



# **KOMPETENSI PUSTAKAWAN DALAM IMPLEMENTASI KECERDASAN BUATAN UNTUK Mendukung Peningkatan Riset di Perguruan Tinggi**

Irhamni Ali, Ph.D

Data Scientist / Pustakawan Muda

Pusat Analisis Perpustakaan dan Pengembangan Budaya Baca

Perpustakaan Nasional RI

Disampaikan pada Seri Webinar

PERPUSTAKAAN UNDIVERSITAS DIPONEGORO

Semarang, 28 Agustus 2024



### **Informasi Personal**

Name : Irhamni Ali

### **Informasi Professional**

Instansi : Pusat Analisis Perpustakaan dan Pengembangan Budaya Baca  
Perpustakaan Nasional

Alamat : Jl.Salemba Raya No.28 A Jakarta ( [fnuirhamni@my.unt.edu](mailto:fnuirhamni@my.unt.edu) )

Posisi : Library Data Scientist / Post Doctoral Fellow

### **Education :**

Bachelor : Ilmu Perpustakaan & Informasi - Universitas Indonesia : 2000-2005

Master : Ilmu Komputer - IPB : 2011-2014

Doctoral : Information Science (Data Science Concentration) University of North Texas : 2020 – 2024

### **Professional Social Media**

- <https://www.linkedin.com/in/irhamni-ali-3985b455/>

### **Pubikasi**

- [https://www.researchgate.net/profile/Irhamni\\_Ali2](https://www.researchgate.net/profile/Irhamni_Ali2)

### **Penghargaan**

- University of North Texas Law and Legal Scholarship Award
- Fulbright Ph.D Scholarship 2020-2023
- Best Promising Future Library Leader In ASEAN By International Network On Emerging Library Inovators - Asean (INELI-ASEAN) • Vietnam • November 2017
- Best Collaborative Project Team Leader By International Network On Emerging Library Innovators - ASEAN (INELI-ASEAN) • Singapore • October 2016
- International Federation Of Library Association And Institution (IFLA) Asia And Oceania Grant In United States Of America. August 2016

# AGENDA

PENDAHULUAN

DAMPAK BIG DATA PADA  
PERPUSTAKAAN

INTRODUCTION TO DATA SCIENCE

ISU TERKINI TERKAIT DATA DI  
PERPUSTAKAAN

PENGEMBANGAN SDM

KESIMPULAN

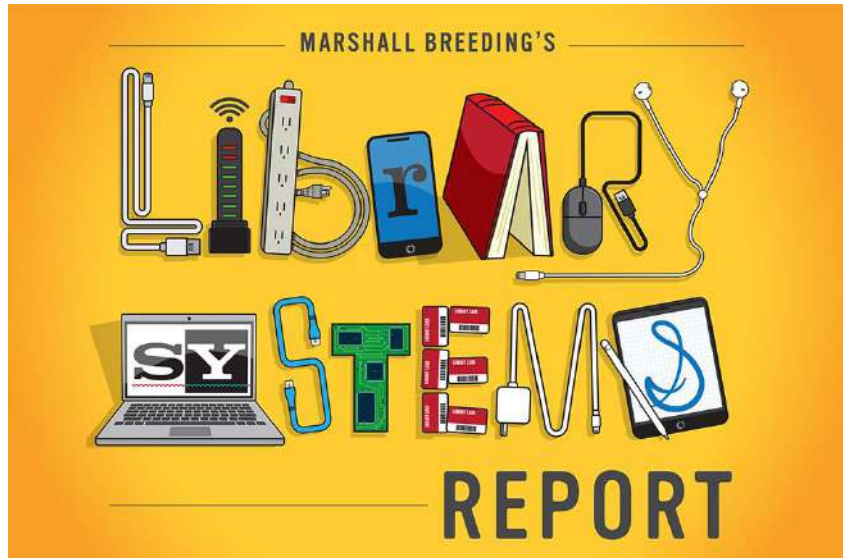
# PENDAHULUAN



"The world's most valuable resource is no longer oil, but data." (The Economist, 2017)

