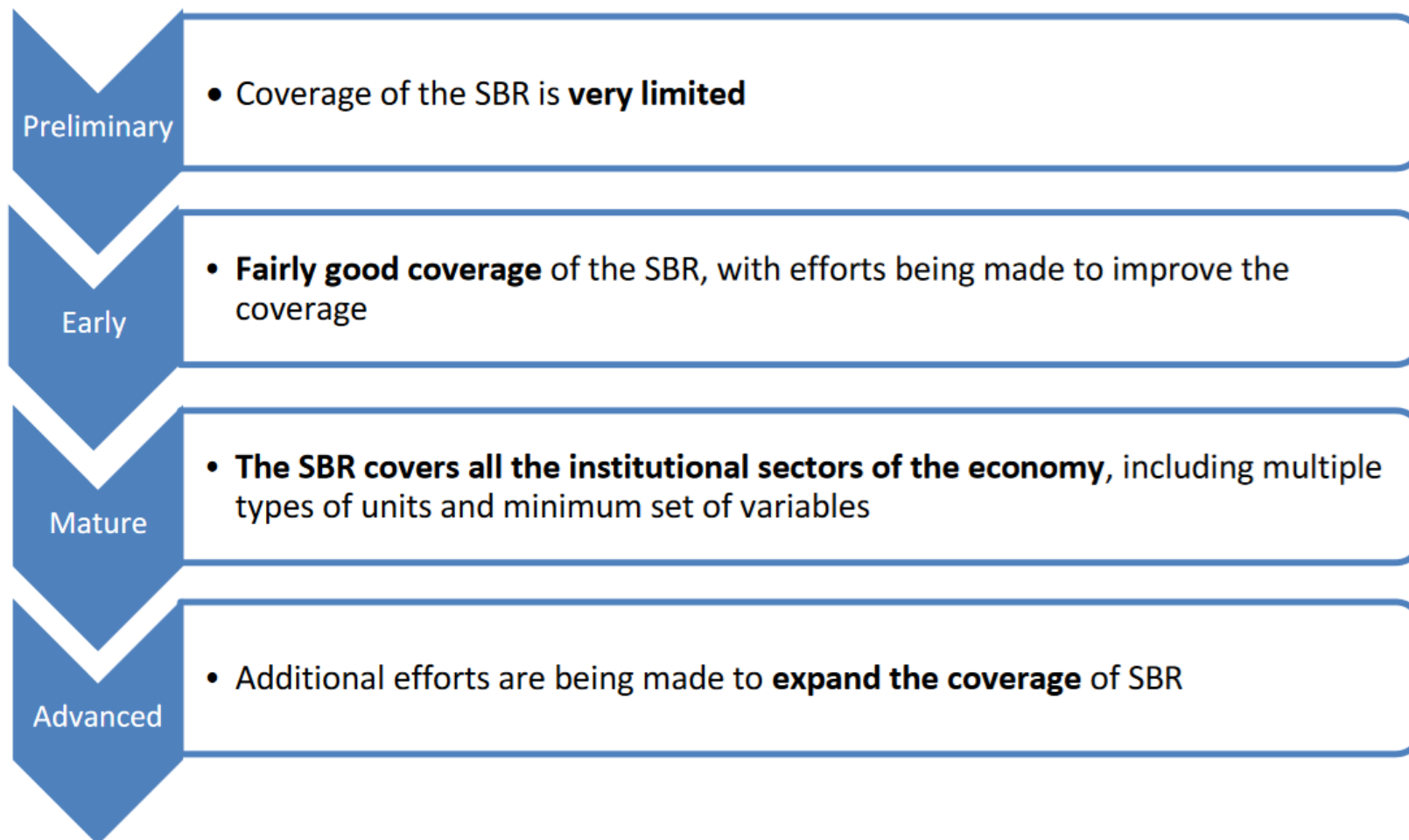
List of variables that should be included in the SBR using the following groupings:

- Identification and contact;
- Demographic;
- Economic/stratification;
- Links and external references.

