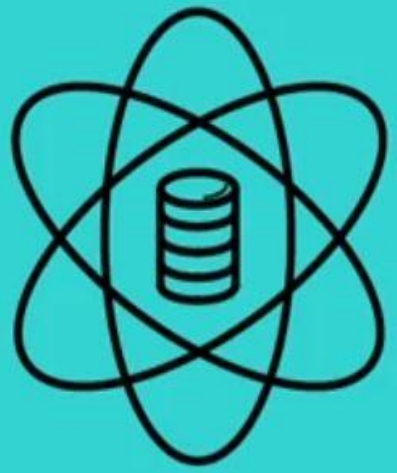
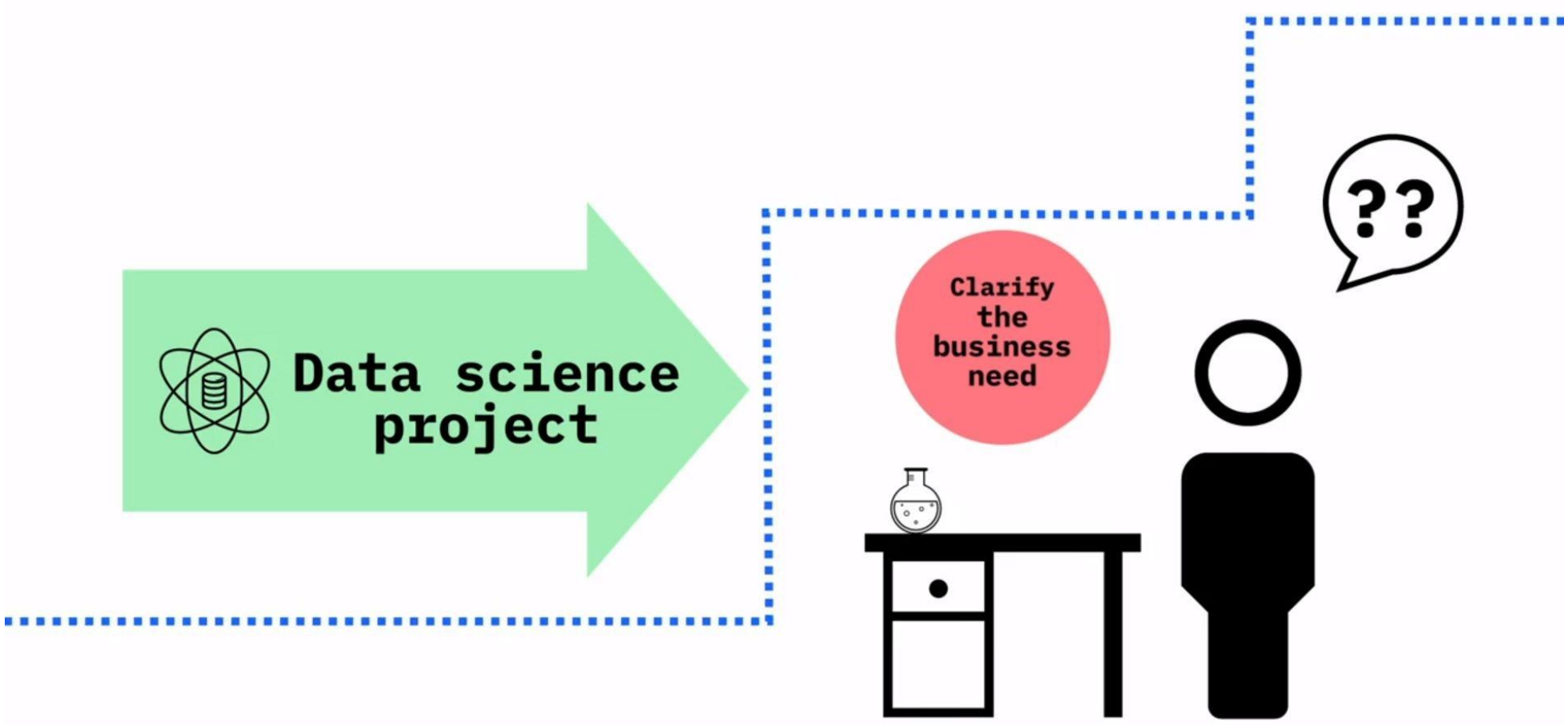
Data science and data mining

ارائه دهندگان:
نیکی نجفی-960164039
فاطمه مهدوی گلوجه-960207862
درس آزمایشگاه پایگاه داده ها
استاد راهنما:
استاد الهام رشید مهرآبادی

Data Science
can help organizations:

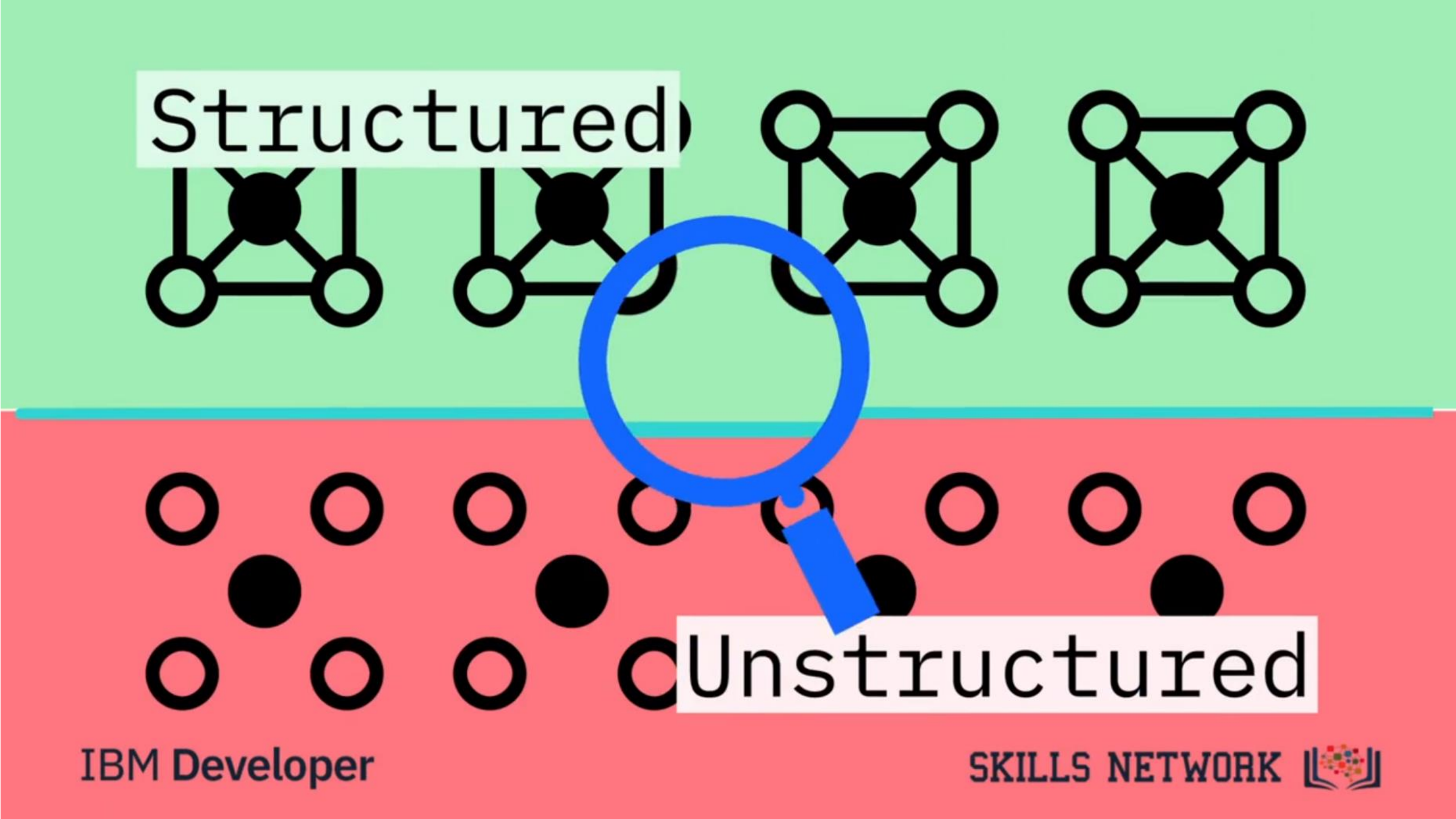


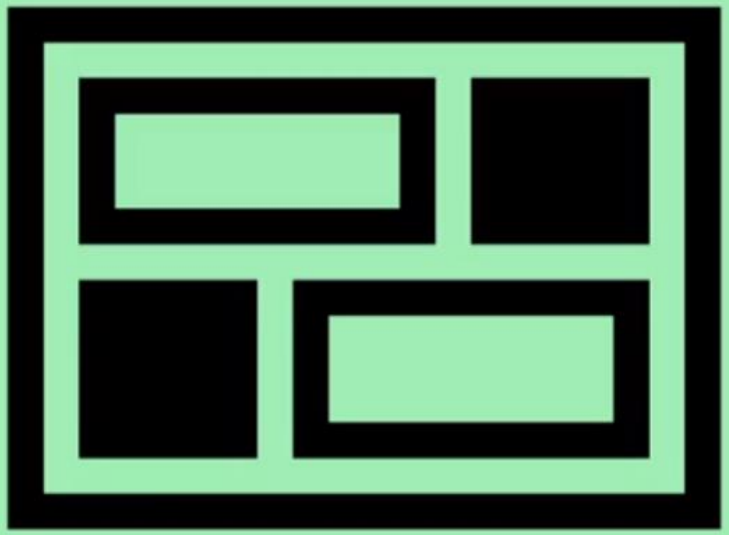
- Understand their environments
- Analyze existing issues
- Reveal previously hidden opportunities

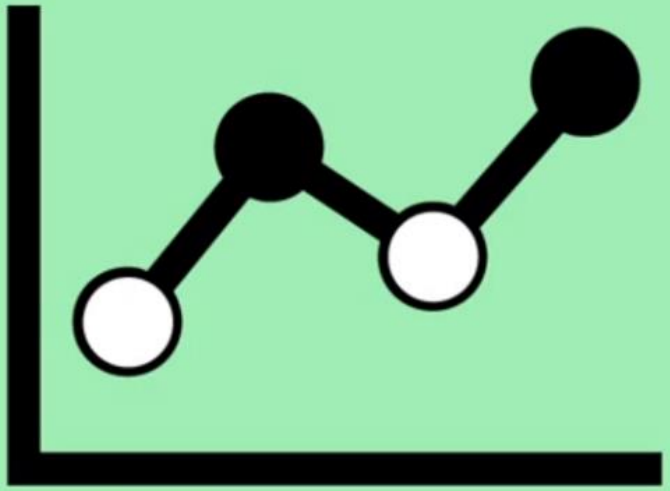

Data science project
Clarify the business need
??

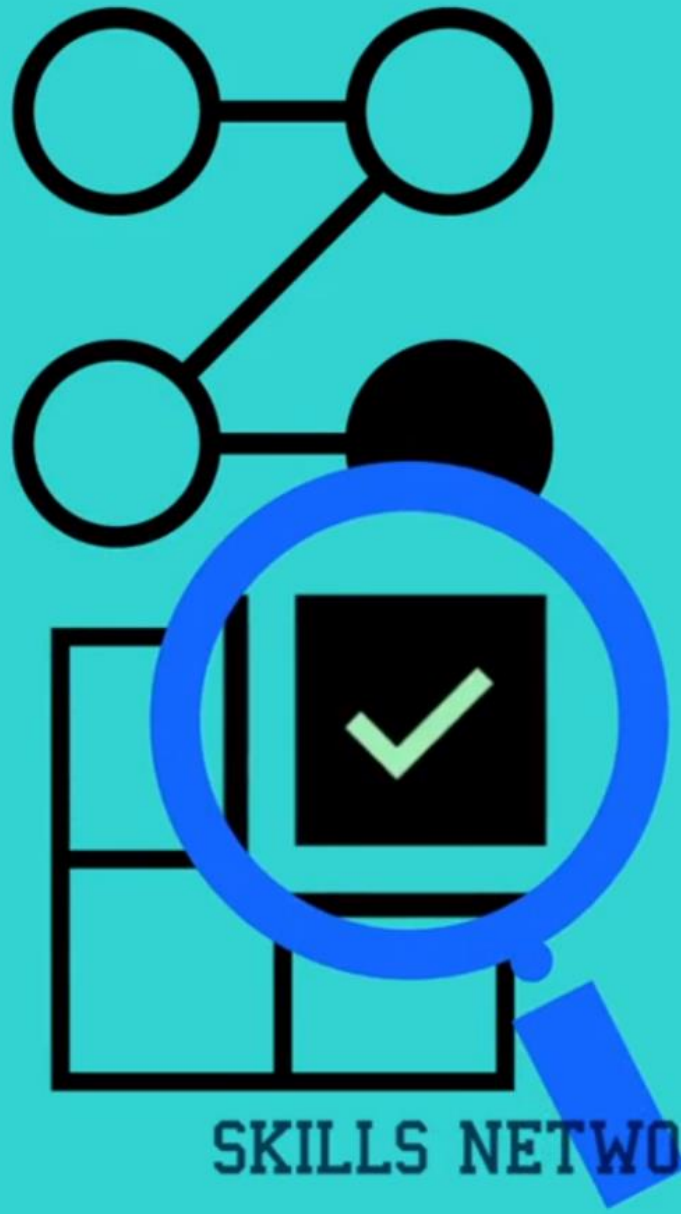
What data do we need to solve the problem?

Where will that data come from?