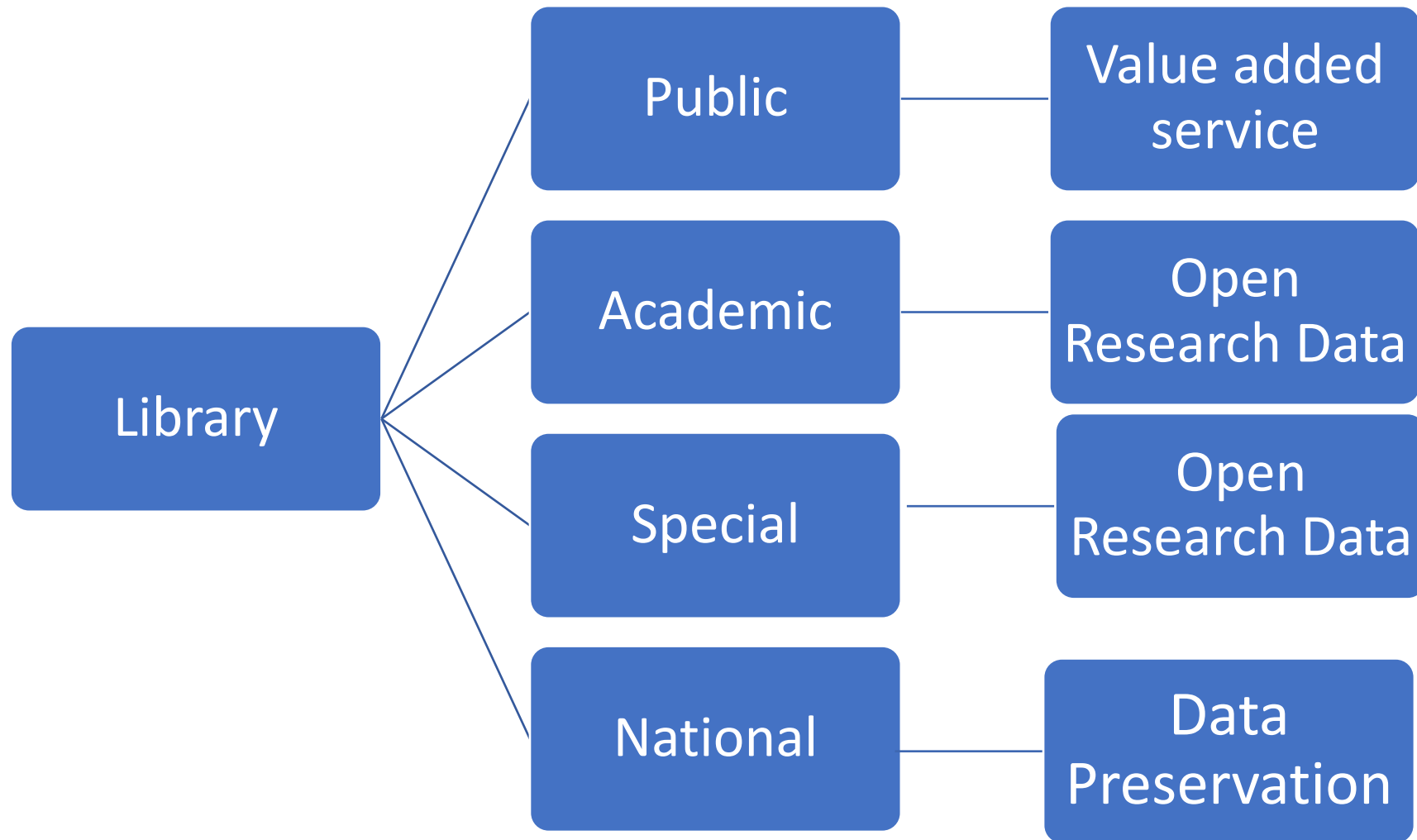
Source :<https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

# DAMPAK BIG DATA PADA PERPUSTAKAAN



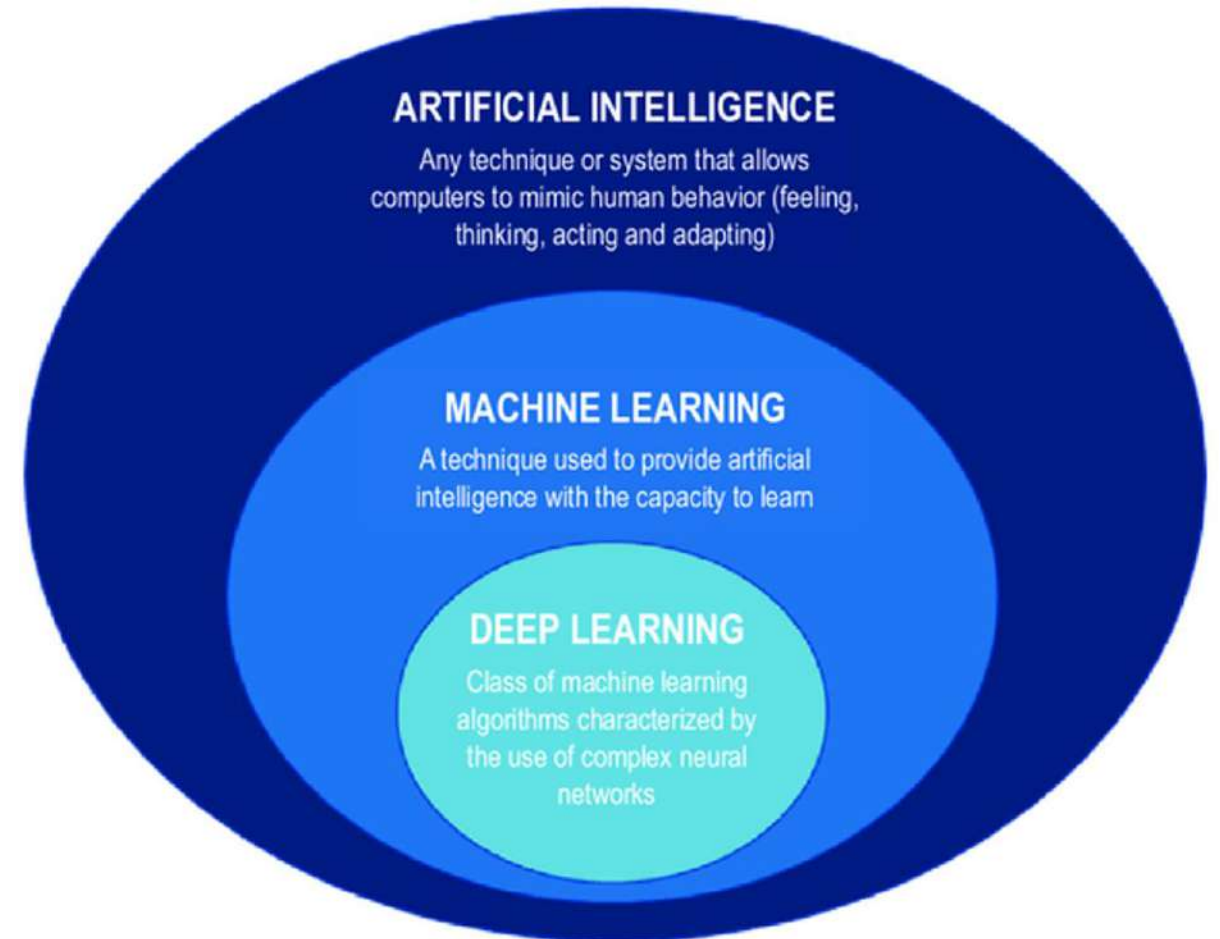
- **Cornell University in Ithaca**, New York, has switched from the Ex Libris Voyager ILS to FOLIO, with support from EBSCO Information Services.
- **Cultuurconnect**, which includes almost all public libraries in the Flanders region of Belgium, has begun implementing OCLC Wise, moving from multiple incumbent systems.
- **Libraries Ireland**, comprising all public libraries in Ireland, has selected Civica Spydus to replace the Sierra ILS that it had been using since the inauguration of the countrywide library management project.
- **Montgomery County (Md.) Public Libraries** has switched from Symphony to Koha and Aspen Discovery, with support from ByWater Solutions. This system, with 21 locations serving more than 1 million people, is the largest in the US using Koha.
- **Queens (N.Y.) Library** chose SirsiDynix Symphony, replacing an Innovative Virtua ILS in place since 2008. The library serves 2.3 million people across 65 locations.
- **St. Louis Public Library (SLPL) and the St. Louis County (Mo.) Library** have selected Innovative's Polaris as a shared ILS. The county library system previously used Sierra; SLPL will migrate from Symphony.
- **University of Chicago Library** migrated from now-defunct Quali OLE to FOLIO, with support from Index Data.
- **University of Missouri Libraries** has implemented FOLIO with support from EBSCO Information Services, moving from an Innovative Sierra ILS.