# Coverage of an SBR

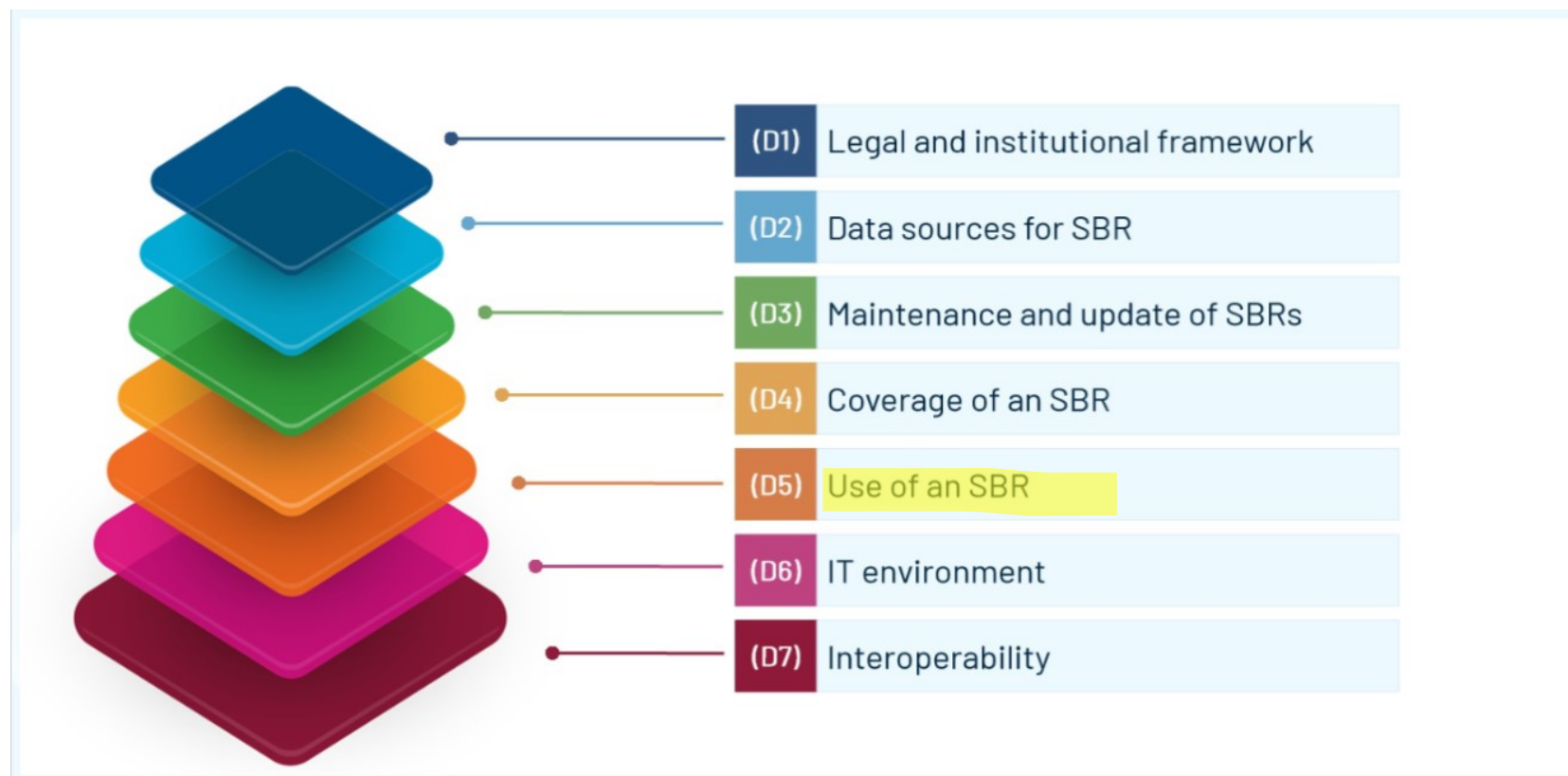
- The coverage of the SBR highly depends on the coverage of the data sources used to update the SBR, and also on the broader legal framework governing the SBR (for example, if it is compulsory for businesses to register and whether the NSO can access administrative data or not). **This dimension, therefore, is very interlinked with other dimensions of the maturity model.**

## Summary of stages of maturity for Dimension 4: Coverage of an SBR





# Seven dimensions



## Dimension 5: Use of an SBR

- This dimension of the maturity model refers to the use of the SBR.
- As the SBR develops, it lends itself to a wider range of uses by users internal to the NSO as well as by external users.

# Use of an SBR

## Considerations for use

- Quality
- Confidentiality

# Use of an SBR

## Internal uses

- Survey frames
- Survey estimation
- Survey response and provider burden measure

# Use of an SBR

## External uses

- Business demography
- List releases of business characteristics
- Micro data linking outputs
- Aggregate data
- Anonymised unit record data
- International registers / exchange

# Use of an SBR

- As the SBR matures, its roles expand for example from being solely in support of different surveys to provide the populations of statistical units with links to administrative units at fixed points in time for specific reference periods, enabling the production of consistent and coordinated business statistics.