Structured
Unstructured
IBM Developer
SKILLS NETWORK

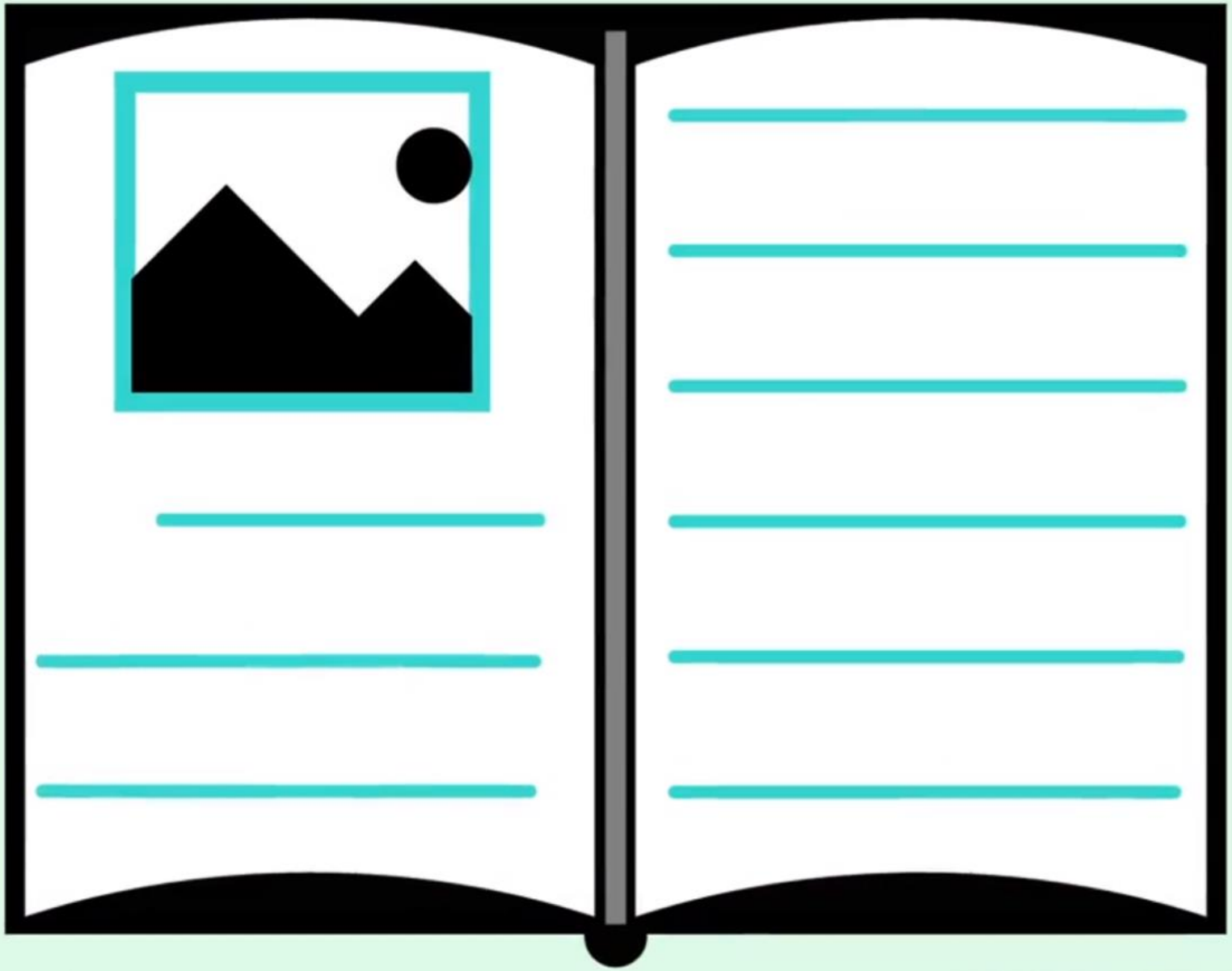






IBM Developer

SKILLS NETWORK



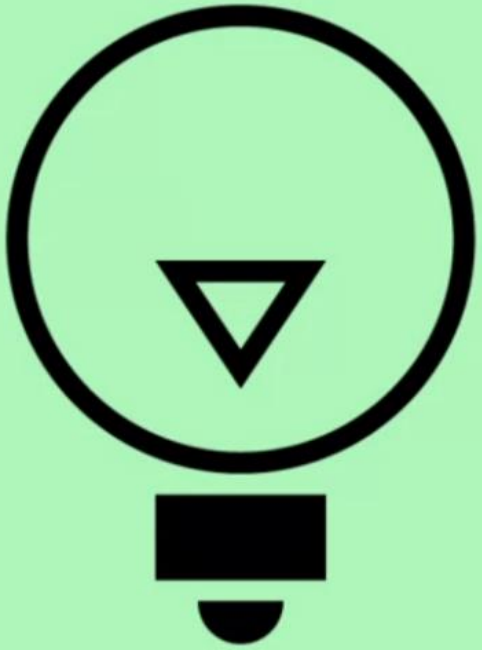
Powerful data visualization tools

Understand the nature of the results

Recommended action to take

Data Science is changing:

- The way we work
- The way we use data
- Our approach to the world

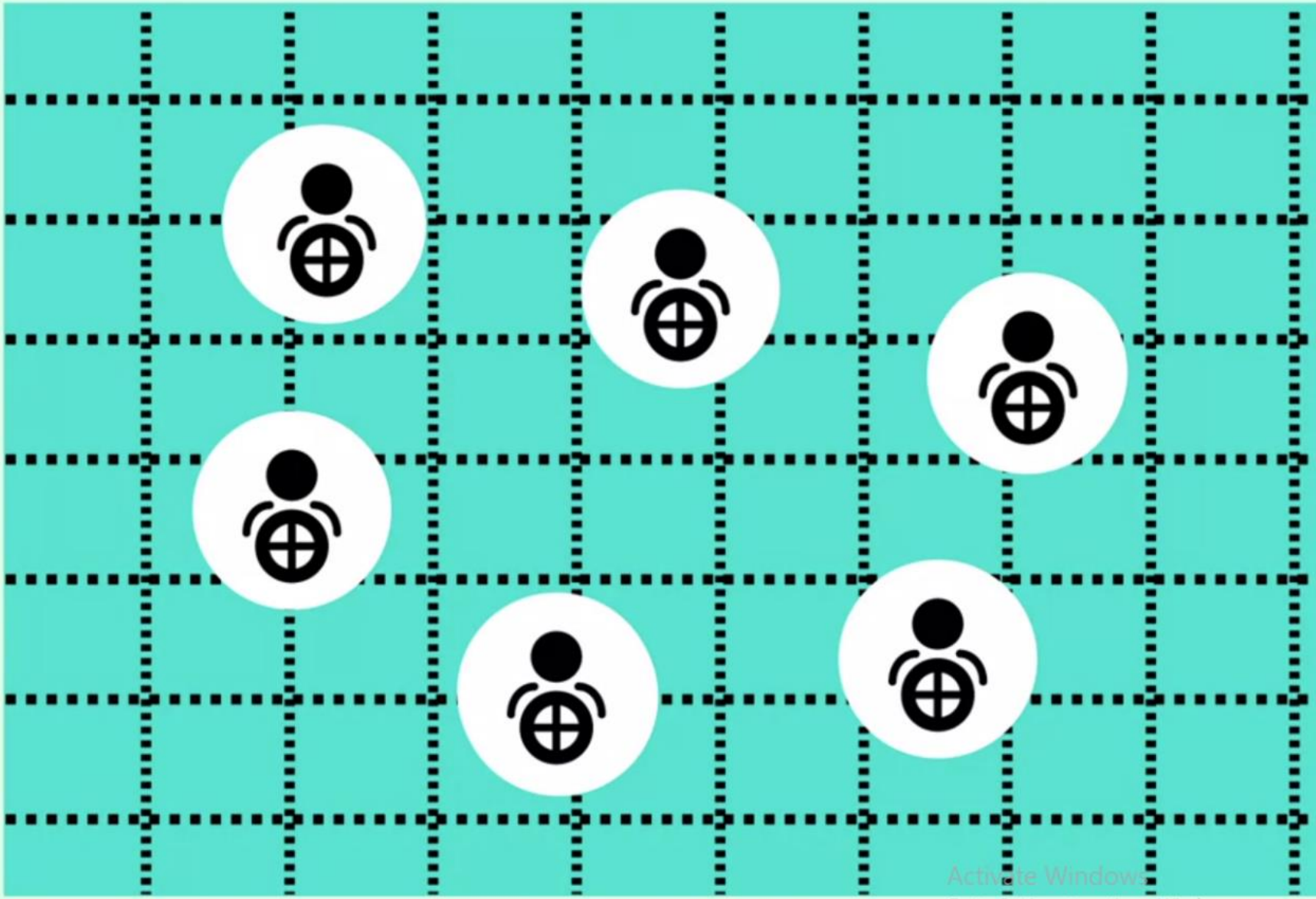


IBM Developer

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.

Uber

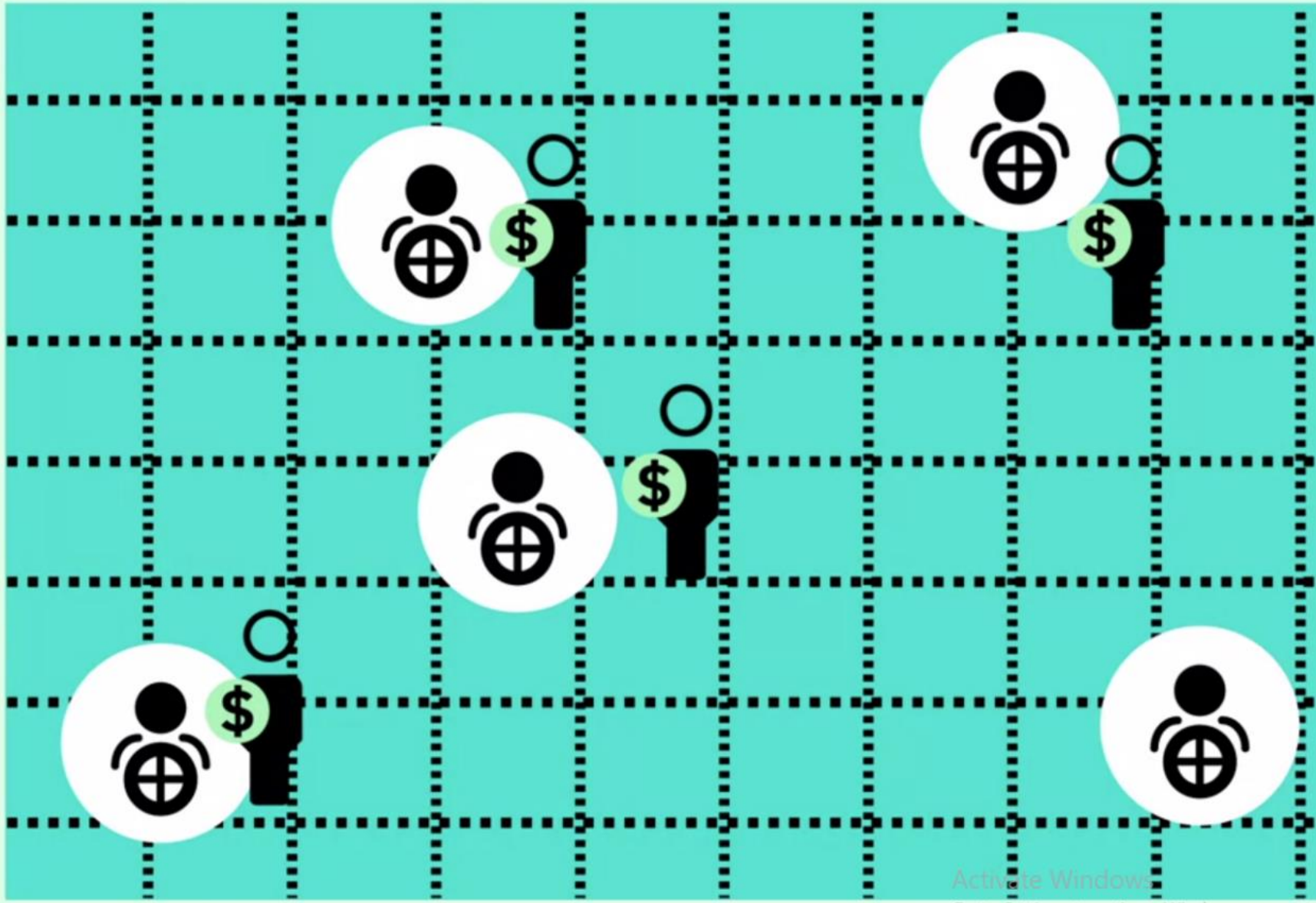


Activate Windows
Go to Settings to activate Windows.

IBM Developer

SKILLS NETWORK

Uber



Activate Windows
Go to Settings to activate Windows.

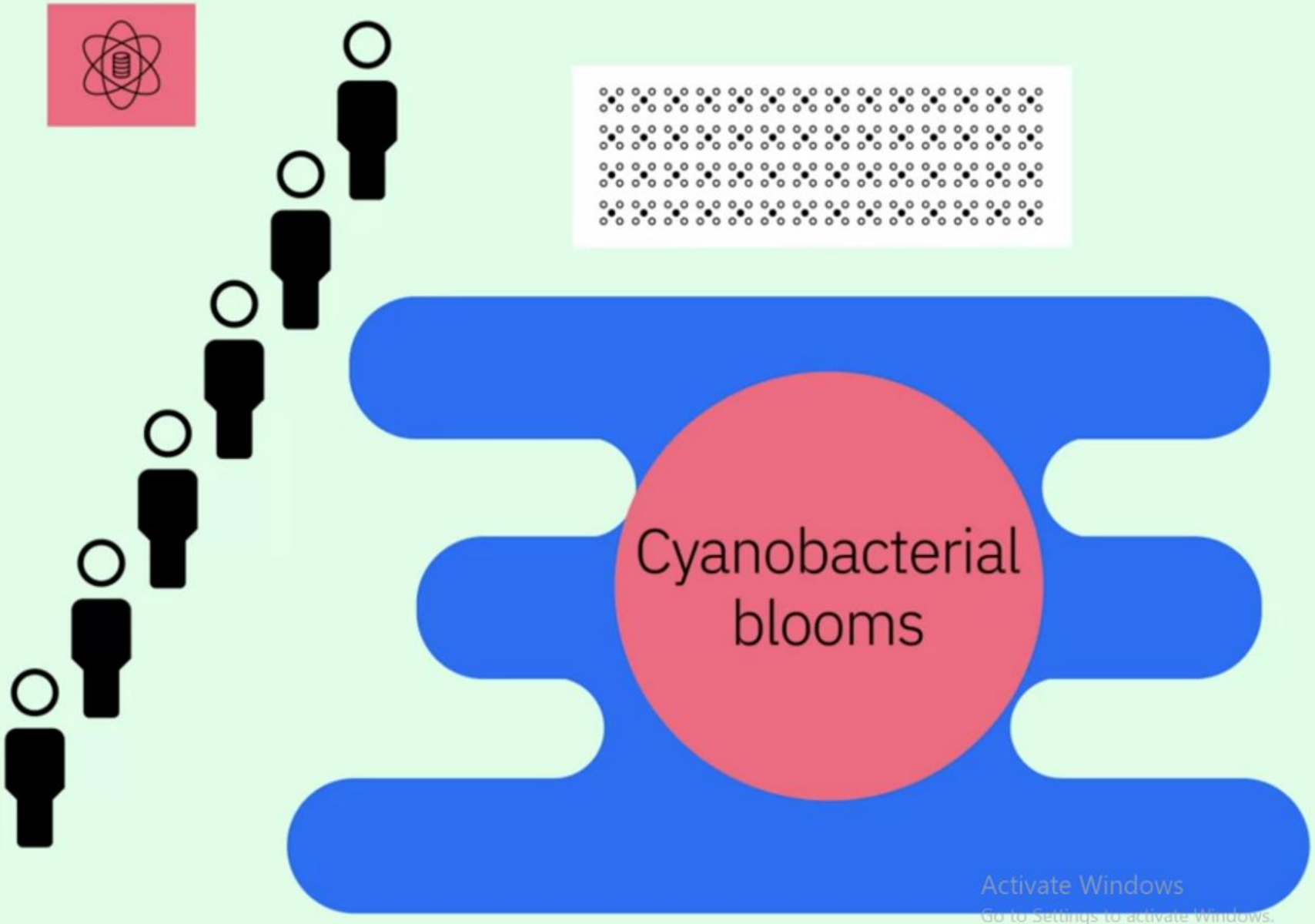
Cyanobacterial blooms

Activate Windows
Go to Settings to activate Windows.