# DAMPAK BIG DATA PADA PERPUSTAKAAN



# DAMPAK BIG DATA PADA PERPUSTAKAAN

- Artificial Intelligence (Kecerdasan Buatan) : Teknik yang memungkinkan computer memiliki kecerdasan seperti manusia
- Apa itu Machine Learning (ML)? *ML* cabang dari kecerdasan buatan yang mampu memantau tren tertentu dengan menganalisis dan mengidentifikasi pola-pola (Das & Islam, 2021).
- Model AI yang sering digunakan adalah model ML yang seperti Multinomial Naive Bayes, Gaussian Naive Bayes, Decision Tree, SGD Classifier, LightGBM Classifier, Random Forest Classifier.
- Untuk Deep Learning sering digunakan CNN (Convolutional Neural Network)

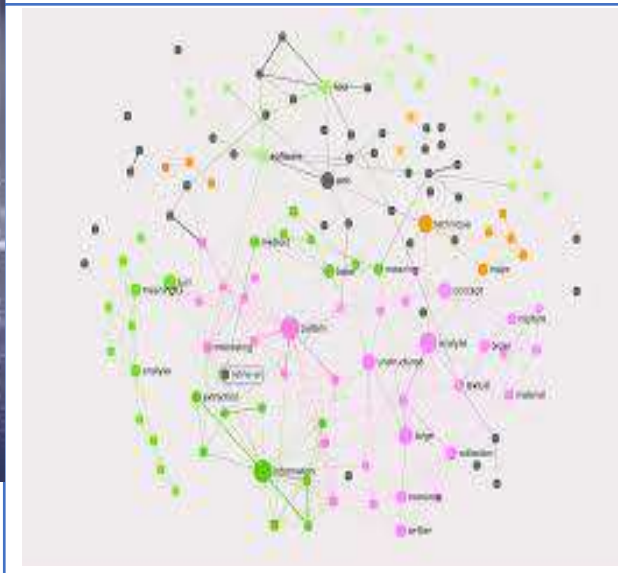


# IMPLEMENTASI KECERDASAN UNTUK MENDUKUNG PENINGKATAN RISET DI PERGURUAN TINGGI

Digitalization



Text Mining



Textual Big Data Analytics



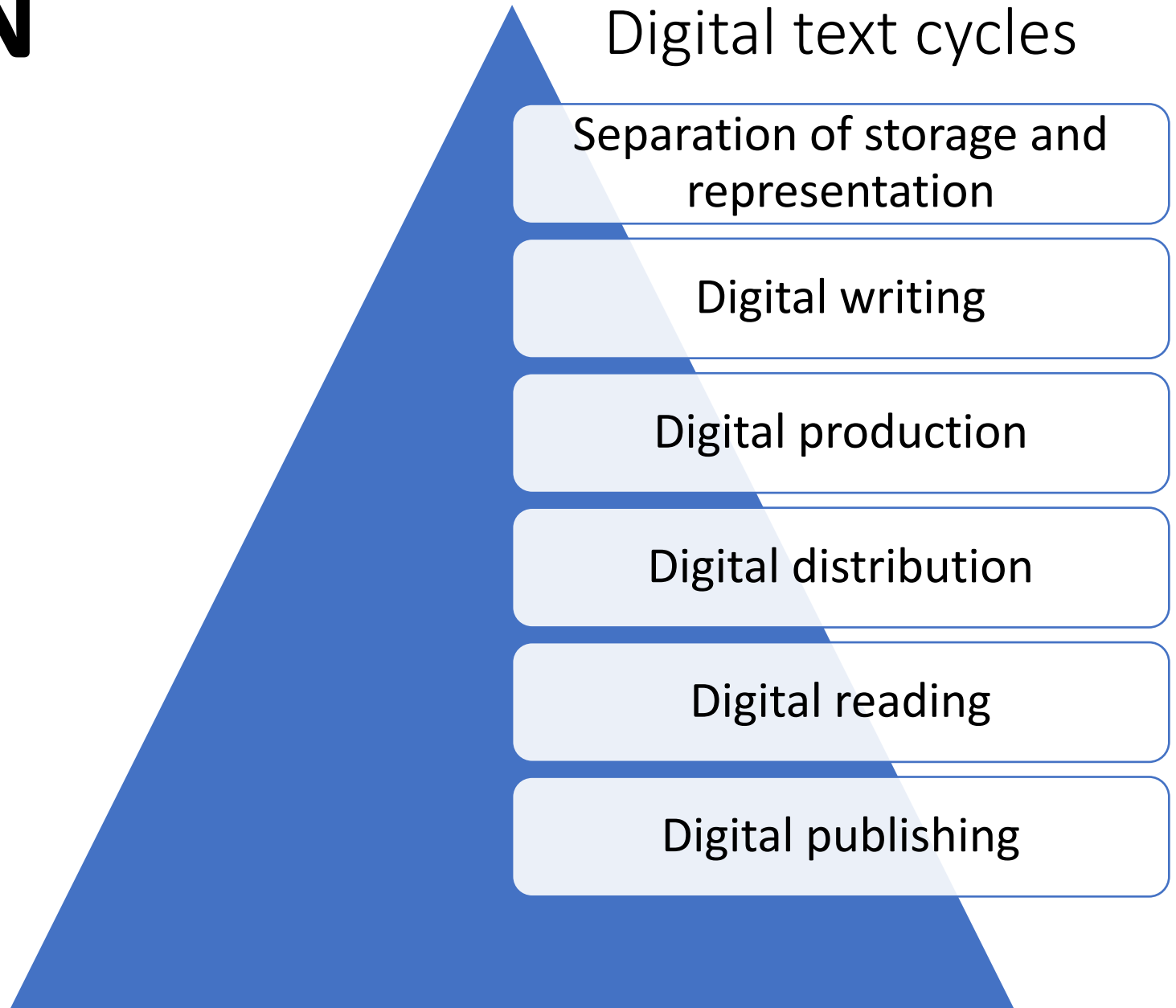
BIG DATA



# DIGITALIZATION

**Digitalization** is a process mainly based on principles of print in a digital environment that represents an effort to use basic digital principles in a more flexible way of producing and distributing verbal texts.

Sources :  
Hillesund, T. (2005). Digital text cycles: from medieval manuscripts to modern markup.