# Use of an SBR

- As the SBR matures, it is used as the basis for the compilation of business demography and other statistics based directly on the SBR.
- Finally, as the SBR matures, it is used as the basis for micro-data linked, for data exchange and contributes to the development of international regional group registers.

## Summary of stages of maturity for Dimension 5: Use of SBR

Preliminary

- Use of SBR is **almost solely to support siloed business surveys**

Early

- The SBR can only be used to **supplement a stove pipe survey frame process**

Mature

- The SBR serves as a **backbone in the production** of business demography statistics and economic statistics

Advanced

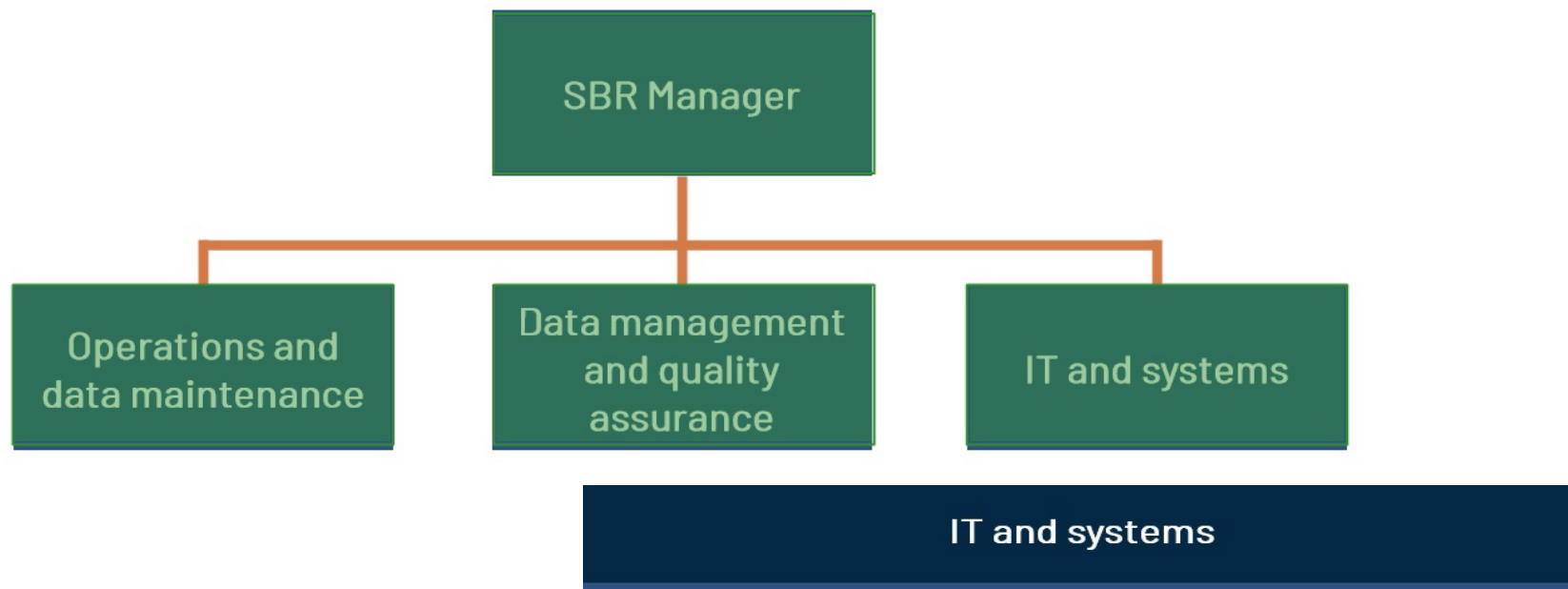
- The SBR is used for **micro-data linking, data exchange** and contributes to the **development of global/regional registers**



# Seven Dimensions of the SBR Maturity Model

- 
- Dimension 6: IT environment
  - Dimension 7: Interoperability

# Dimension 6: IT Environment



- Ensuring that the systems and software needed for extracting data are properly programmed and optimized.
- Maintaining, and possibly advances or further developing, the software that enables SBR staff to easily access and update the SBR content.
- Administering the database processes and tables that make up the SBR, and ensuring that the tables are accessible and available for production of the outputs.

# Dimension 6: IT Environment

It's important to keep in mind several aspects in the development of an IT infrastructure for SBRs so that it is in line with the resources and priorities of the NSO.



Effective partnerships with internal and external stakeholders are crucial.



Implementation of the SBR must focus on achieving its primary role.