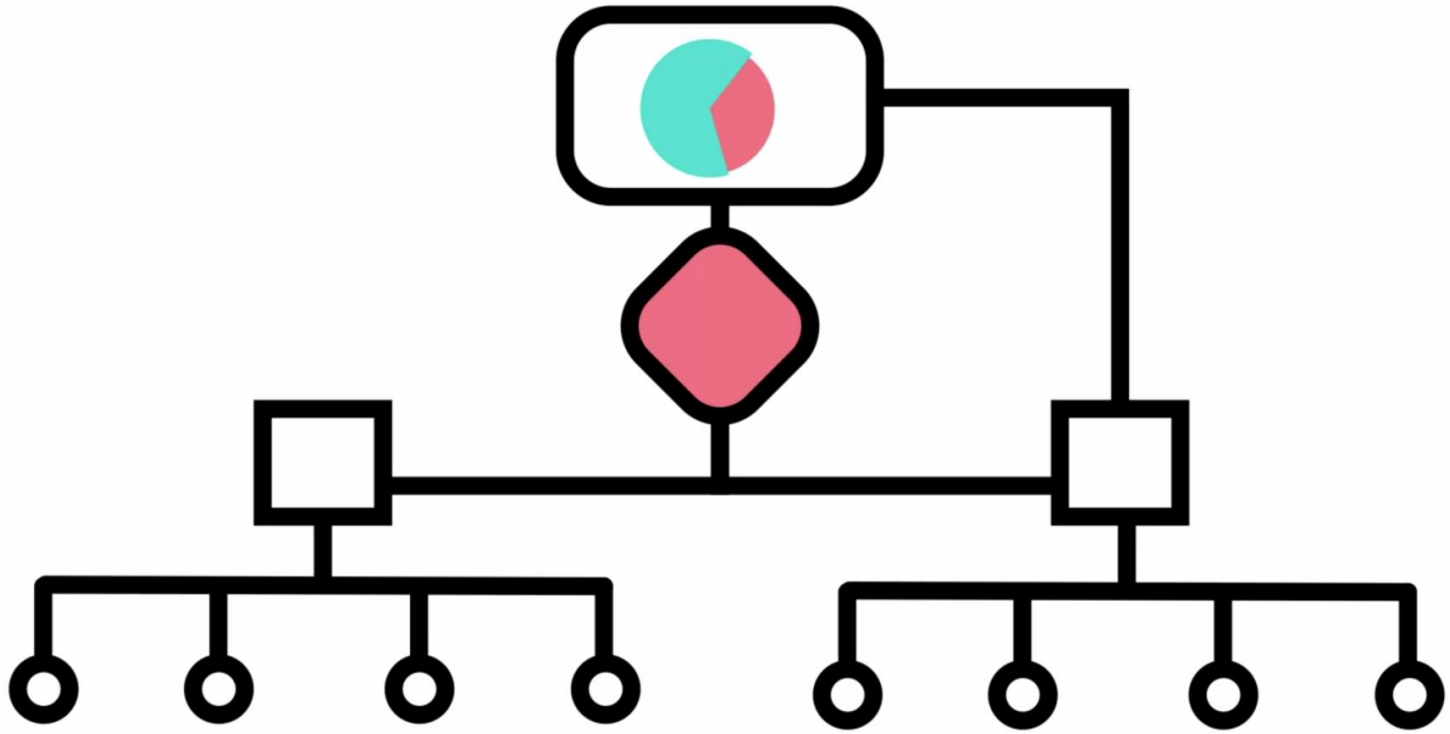
Physical

Chemical

Biological



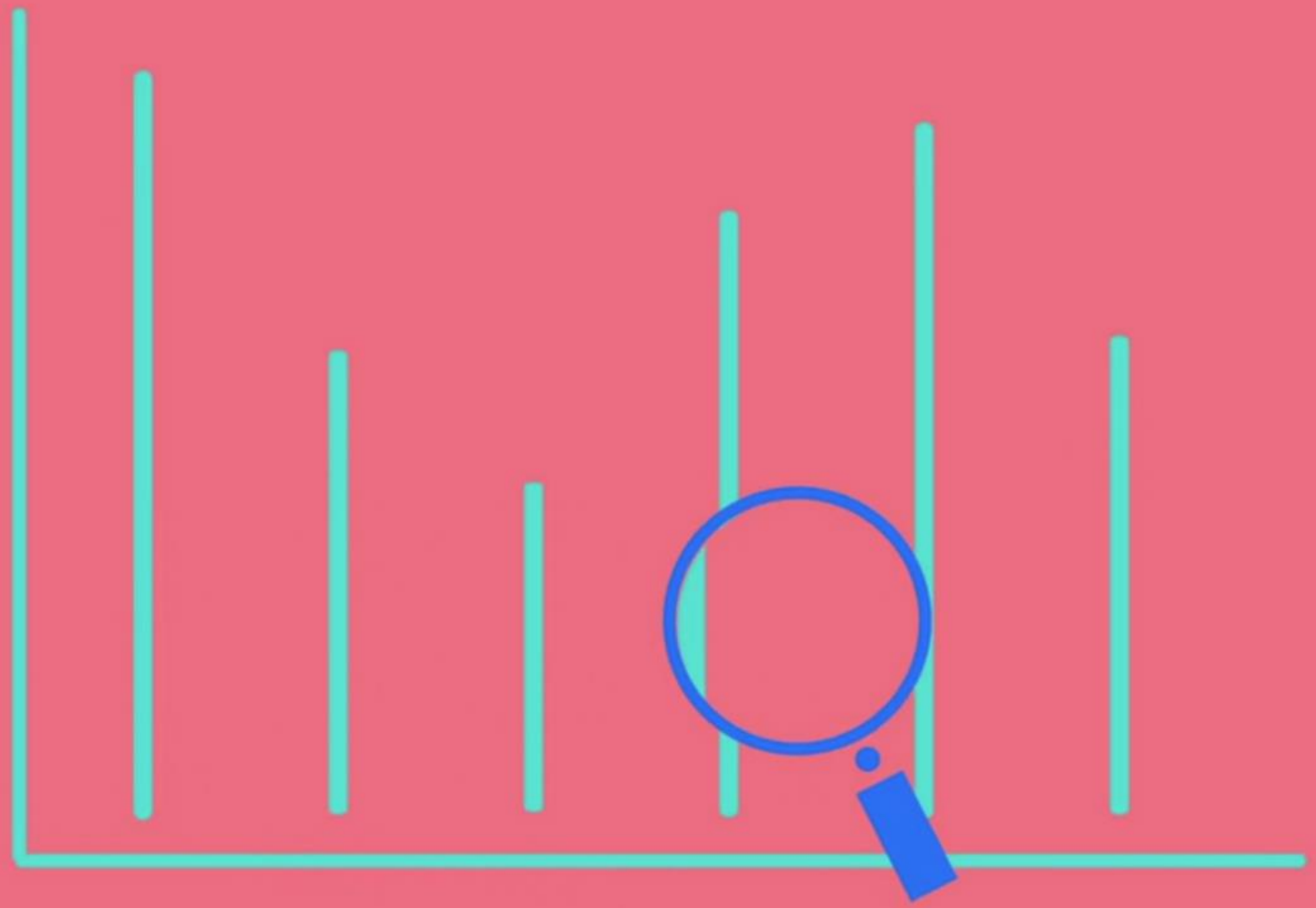
Activate Windows
Go to Settings to activate Windows.



IBM Developer

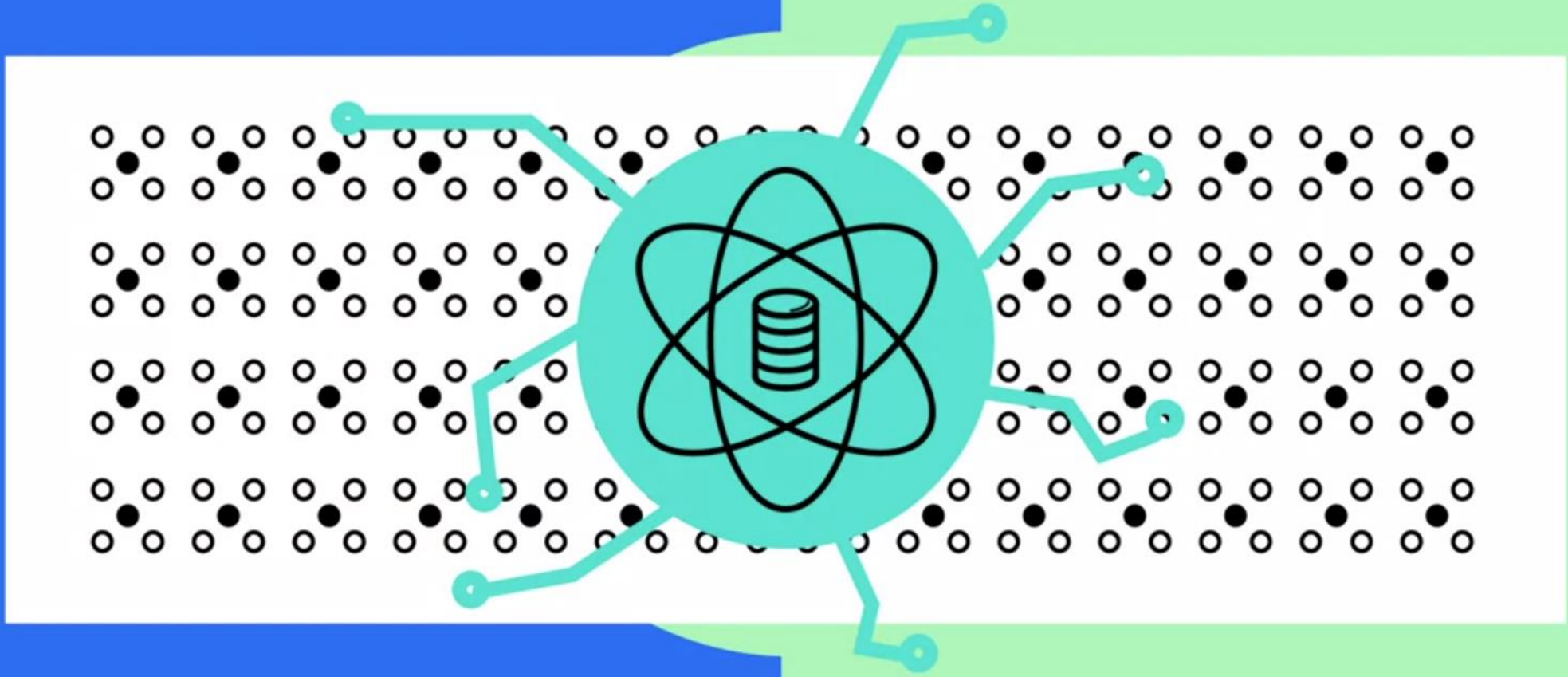
SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.



Activate Windows
Go to Settings to activate Windows.

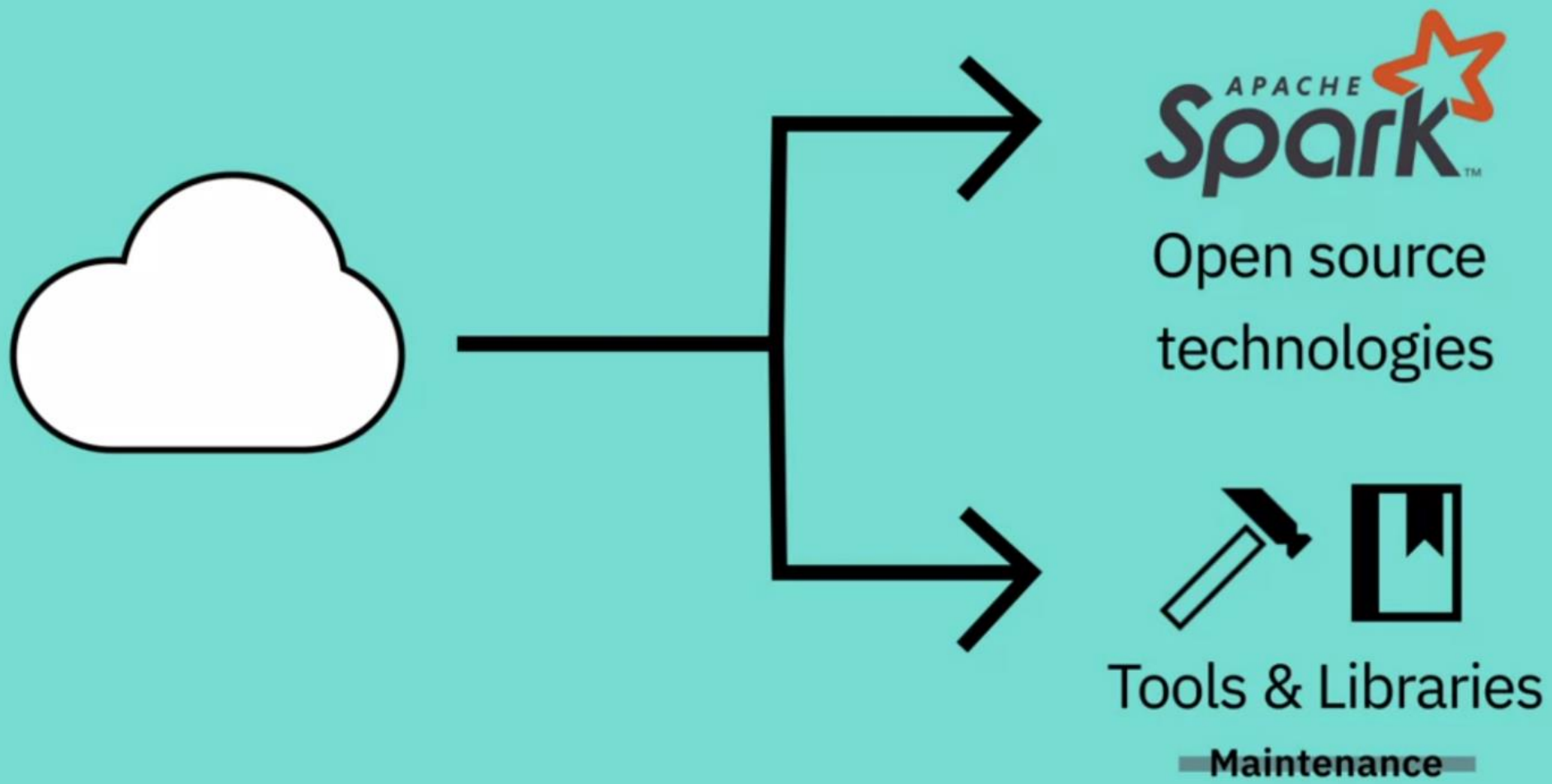
Activate Windows
Go to Settings to activate Windows.