# TEXT MINING

Text mining is a new field that attempts to glean meaningful information from natural-language text. In Text mining sometimes it may be loosely characterized as the process of analyzing text to extract information that is useful for particular purposes. It most commonly targets text whose function is the communication of factual information or opinions, and the motivation for trying to extract information from such text automatically is compelling – even if success is only partial. “Text mining” (sometimes called “text data mining”

Sources : Witten, I. H., Don, K. J., Dewsnip, M., & Tablan, V. (2004). Text mining in a digital library.

## Text Mining Methods

Information Extraction (IE)

Information Retrieval (IR)

Natural Language Processing

Clustering

Categorization

# Text Mining in Practices

ABOUT PEOPLE PROJECTS PUBLICATIONS FOR SCHOLARS FOR STUDENTS RESOURCES

Program in Islamic Law

ISLAMIC LAW CALENDAR PIL NEWS HARVARD EVENTS JOBS, GRANTS, CFPS RELATED LINKS & ORGS

NEWS PREVIOUS NEWSLETTERS

OPPORTUNITIES

## Islamic Law, Data Science, and AI Lab | Spring 2021 | Harvard Law School

Posted on January 29, 2021

**ALL FOR APPLICATION**

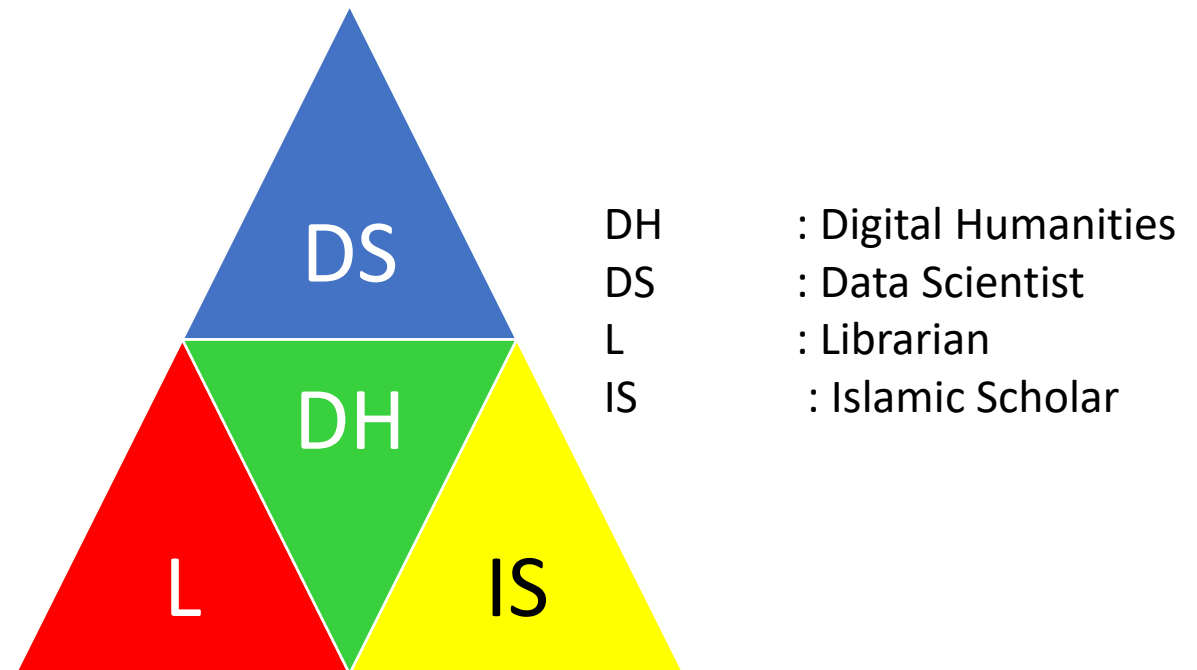
Islamic Law, Data Science, and AI Lab | Spring 2021 | Harvard Law School