Seek to minimize unnecessary conceptual and technical complexities.



# Dimension 6: IT Environment

## Software development methodology

### Agile implementation:

- Demonstrating value to shareholders quickly
- Managing change and risk continuously
- Frequent evaluation and realignment of scope, cost, schedule, quality

### 2 phases of IT development:

- Acquiring initial db infrastructure
- Then programming and process development

Follow organizational IT standards

If one is not already implemented, should adopt an appropriate one

# Dimension 6: IT Environment

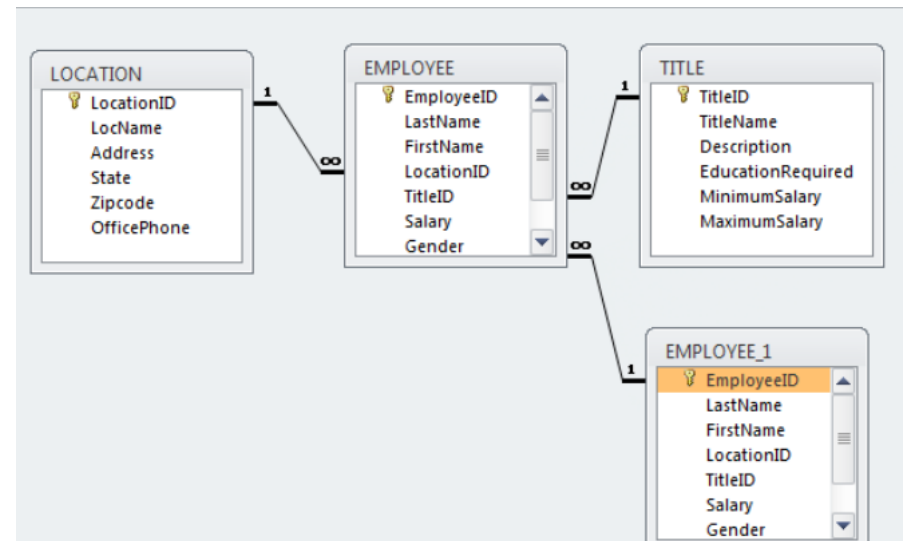
## Database management systems

- Security
- Scalability
- Concurrent users
- Flexible and standardized query language

## Relational database management systems (RDBMS)

Additionally:

- Referential and data integrity
- Efficient storage and performance
- Modularity



# Dimension 6: IT Environment

## Database options

- Low cost: Microsoft Access (no concurrent users, limited capacity)
- Others: Oracle, Microsoft SQL Server, MySQL, MariaDB and PostgreSQL

## Technologies should be:

- Accessible, staff able to be trained
- Long shelf life
- Able to integrate external systems, administrative data especially, through APIs
- Use clear definitions and concepts

# Dimension 6: IT Environment

## Other elements

- Should be a place (within or external to db) for storing and accessing data on: frames, snapshots, reporting burden, etc.
- In primary role of providing survey frames, db should include data on collection process: units sampled, how and where data about units will be collected and results of collection efforts -> inputs into “respondent burden module” for use across agency/statistical office, including efforts to mitigate that burden



# Dimension 6: IT Environment

## Other elements

### Unique Identifiers

Essential, will discuss more in the session on the Global Initiative on Unique IDs

### Data retention policy

- Don't just change the information in the system
- Instead add information about the change, including the time when change occurred (or was implemented), and maintain the old value
- Can help in reconstructing the register and maintaining the historical register

### Documentation

User manual/wiki/technical documentation

# Dimension 6: IT Environment

## ADB and Statistics Norway-developed systems

Adaptable, NSOs maintain ownership of the system and code

### Statistics Norway: **STATBUS** [Statbus.org](http://Statbus.org) + [github.com/statisticsnorway/SBR](https://github.com/statisticsnorway/SBR)

- Runs on Microsoft's IIS (Internet Information Server) and is written in ASP.NET
  - Choice of underlying database – Microsoft SQL Server, PostgreSQL, MySQL
  - Data can be uploaded in CSV or XML format
- 
- ✓ Includes data quality checks, system user management
  - ✓ Different languages, different classification systems
  - ✓ Define, construct, maintain statistical units
  - ✓ Classify units by sector/activity
  - ✓ Produce survey frames and register-based statistics
  - ✓ Upload from different sources
  - ✓ Multiple ID fields including computer generated; up to 3 languages

# Dimension 6: IT Environment

## ADB and Statistics Norway-developed systems

### ADB SBR System

Many of the same features as STATBUS:

- Web-based application
- User accounts
- Create frames/snapshots, historical data
- API integration

### Live Register + Data Integration



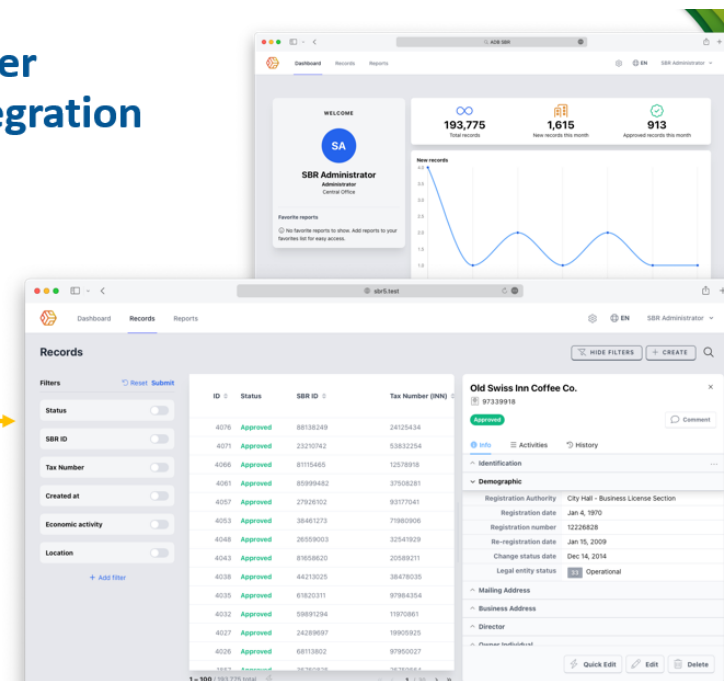
Direct input  
(encoding, profiling)



Flat file import  
(LSX, CSV, XML, etc)



External systems  
via APIs



### These systems are not “plug-and-play”!

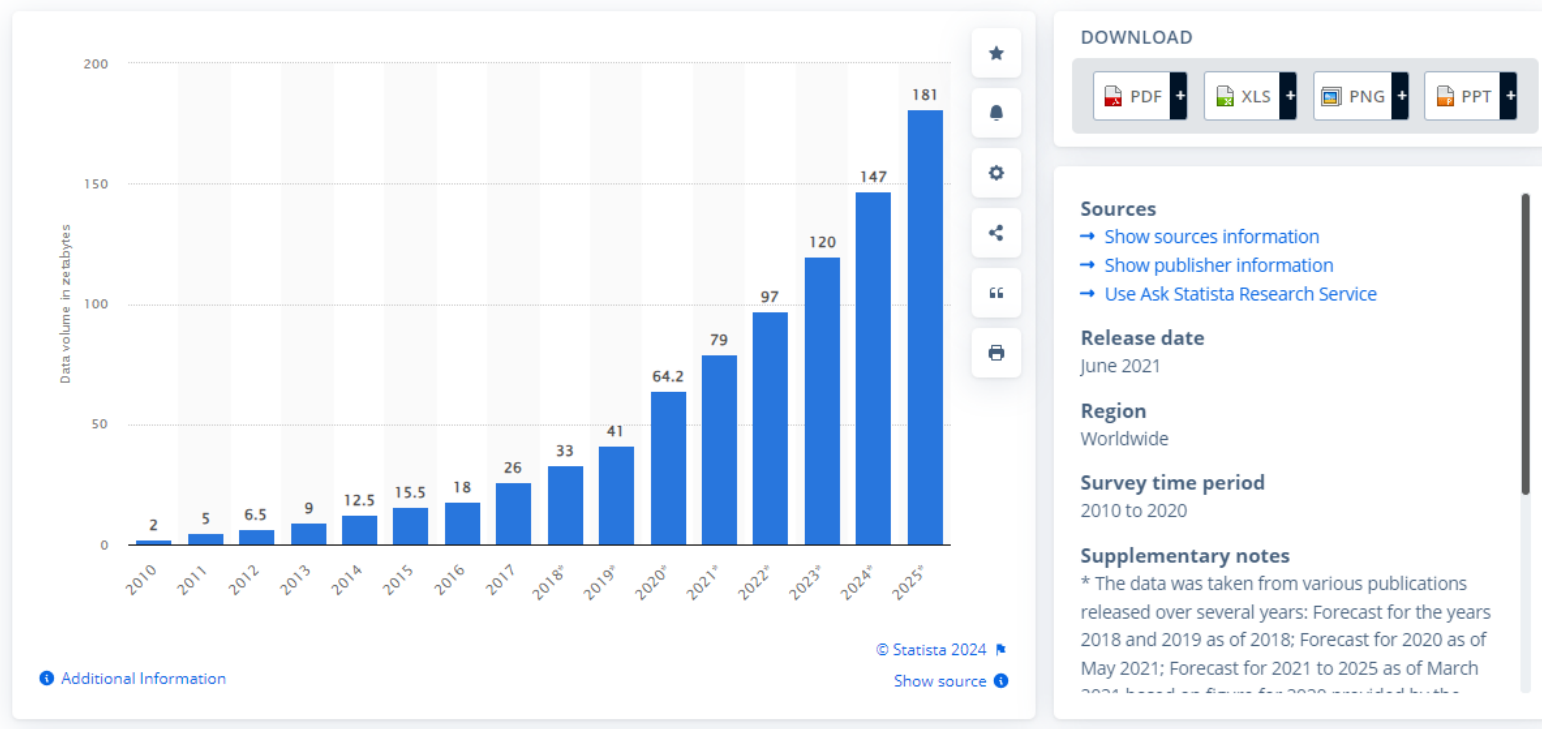
- Require collaboration with these organizations
- Internal planning and management and maintenance are still needed