IBM Developer

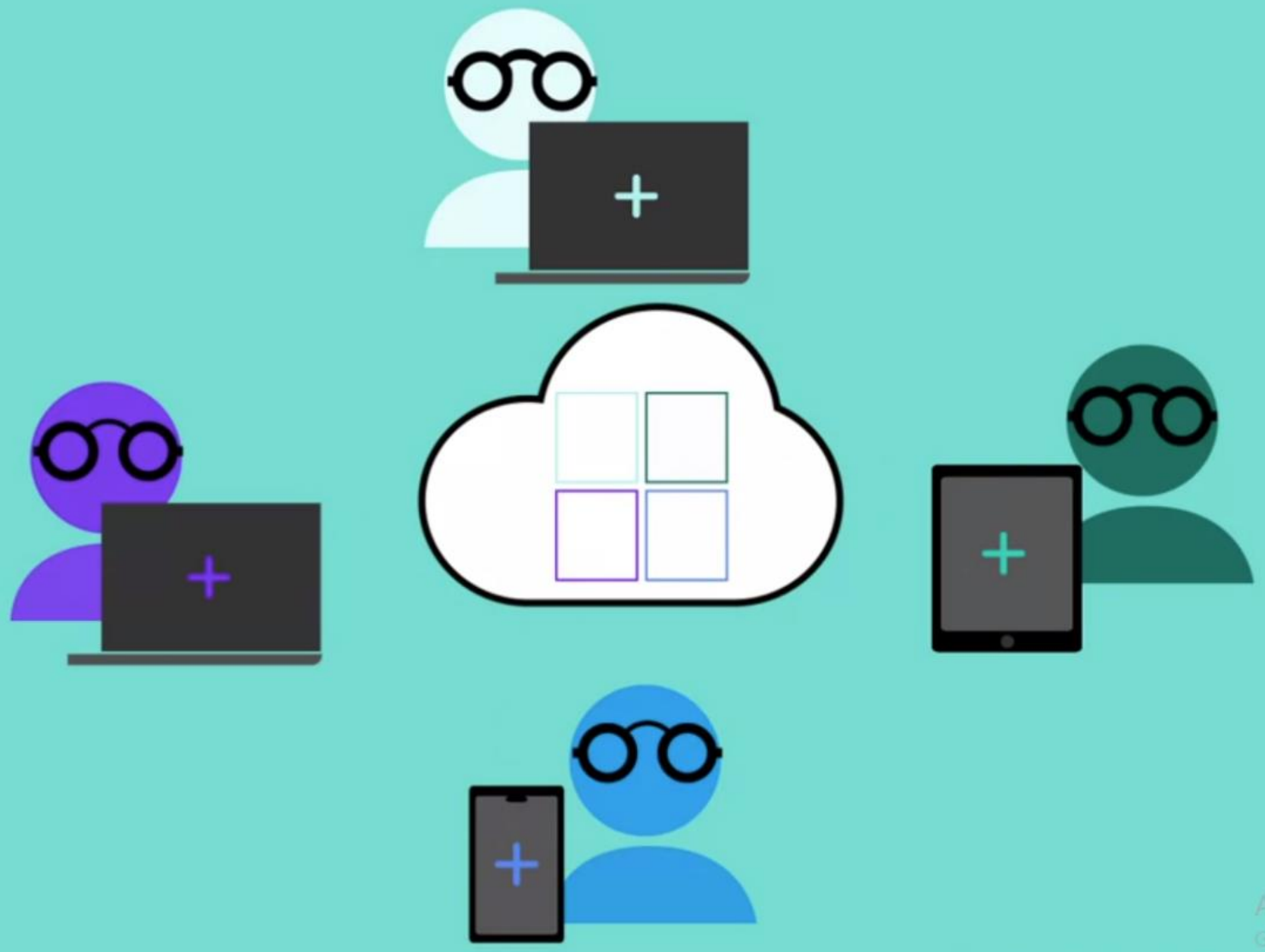
SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.



Activate Windows
Go to Settings to activate Windows.

Activate Windows
Go to Settings to activate Windows.



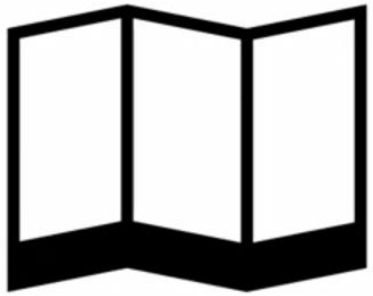
Activate Windows
Go to Settings to activate Windows.

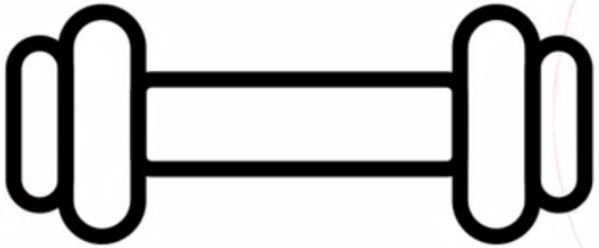
Some big tech companies offer cloud platforms

Become familiar in a pre-built environment

Activate Windows
Go to Settings to activate Windows.

Foundations of Big Data





Activate Windows
Go to Settings to activate Windows.

IBM **Developer**

SKILLS NETWORK

Foundations of Big Data

Activate Windows
Go to Settings to activate Windows.

Big Data

"Big Data refers to the dynamic, large and disparate volumes of data being created by people, tools, and machines. It requires new, innovative, and scalable technology to collect, host, and analytically process the vast amount of data gathered in order to derive real-time business insights that relate to consumers, risk, profit, performance, productivity management, and enhanced shareholder value."

Ernst and Young

Activate Windows
Go to Settings to activate Windows.

The V's of Big Data

Velocity

Volume

Variety

Veracity

Value

Activate Windows
Go to Settings to activate Windows.

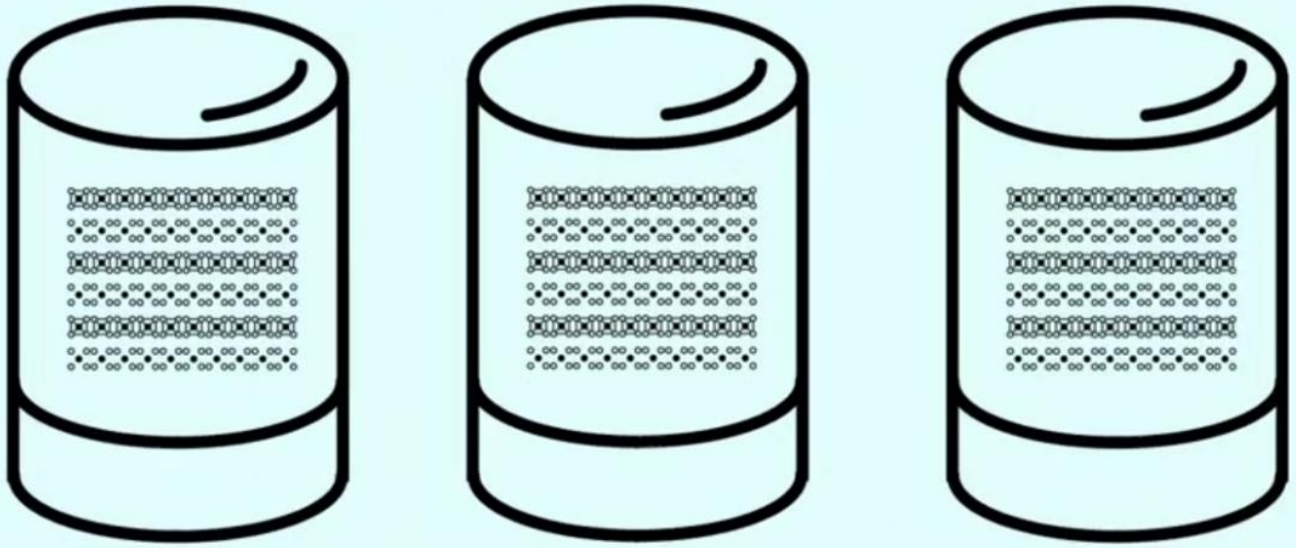
Velocity

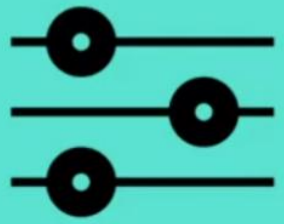
IBM Developer

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.

Volume





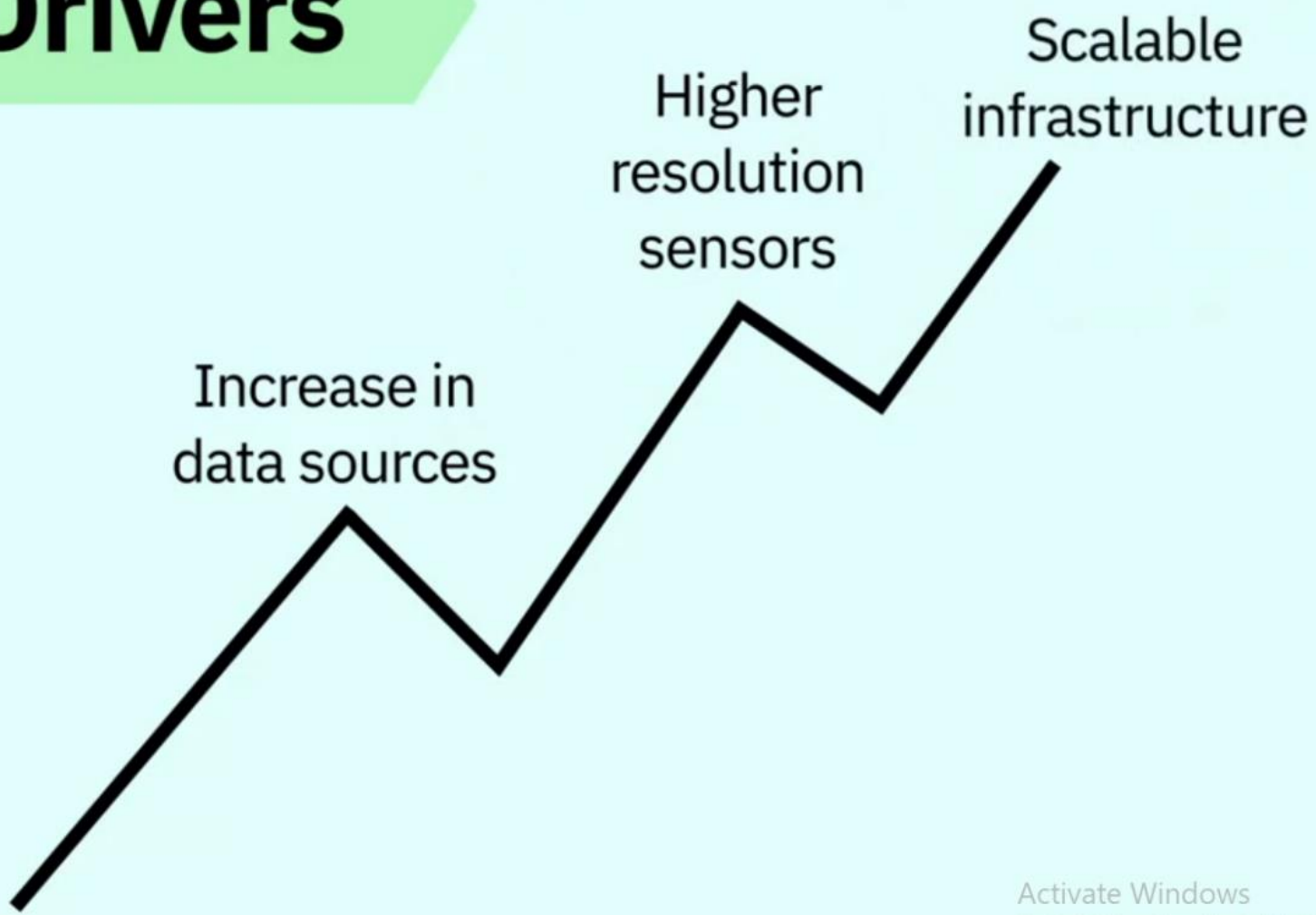
IBM Developer

Activate Windows
Go to Settings to activate Windows.

SKILLS NETWORK

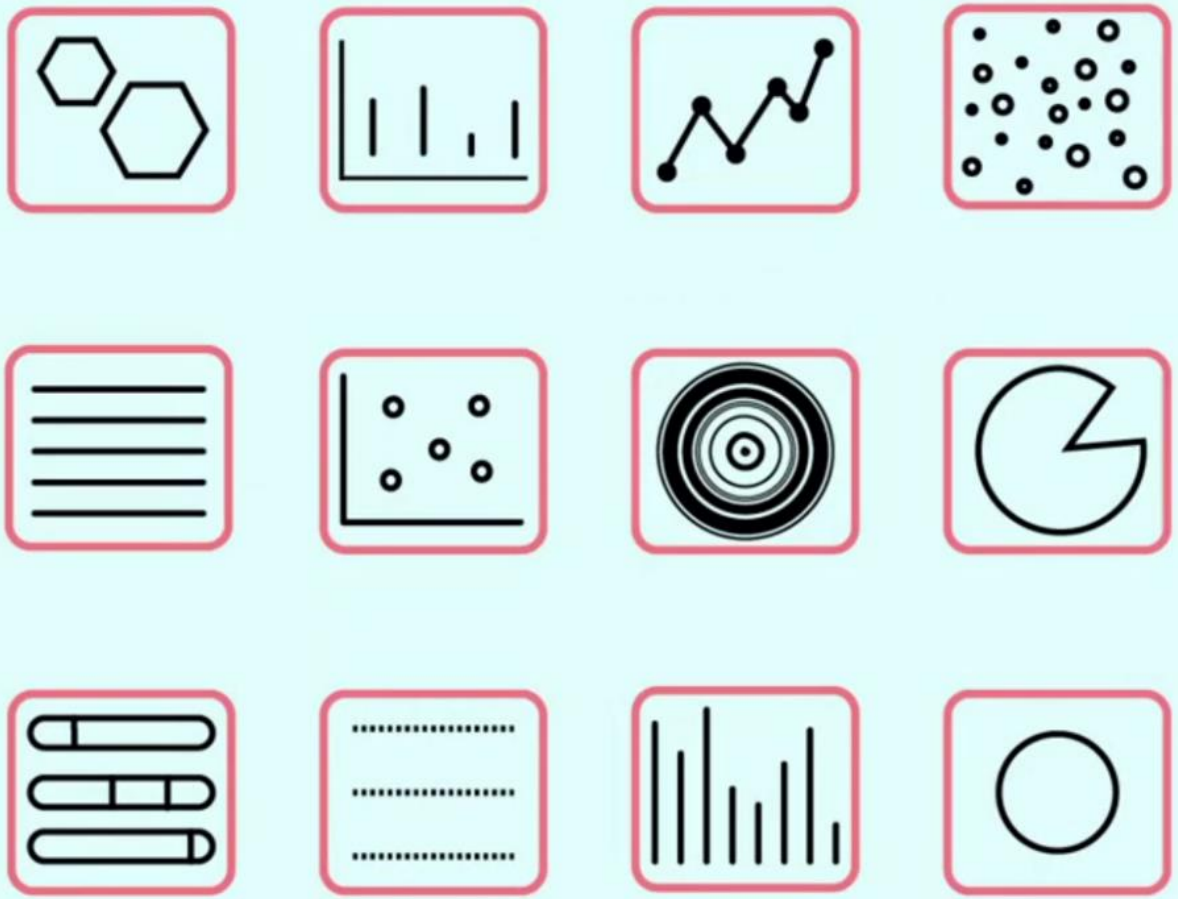
Volume

Drivers



Activate Windows
Go to Settings to activate Windows.

Variety



Activate Windows
Go to Settings to activate Windows.

IBM Developer

SKILLS NETWORK

Variety

IBM Developer

Activate Windows
Go to Settings to activate Windows.

SKILLS NETWORK

Variety

Drivers

- Mobile technologies
- Social media
- Wearable technologies
- Geo technologies
- Video
- Many more

Activate Windows
Go to Settings to activate Windows.