Convened by Intisar Rabb & Zahra Takhshid, Harvard Law School

The PIL Islamic Law Lab seeks applicants interested in presenting or developing works in progress at the intersection of Islamic law and data science for the Spring 2021 Lab, with respect to their own research or to ongoing projects within our Lab. Specifically, our Lab will focus on emerging tools in the Islamic digital humanities / data science space, and to developing new components of our in-house data science tools: Courts & Canons: (1) creating historical gazetteers for the Islamic world, (2) 'hacking the library' to make bibliographic library searches in our field more useable and precise, and (3) better understanding and mapping Islamic legal genres. This Lab is also launching a research project on "AI Principles in Islam," by which we aim to assess and develop ethical and legal guidelines that arise in AI from Islamic perspectives. Generally, research themes for which we seek proposals include research queries on Islamic law that make use of a data science/digital humanities tool, method, or a data set to answer novel questions in the fields of Islamic law or history; or algorithms, digital corpora, and other DH / data science tools (e.g., KITAB-project.org, e-Shia.ir, or others). We are open to a diverse set of research inquiries within these specific or general areas. Public lab sessions presenting research will last for one hour, consisting of a 15-20 presentation and 40-45 minutes Q&A. Additional lab lessons (by invitation) will be devoted to building tools native the Lab. We hope the Lab will facilitate students and collaborating scholars to produce publishable papers and data science tools over the course of the Spring semester. Selected work may be considered for publication opportunities on the Islamic Law Blog or the Journal of Islamic Law (research), and the SHARIASource portal (tools).

The Lab will focus on emerging tools in the Islamic digital humanities / data science space, and to developing new components of our in-house data science tools from Court (Fatwa)/ Qanun :

- (1) creating historical gazetteers for the Islamic world,
- (2) 'hacking the library' to make bibliographic library searches in our field more useable and precise, and
- (3) better understanding and mapping Islamic legal genres.



# TEXTUAL BIG DATA ANALYTICS

Textual Big Data Analytics in Cultural and science historians, information scientists and text mining experts are currently addressing these questions in how to implement the text mining tools to study long-term developments and transformations a discourse in a systematic, longitudinal, and **quantifiable way.**

## Big Data Textual Analytics Methods

Heuristics

Computational Analytics

Quantitative trans-disciplinarity

Reproducibility

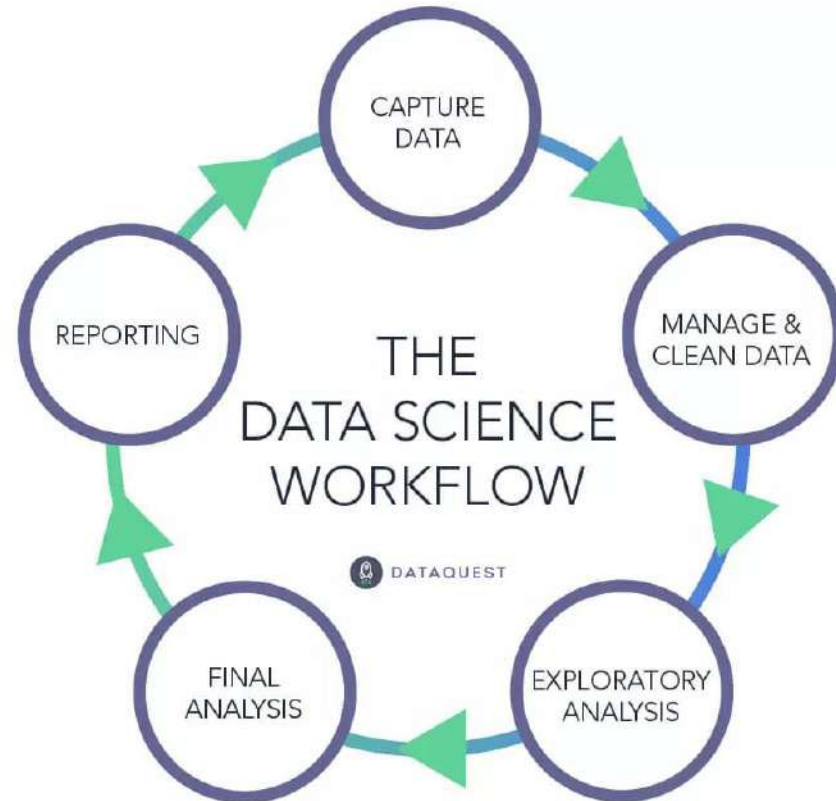
# INTRODUCTION TO DATA SCIENCE

## What is Data Science ?

Data science is a discipline that provides principles, methodology and guidelines for the analysis of data for tools, values, or insights (Yan, D., & Davis, G.E. (2019). A First Course in Data Science. *Journal of Statistics Education*, 27, 99 – 109)