# Dimension 6: IT Environment

## Big Data, Artificial intelligence (AI) and Machine Learning (ML)

Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025

(in zettabytes)



# Dimension 6: IT Environment

## Big Data, Artificial intelligence (AI) and Machine Learning (ML)

- Many web scraping tools available online
- Companies providing data -> now also providing web scraping services
- Can be used to help identify web stores/online retail
- AI can help write code for scraping, etc.
- Language recognition is a major component

# Dimension 6: IT Environment

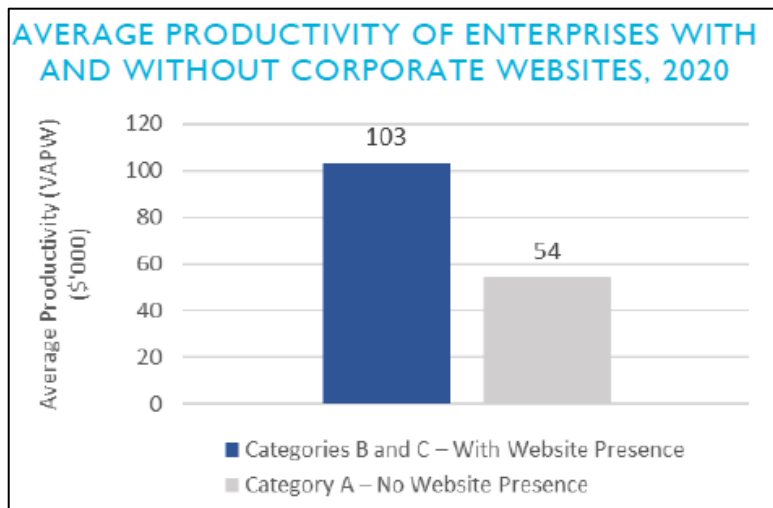
## Big Data, Artificial intelligence (AI) and Machine Learning (ML)

### Singapore example

PERFORMANCE OF CLASSIFIERS EXPLORED	
Classifier	Test Set Accuracy
Random Forest	79%
Gradient Boosting Machine	77%
Voting Classifier	77%
Logistic Regression	72%
Neural Network	71%
AdaBoost	70%
Support Vector Machine	68%
Naïve Bayes	57%

<https://www.singstat.gov.sg/-/media/files/publications/economy/ssn222-pg1-4.ashx>