Veracity

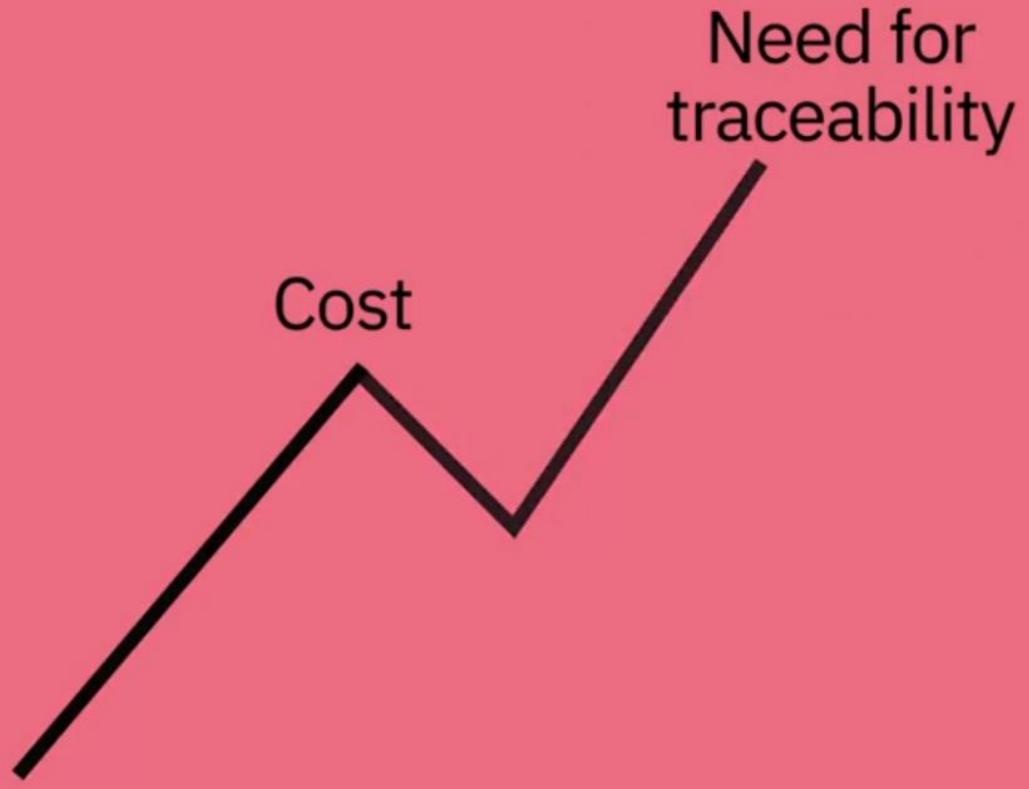
Attributes

- Consistency
- Completeness
- Integrity
- Ambiguity

Activate Windows
Go to Settings to activate Windows.

Veracity

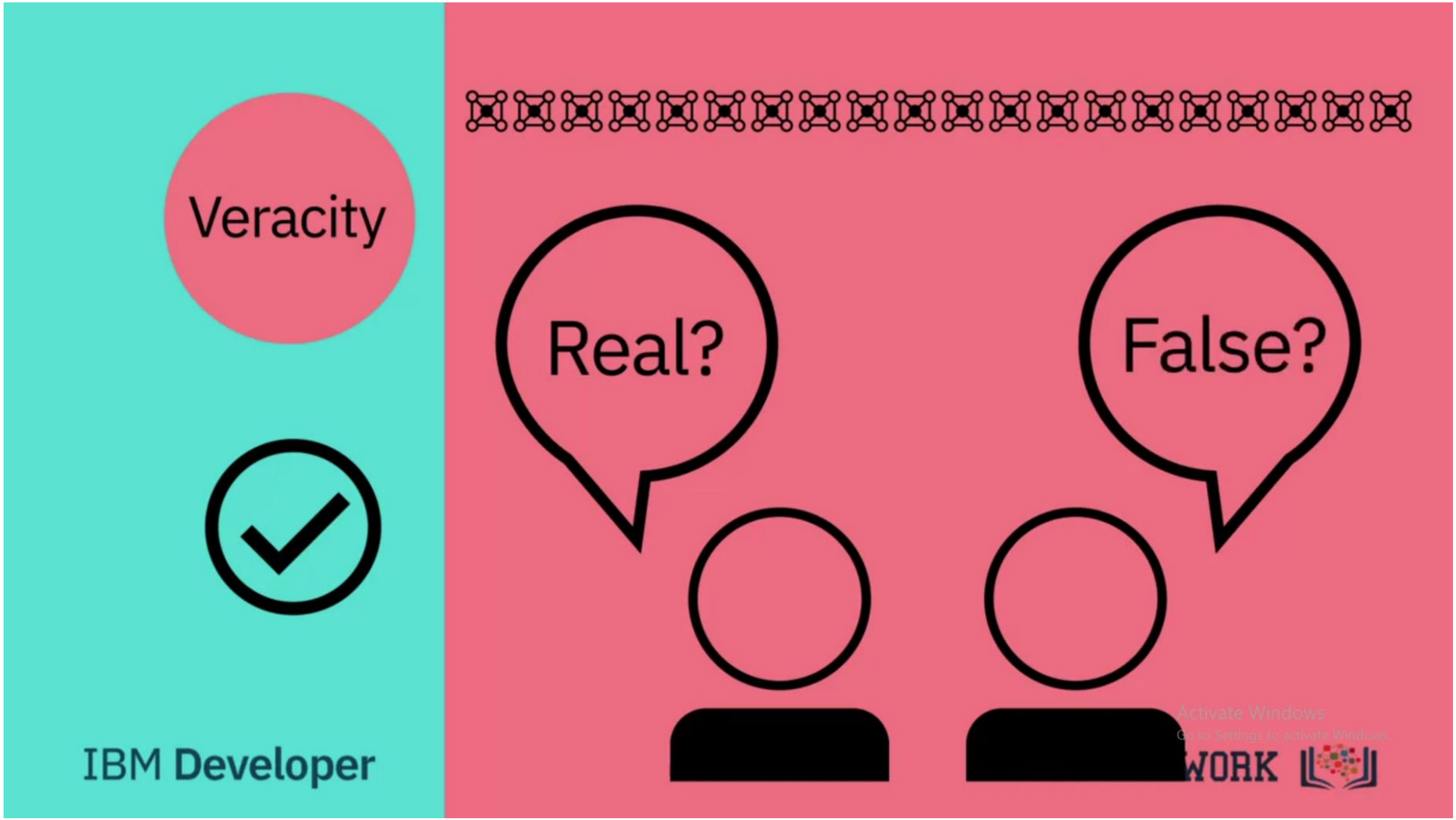
Drivers



IBM Developer

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.


Veracity
Real?
False?
IBM Developer
Activate Windows
Go to Settings to activate Windows.
WORK

Value

IBM **Developer**

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.

Value

Activate Windows
Go to Settings to activate Windows.

Value

IBM Developer

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.