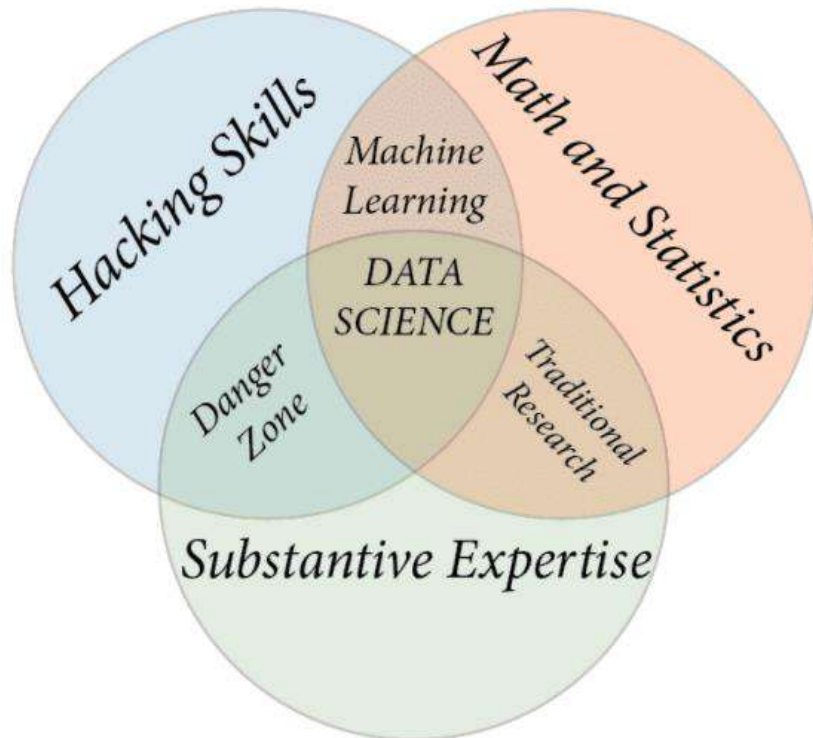
Source : <https://www.youtube.com/watch?v=xC-c7E5PK0Y>

# Data Science Workflow

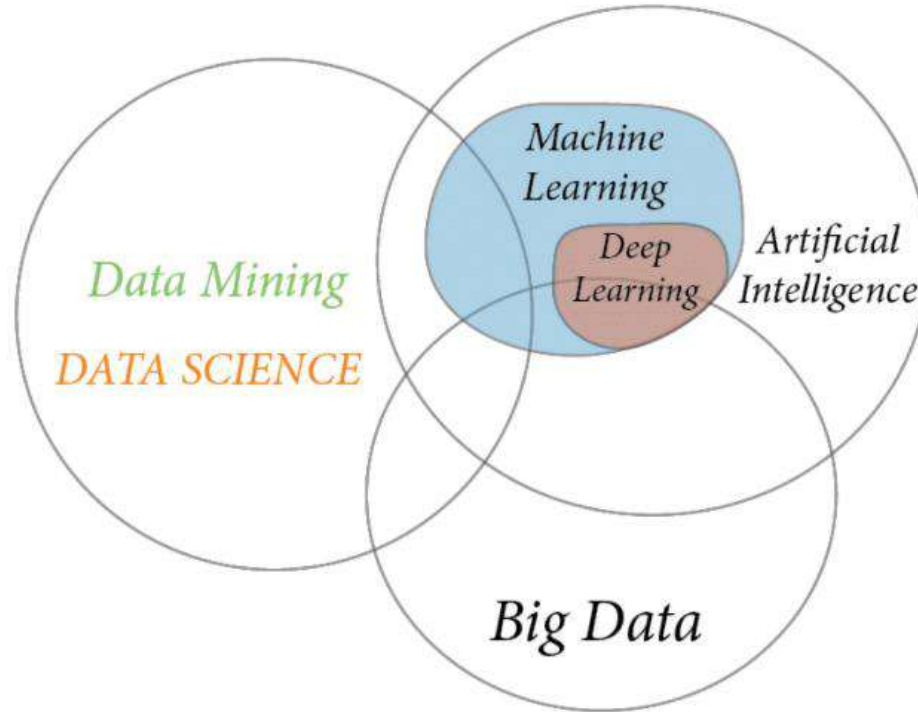


# CONCEPT OF DATA SCIENCE

## What's Data Science?



2010



2015

Source :

<https://www.kdnuggets.com/2016/10/learn-data-science-8-steps.html>

# Skill Untuk Data Science

## Top 10 skills in 2020

- Complex problem solving.
- Critical thinking.
- Creativity.
- People management.
- Co-ordinating with others.
- Emotional intelligence.
- Judgment and decision making.
- Service orientation.
- Negotiation.
- Cognitive flexibility.