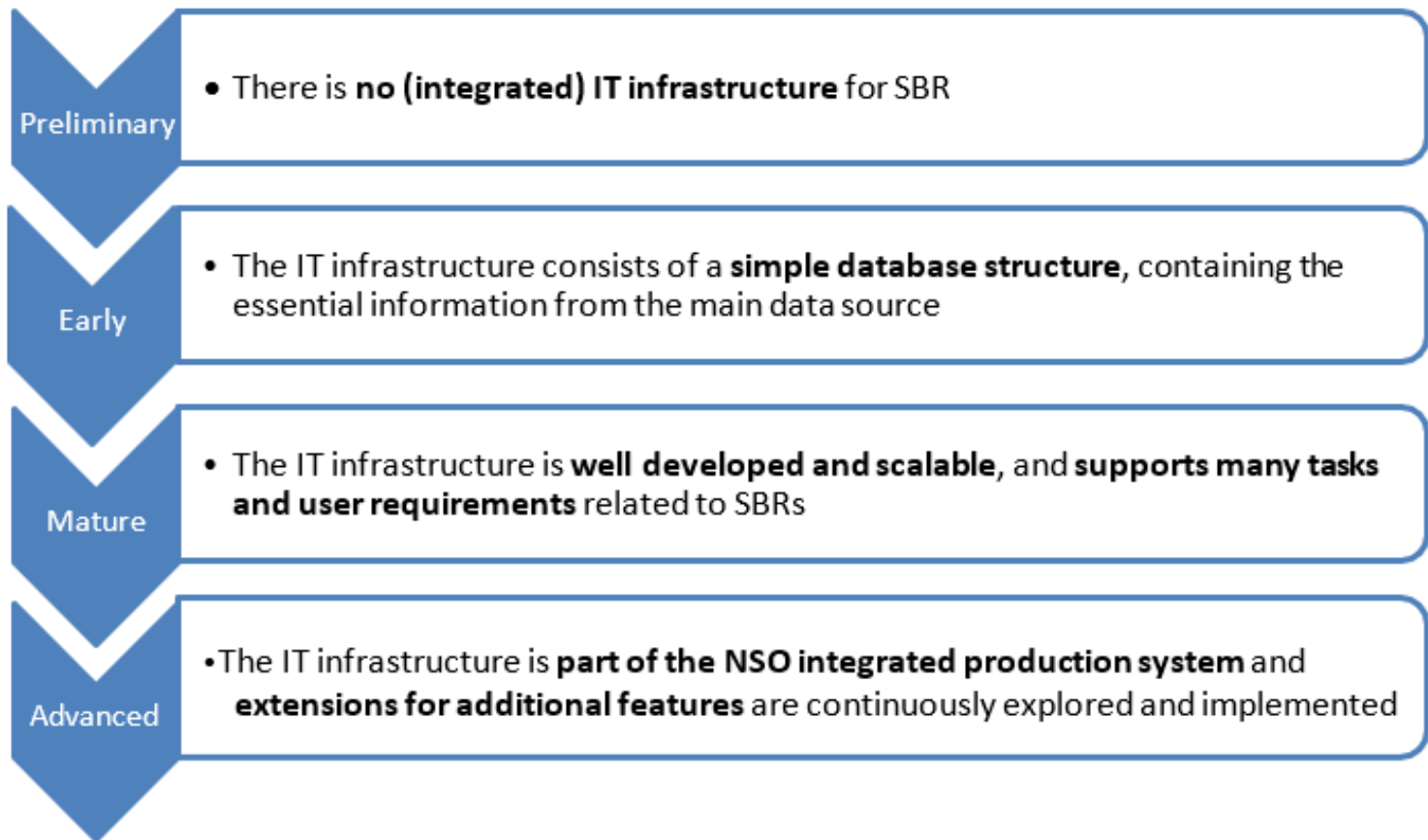
FEATURE IMPORTANCE OF SELECTED WORDS	
Feature Word	Feature Importance
Shop	0.044
Cart	0.041
Price	0.027
Facebook	0.021



# Dimension 6: IT Environment

**Figure 8**

**Summary of stages of maturity for Dimension 6: IT environment**



# Dimension 6: IT Environment

## Preliminary →

- No integrated infrastructure
- Records kept manually, i.e. Excel
- Basic maintenance strategy

## Early →

- Simple db structure with essential data
- May have other applications, but not integrated

## Mature →

- Managed like a project, maintained, dedicated staff
- Modular, scalable, supports historical register, unique IDs

## Advanced

- Part of integrated system, continuously improved
- New technologies: big data, web scraping, portals for businesses to enter info



# Dimension 6: IT Environment

## Questions:

1. Does your SBR use a relational database?
2. Have you had positive/negative experience with specific types of software?
3. Have you had any experience with ADB/STATBUS products?