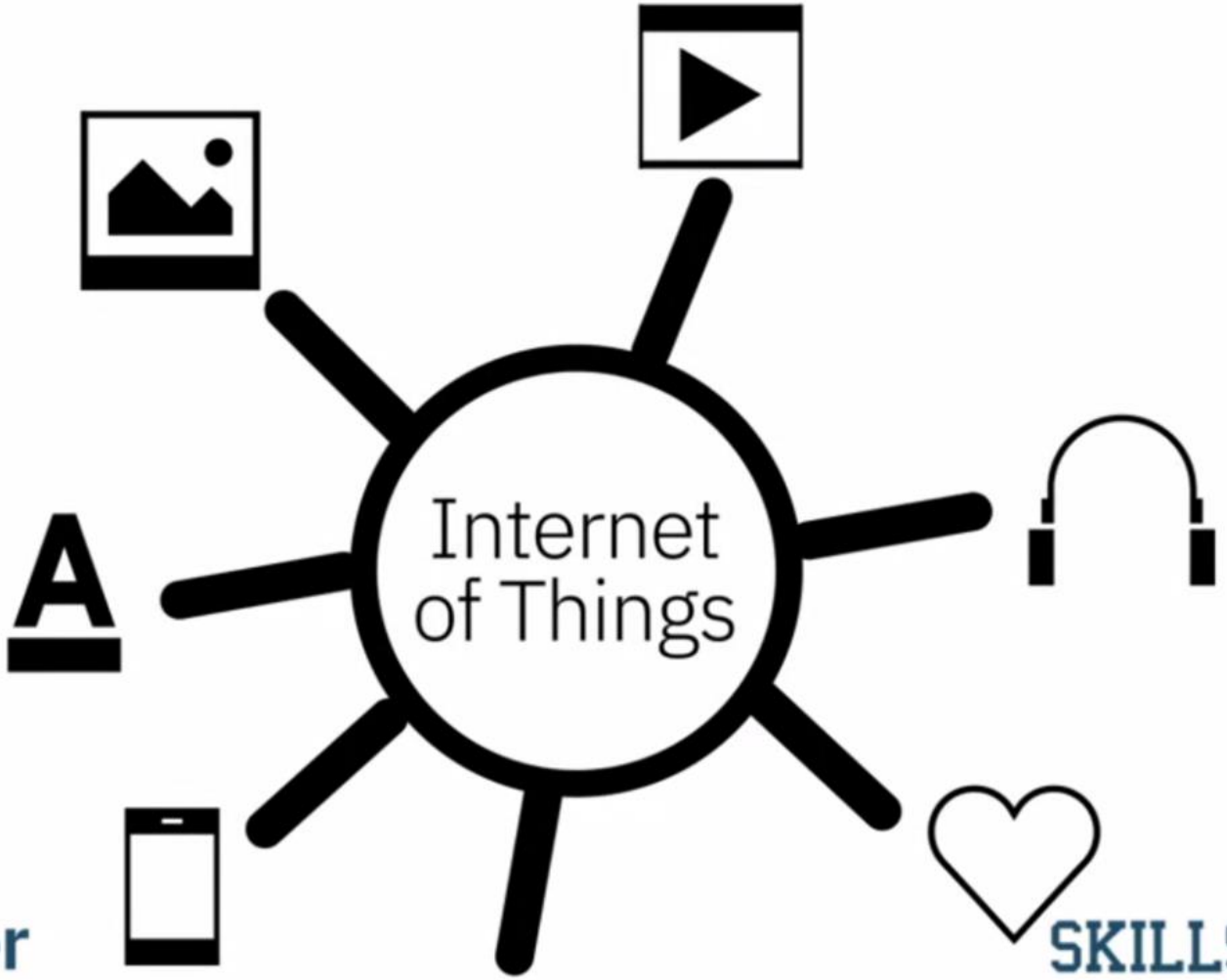
Velocity

2

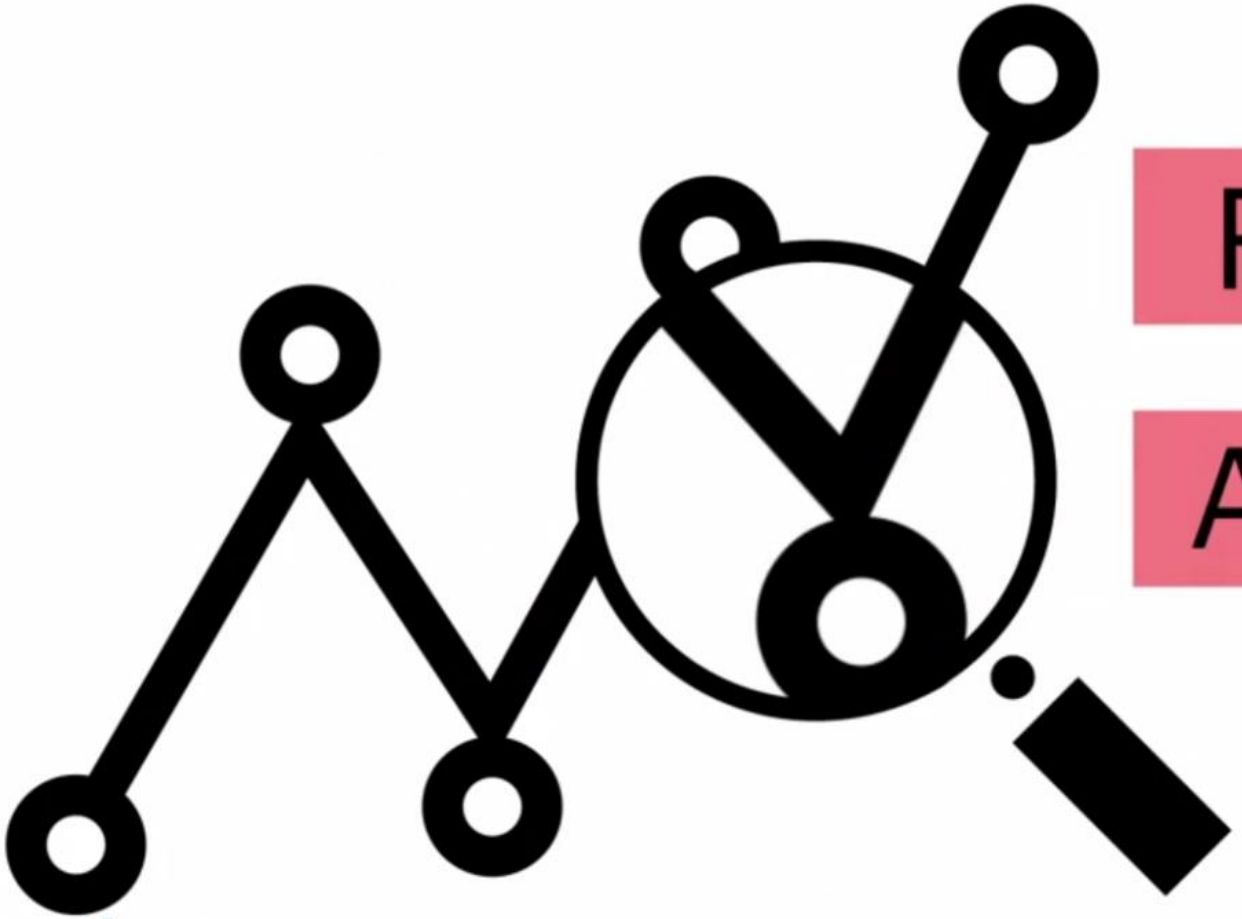
Volume

Volume

Variety



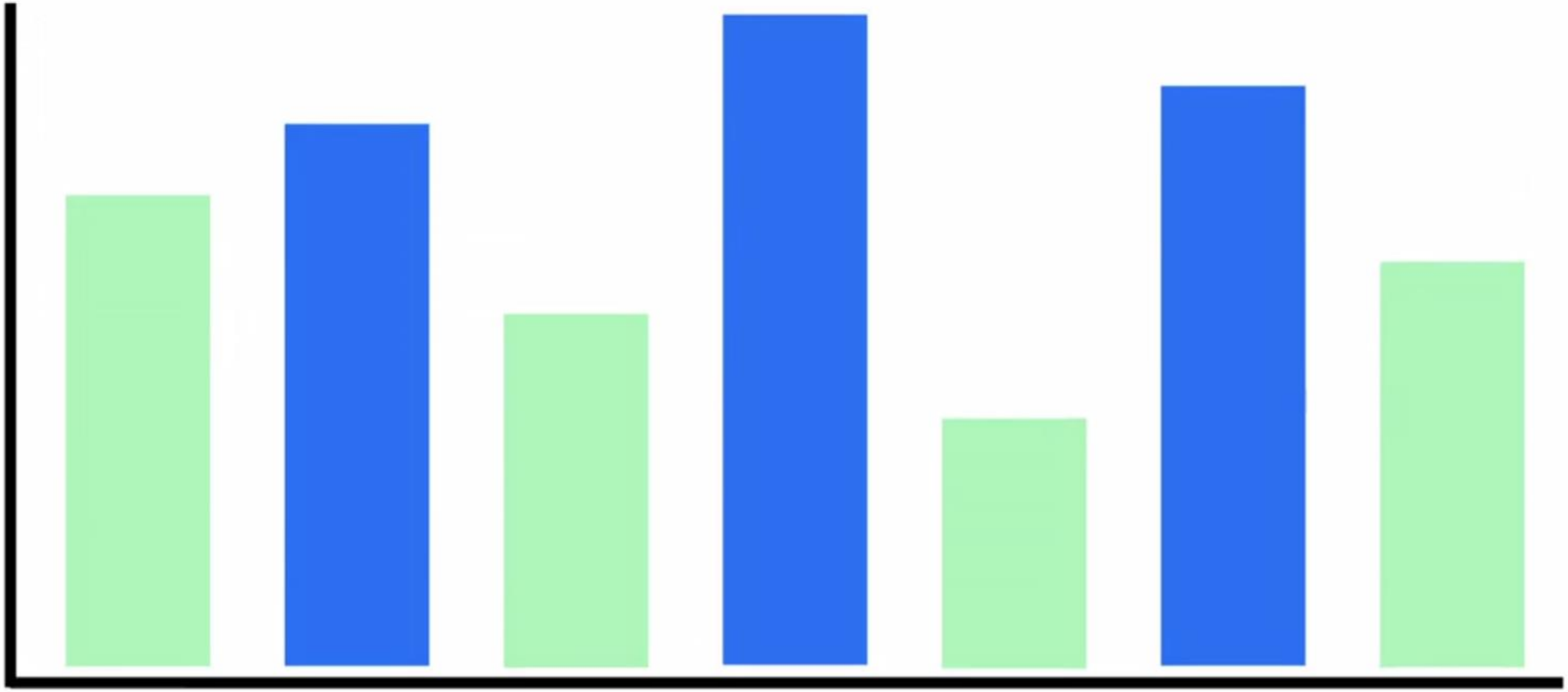
Veracity



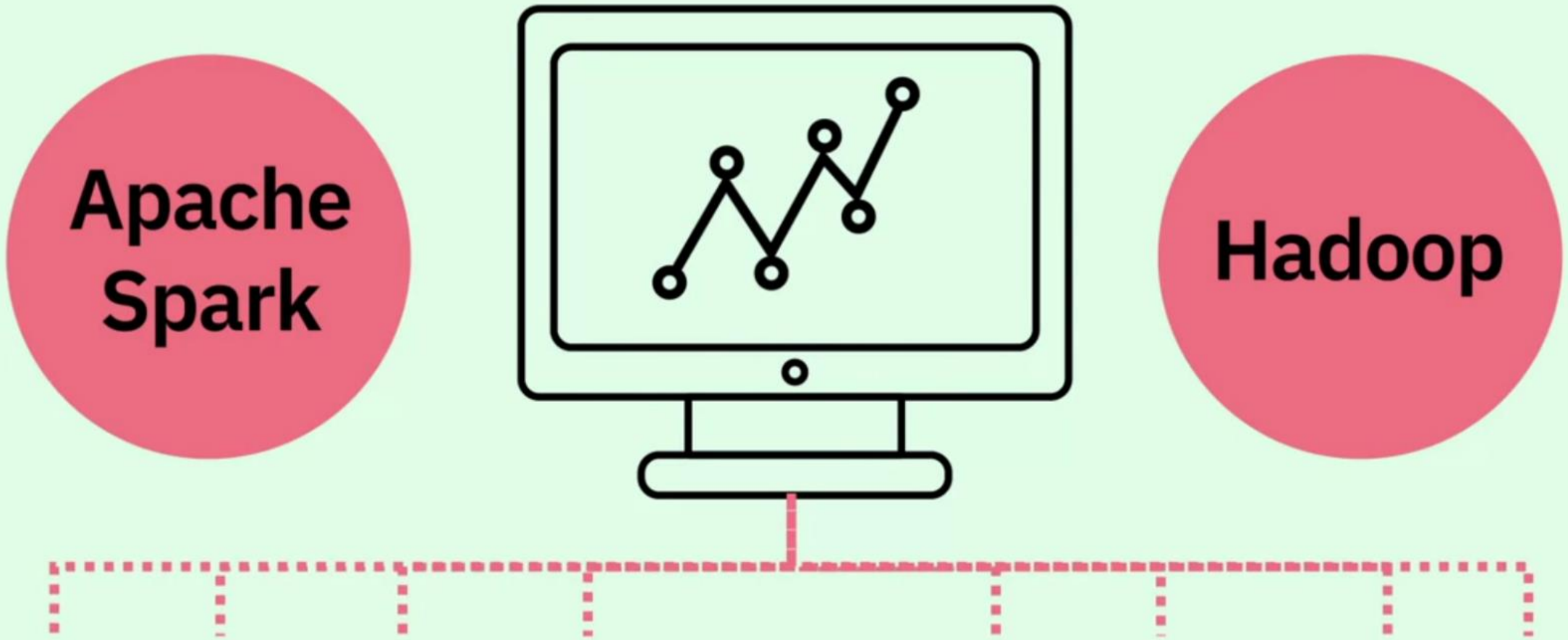
Reliable

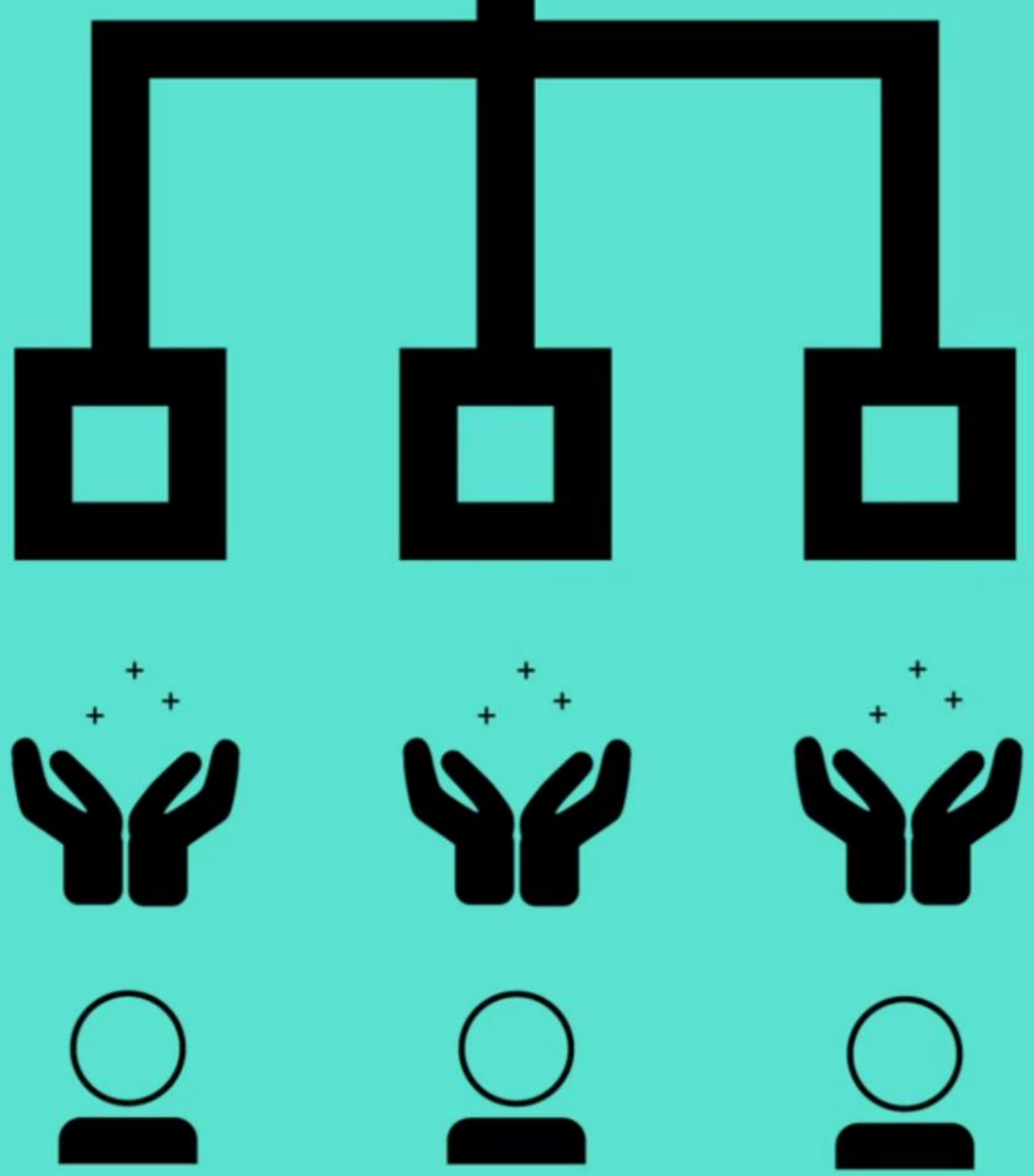
Accurate

Veracity



Alternative tools





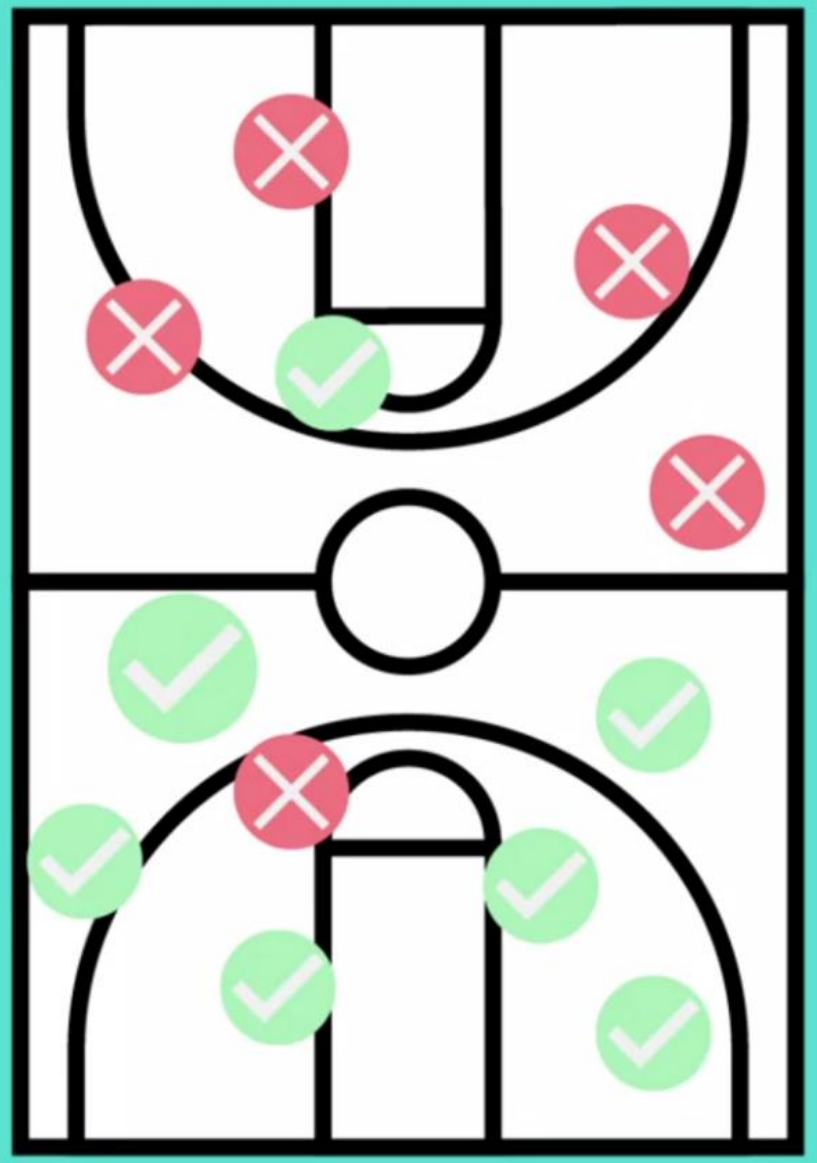
IBM Developer

SKILLS NETWORK

Netflix

IBM Developer

SKILLS NETWORK



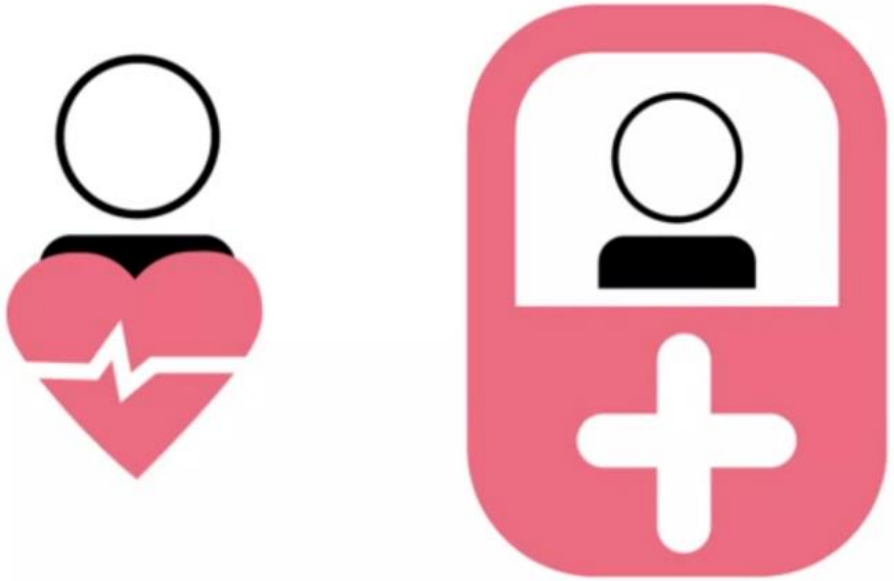
IBM Developer

SKILLS NETWORK

DexLab

A QUICK GUIDE TO DATA MINING

Data Science techniques



IBM Developer

SKILLS NETWORK

Data Science in Healthcare

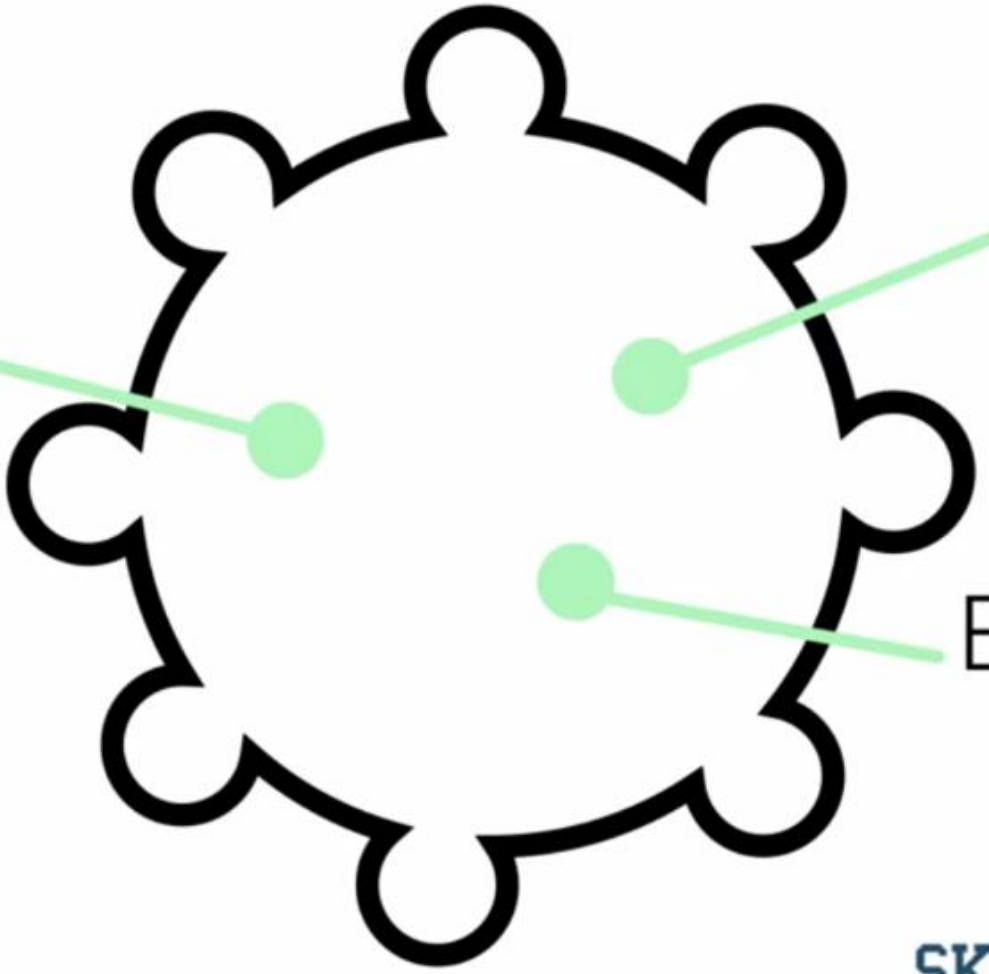
Data mining

Data modeling

Statistics

Machine learning

Data Science in Healthcare

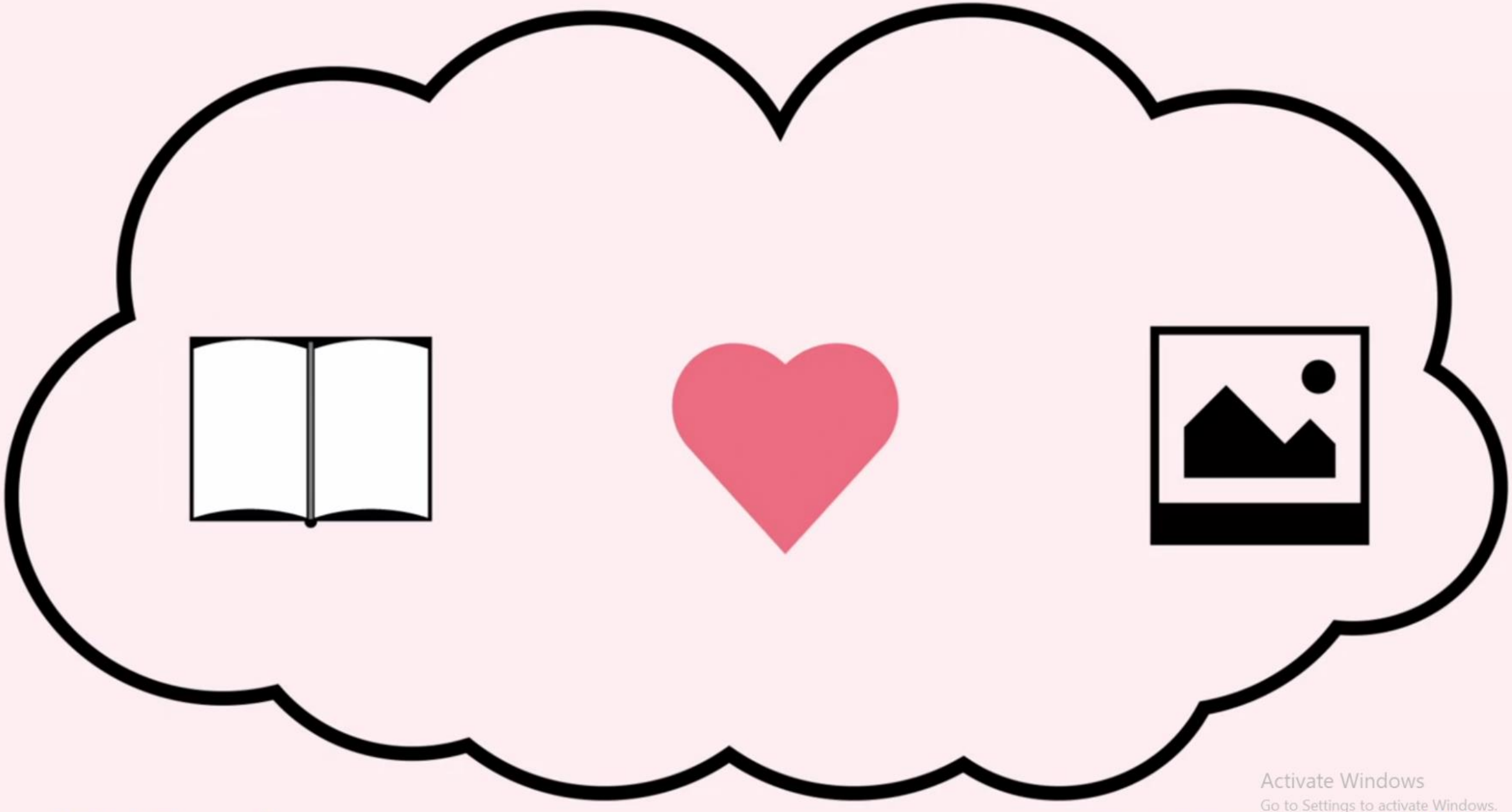


Activate Windows