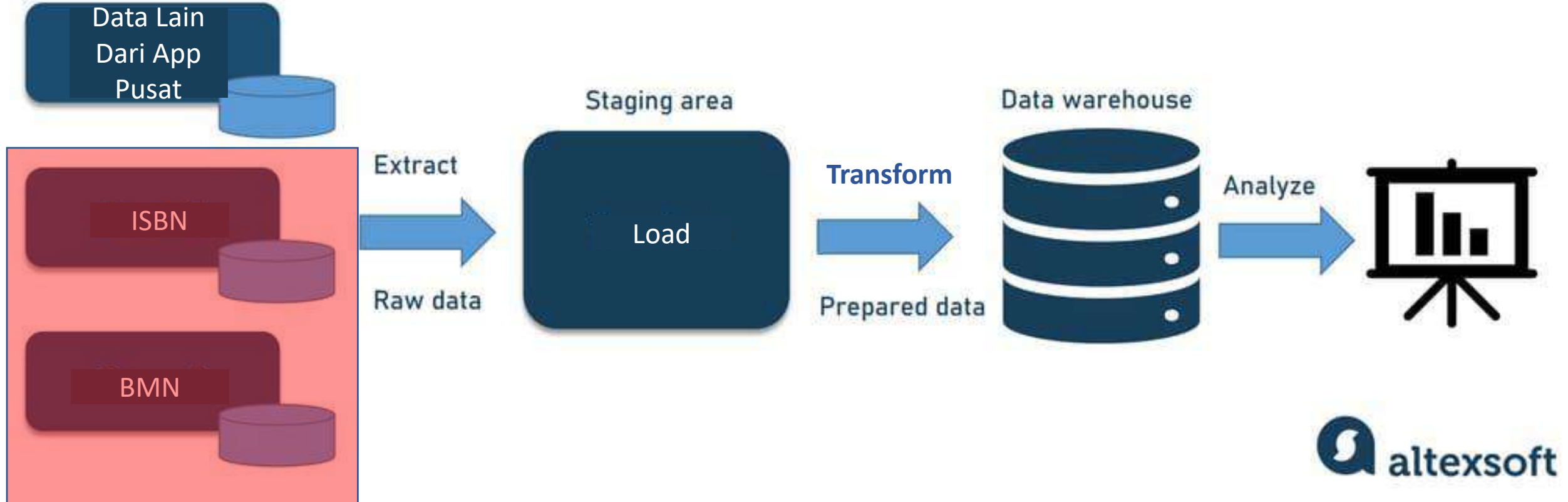
## Top 10 skills in 2005 (compared with 2020 statistics)

- Complex problem solving.
- Co-ordinating with others (-3).
- People management (-1).
- Critical thinking (+2).
- Negotiation (-4).
- ~~Quality control.\*~~
- Service orientation (-1).
- Judgment and decision making (+1).
- ~~Active listening.\*~~
- Creativity (+7).