# Dimension 7: Interoperability

## Interoperability

The ability of computer systems or software to exchange and make use of information in a coordinated manner

## In SBRs

The ability of an SBR to communicate and exchange standardized data with other registers, be they domestic, regional, or global

Data can be easily re-used and processed in different applications, allowing different information systems to work together

# Dimension 7: Interoperability

## Interoperability

- Related to IT infrastructure but important enough to be separately assessed
- Increases efficiency

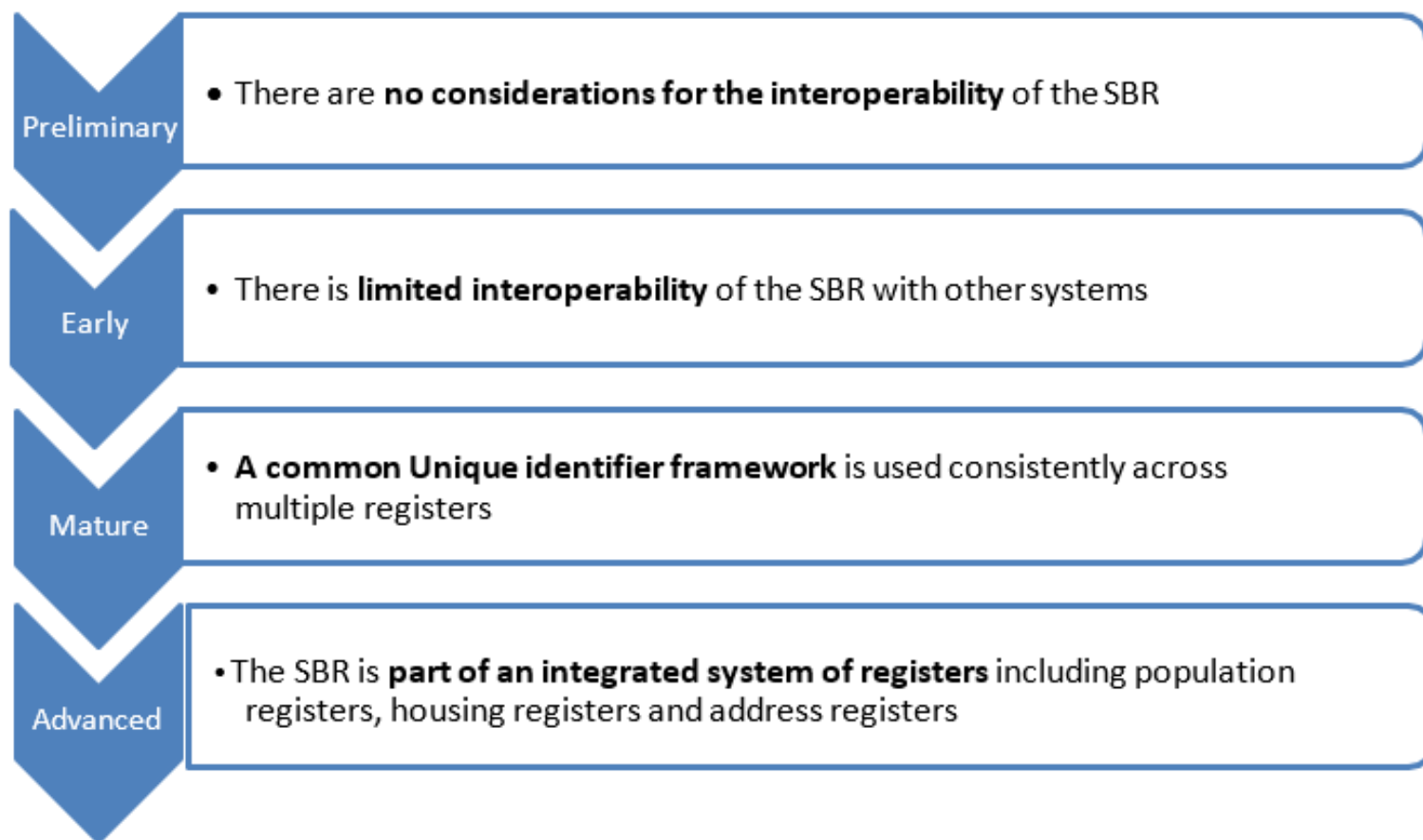
## Requires

- A unique ID
- Standardized language/classification/characteristics

# Dimension 7: Interoperability

Figure 9

Summary of stages of maturity for Dimension 7: Interoperability



# Dimension 7: Interoperability

## Preliminary →

- Matching done manually
- No considerations for interoperability

## Early →

- Some automated routines
- Some interoperability with administrative sources

## Mature →

- Common unique ID to link administrative & statistical registers
- Compatibility with international standards

## Advanced

- Microdata linking implemented
- Interoperable with international sources using global identifiers
- Integration with population, housing, other registers

# Dimension 7: Interoperability

## Questions:

1. Does your SBR meet the minimum requirements, namely:  
Use of a unique identifier +  
A standardized language/classification/characteristics?
2. Which, if any, other registers can be linked with your SBR? Could this be done in a more efficient way?
3. Has your office implemented micro data linking?

Thank you!



Shared Prosperity Dignified Life