Go to Settings to activate Windows.

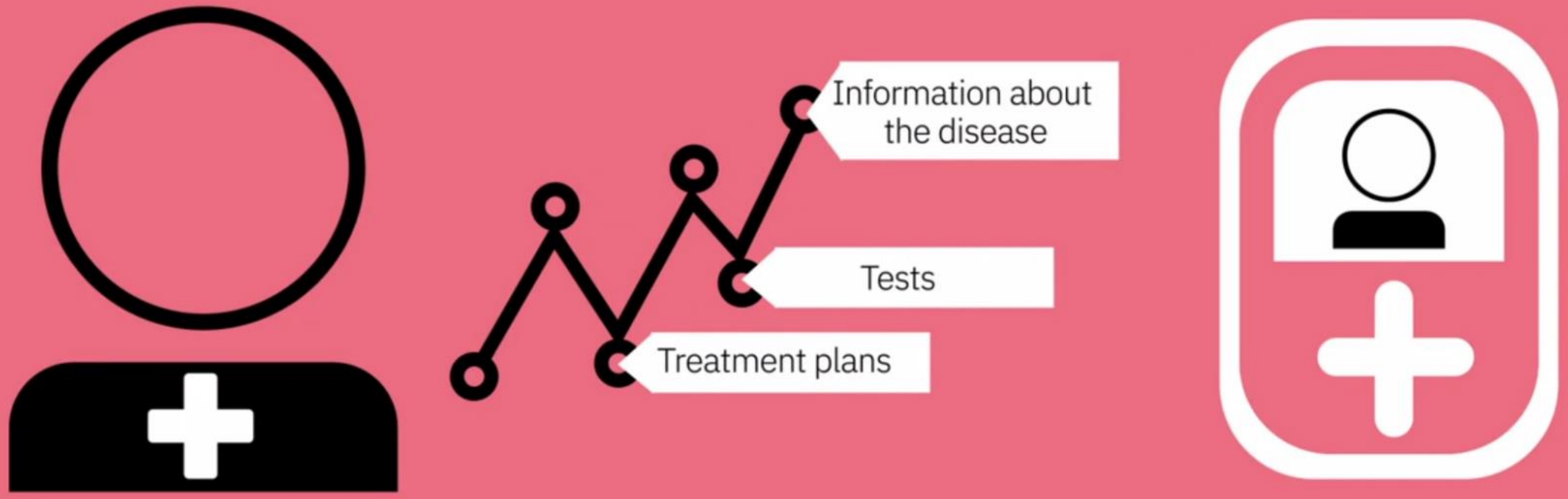
Data Science in Healthcare

- Appropriate tests
- Suitable trials
- Suggested treatments

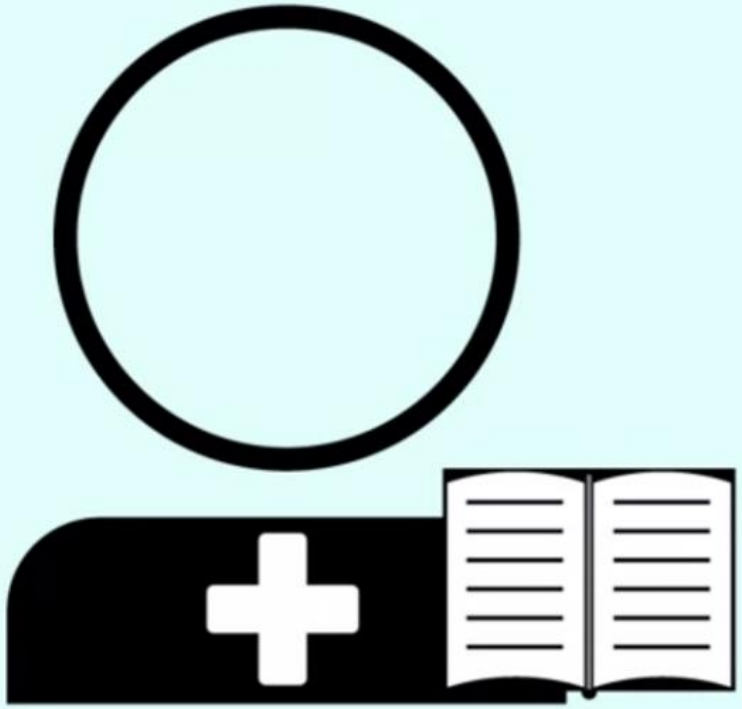
Activate Windows
Go to Settings to activate Windows.

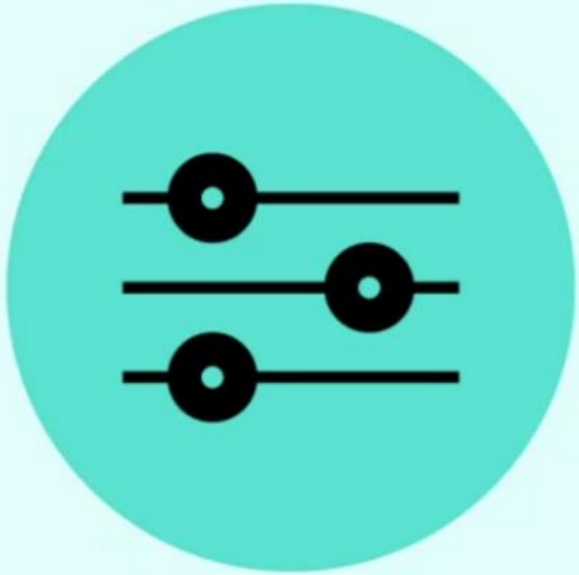


Activate Windows
Go to Settings to activate Windows.

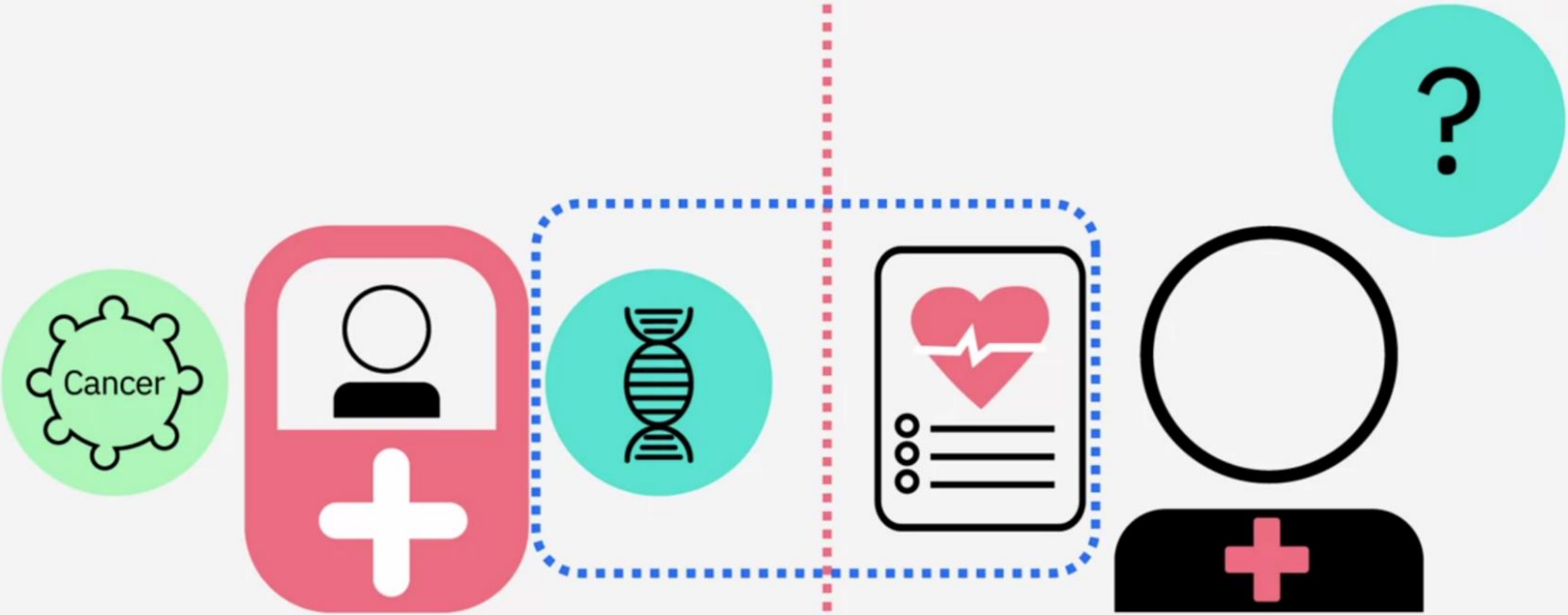


Activate Windows
Go to Settings to activate Windows.





Activate Windows
Go to Settings to activate Windows.



Activate Windows
Go to Settings to activate Windows.

Data Mining

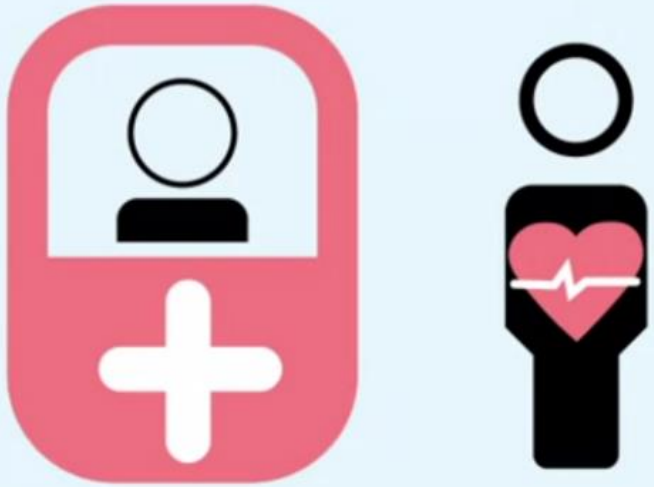
NorthShore University HealthSystem

Chicago

Electronic

Medical

Records



Activate Windows
Go to Settings to activate Windows.

Data Mining

Activate Windows
Go to Settings to activate Windows.