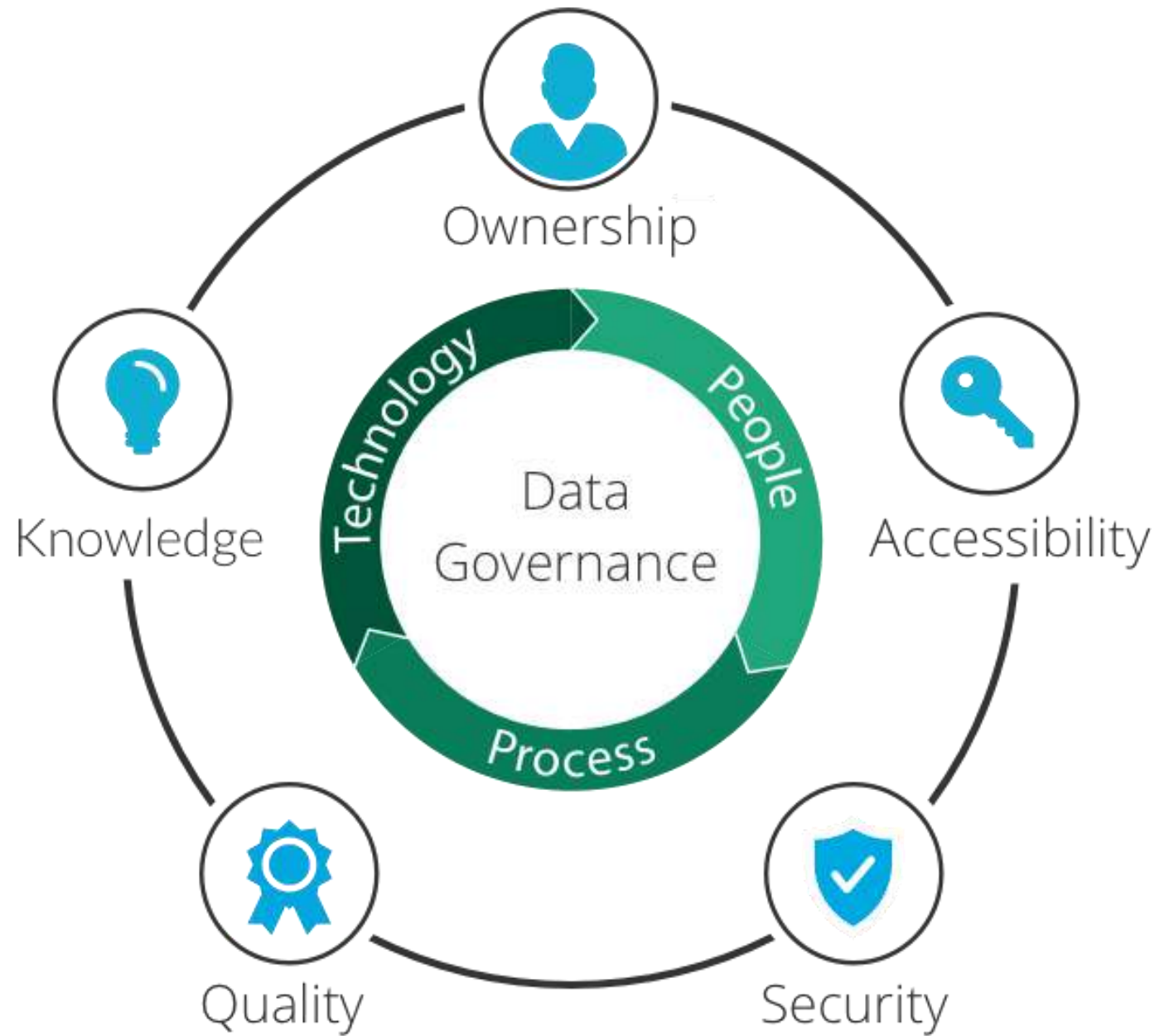


# Infrastruktur Data Science

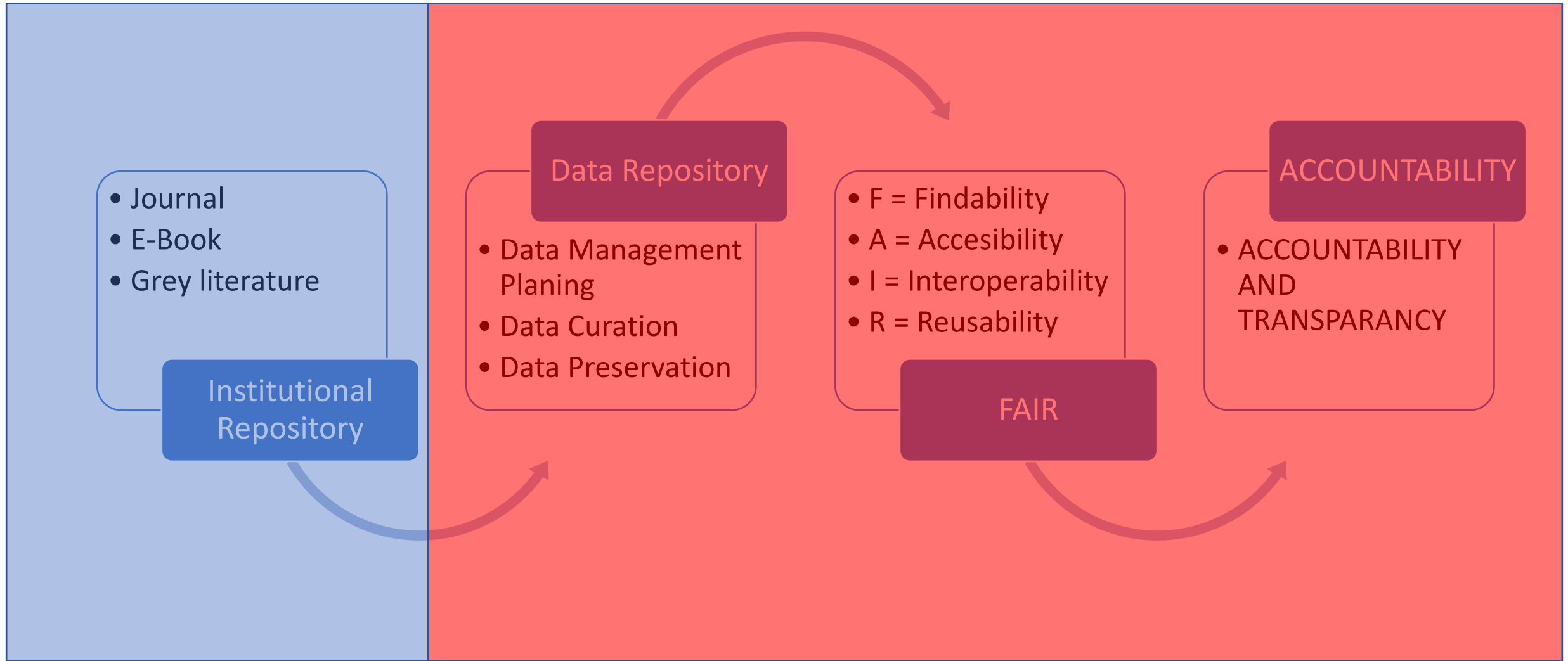
## ETL PIPELINE



# Data Governance



# ISU TERKINI TERKAIT DATA DI PERPUSTAKAAN



# PENGEMBANGAN SDM

## Jalur Pembelajaran AI untuk Pustakawan

### 1. Dasar Pemrograman dan Matematika

Pelajari Python, tingkatkan pemahaman tentang aljabar linear, probabilitas, dan statistik.

### 2. Pengenalan AI dan Pembelajaran Mesin

Pahami konsep dasar AI dan pembelajaran mesin.

### 3. Mendalami Pembelajaran Mesin

Perdalam pemahaman tentang pembelajaran mesin.

### 4. Proyek AI Praktis

Implementasikan proyek AI praktis dan ikuti kompetisi.

### 5. Spesialisasi dan Eksplorasi Topik Lanjutan

Eksplorasi pembelajaran mendalam dan NLP.

### 6. Pembelajaran Berkelanjutan dan Keterlibatan Komunitas

Tetap update dengan berita AI terbaru, bergabung dengan komunitas, dan jaringan.

### 7. Etika dan Implikasi AI

Pahami implikasi etis dan dampak sosial AI.

# Data Science & Computational Thinking

*Decomposition:* Breaking down data, processes, or problems into smaller, manageable parts

*Pattern Recognition:* Observing patterns, trends, and regularities in data

*Algorithm Design:* Developing the step-by-step instructions for solving this and similar problems

*Abstraction (Story telling) :* Identifying the general principles that generate these patterns by focusing on the important information only, ignoring irrelevant details; applying to similar problems

# KESIMPULAN

- Pustakawan Harus up to date
- Perpustakaan perlu mengubah orientasi dari Information Provider menjadi Data Provider
- Perpustakaan Harus berbenah di segala aspek (infrastruktur data & SDM) untuk melakukan Inovasi
- Investasi di Pengembangan SDM menjadi hal utama





Q & A  
Sessions