Disaster

Preparedness

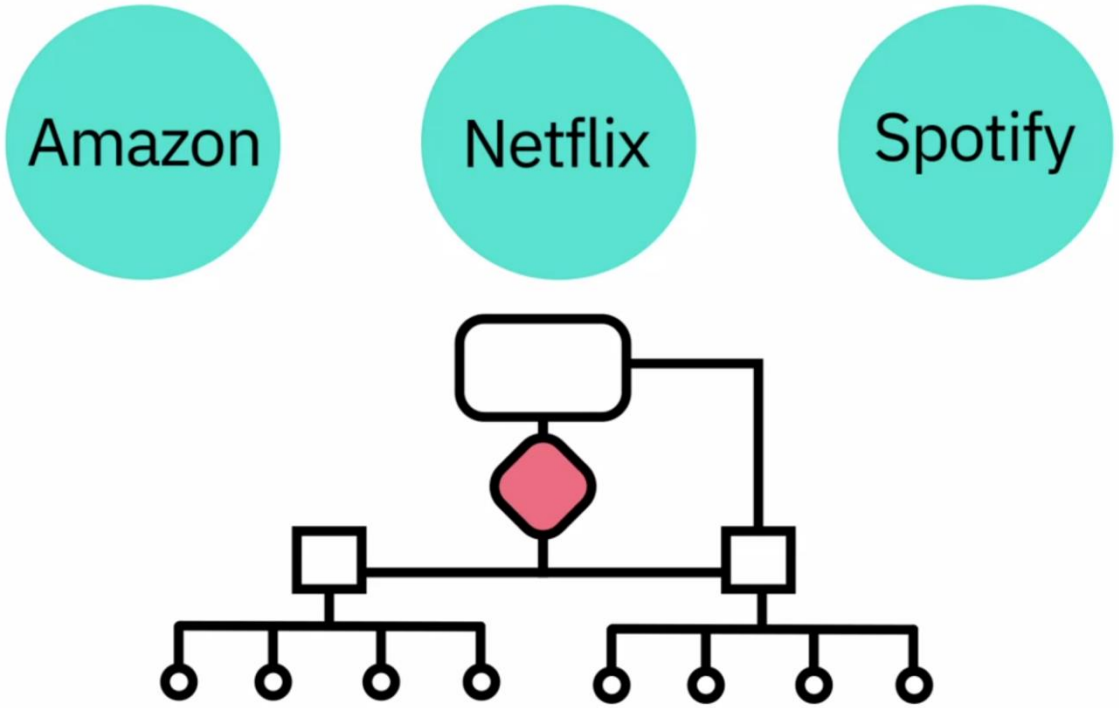
Activate Windows
Go to Settings to activate Windows.

- Earthquakes
- Hurricanes
- Tornados
- Floods
- Volcanic eruptions

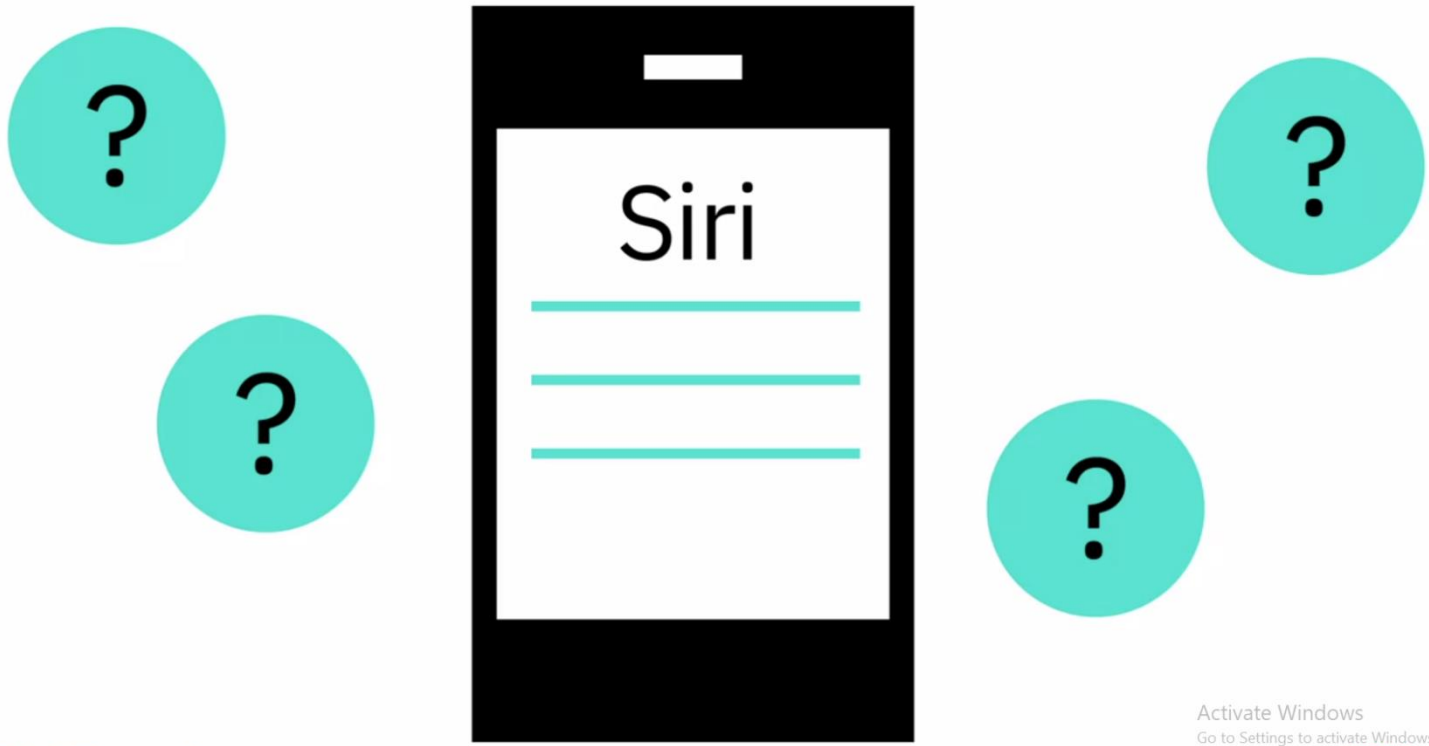
Activate Windows
Go to Settings to activate Windows.

Recommendation Engine

Activate Windows
Go to Settings to activate Windows.



Activate Windows
Go to Settings to activate Windows.

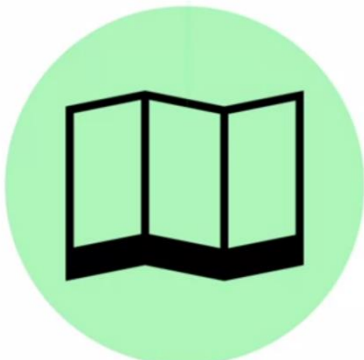


IBM Developer

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.

Google



IBM **Developer**

Activate Windows
Go to Settings to activate Windows.

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.

Business Impact

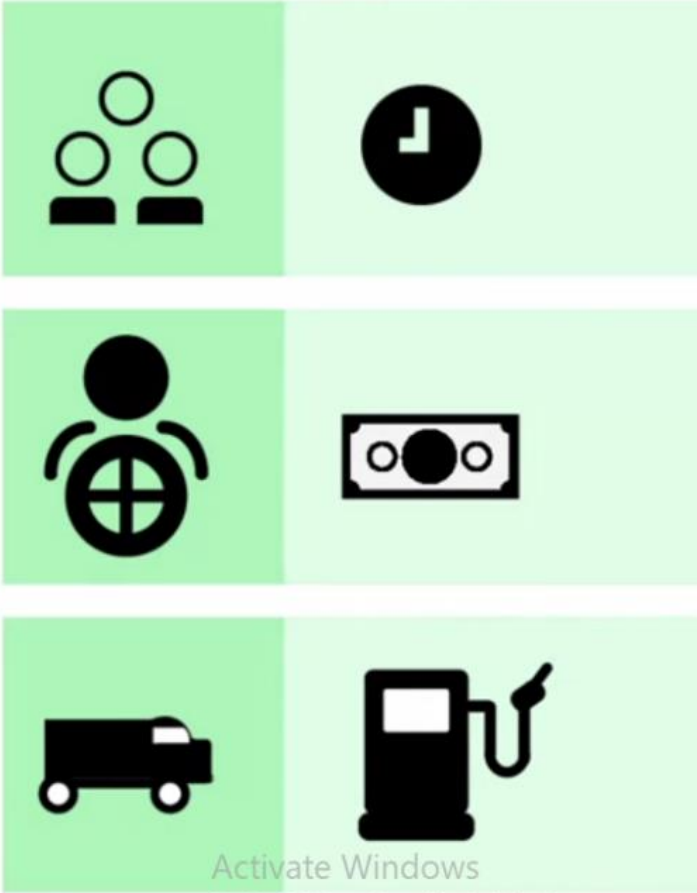
UPS

2013

Route

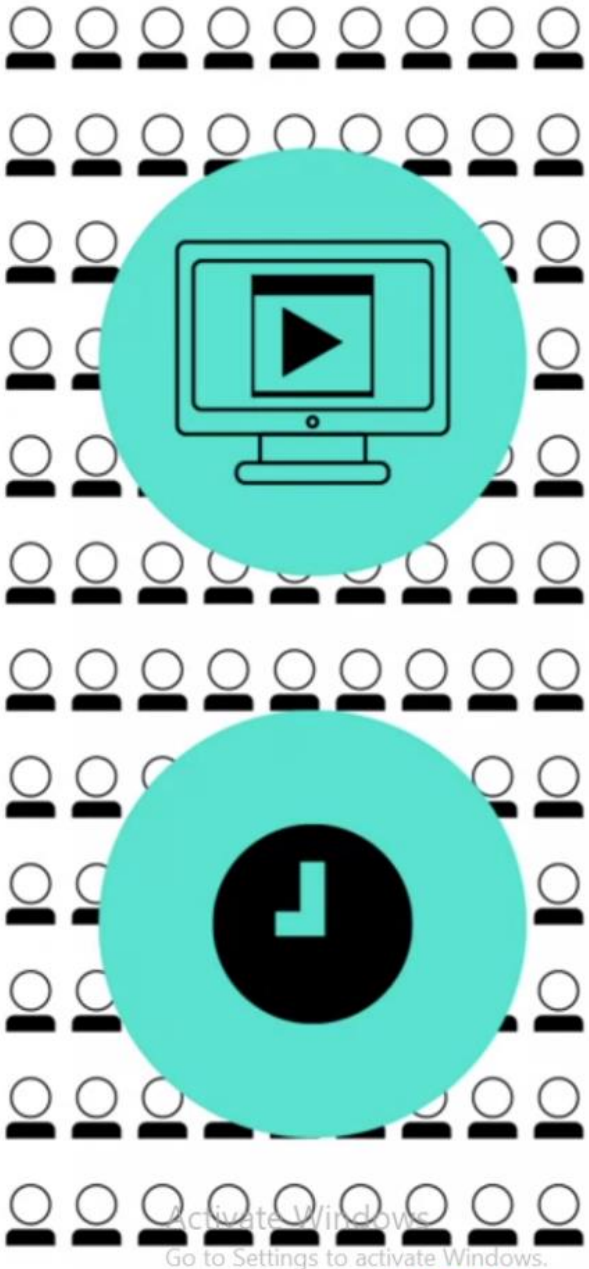
Guidance

System



Activate Windows
Go to Settings to activate Windows.

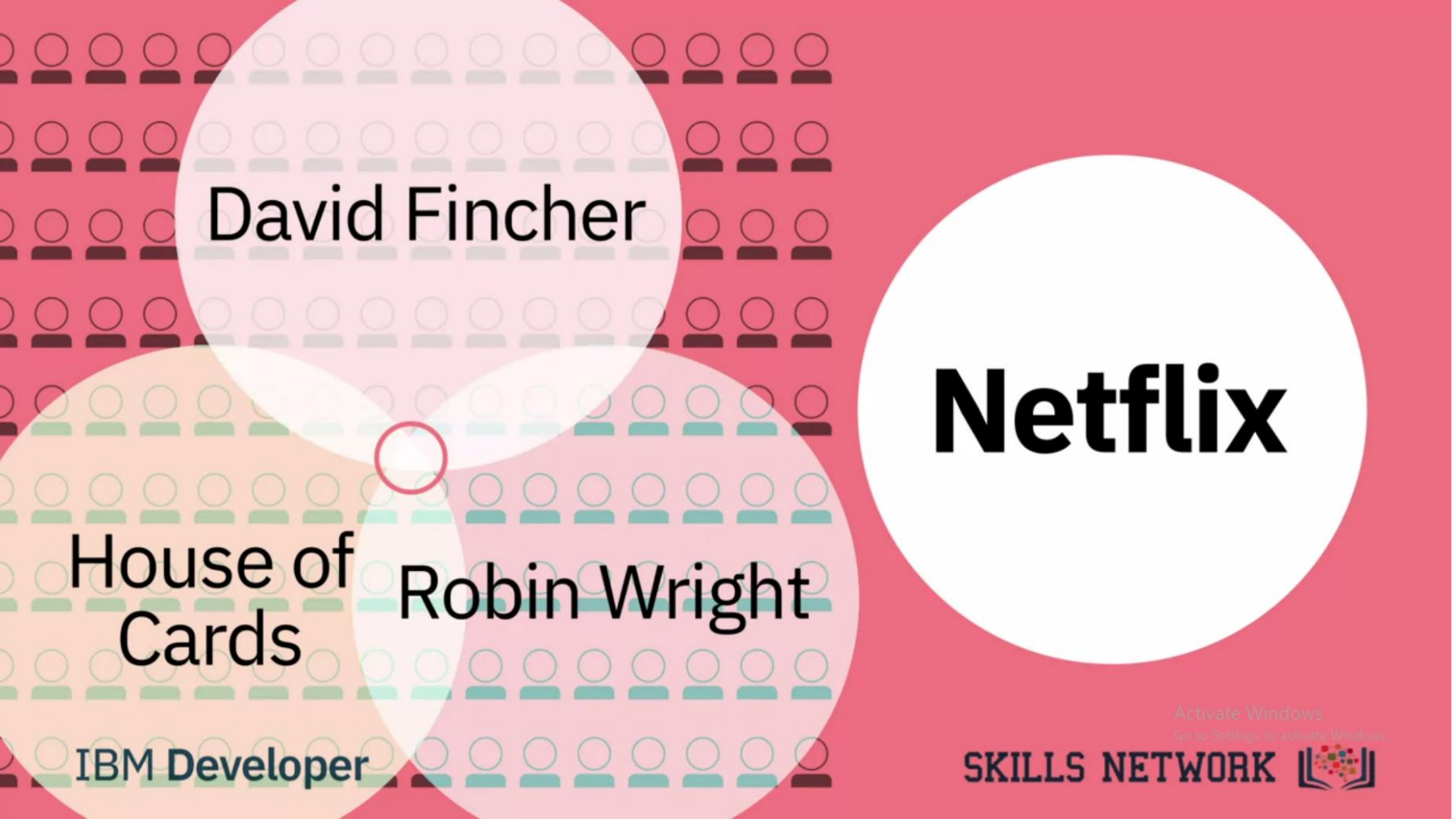
Netflix



IBM Developer

SKILLS NETWORK

Activate Windows
Go to Settings to activate Windows.


David Fincher
Netflix
House of Cards
Robin Wright
IBM Developer
SKILLS NETWORK