





For Whom We Should Create the Bibliographic Data ! The Rise of Machine

By

Mo'men S. Elnasharty Ph.D.

Library, Documentation and Information Dept. Data Science & AI Dept. Faculty of Graduate Studies of Statistical Research, MDS Cairo University, Egypt Navigator001@cu.edu.eg

> Brussels 18th Aug 2023

Agenda

- Title: For Whom We Should Create the Bibliographic Data: The Rise of Machine
- Location: KBR.
- **Duration:** 1 Days (10:10 pm 10:35 am).
- Speaker: Moamen Elnasharty, PhD.
- Audience: Librarians, Knowledge Managers and Academic Staff.
- Organizer: IFLA WLIC 2023.



Outline

Bibliographic Control.
Al & Machine Learning.
Generate Al Based on BC,





In the

Past

The Human Being Was The Primary and The Only Source Of Data Production by:

Authorship, Discovery, Creativity, Extrapolation, Conclusion, Invention ...

The Products was

Manuscripts, Printed Books and Journals, Audiovisual, Digital Resources !



So, the Metadata and Bibliographic Control Co

 $\frac{1}{2}$

.....

 \sum

 \sum

Metadata

- Metadata is basically data that describes Entites or Information Resources.
- It helps us understand the origin, structure, nature, and context of data.
- As a result, we can categorize, organize, and then easily retrieve information.



https://dataedo.com/kb/data-glossary/what-is-metadata

Bibliographic Control

A broad term encompassing all the activities involved in Creating, Organizing, Managing, and Maintaining the file of Metadata Records representing the items held in a library, Bibliographies, Databases, Museums or archival collection, to Facilitate Access to the information contained in them.





The first Bibliographic Record Backes to the Ancient Egyptian Civilization, Amenhotep III, The first Bibliographic Record Backes to the ancient Egyptian civilization, Amenhotep III, by hieroglyphic inscription, it had been about "the book of the moringa tree",







The **Pinakes** is a lost bibliographic work composed by **Callimachus** that is the first library catalog in the world; its contents were based upon the holdings of the Library of Alexandria during Callimachus' tenure there during the third century BCE.

At 990 C.E

Kashf al-Zunun 'an Asami al-Kutub wa al-Funun (The Removal of Doubt from the Names of Books and the Arts) is a bibliographic encyclopedia of books and sciences compiled by Turkish writer Kâtip Çelebi. It was written in Arabic.

Cataloged titles was of approximately 15,000 books; 9,500 names of authors; and 300 sciences and arts. The work is seen as a significant example of and contribution to Ottoman historiography. The Fihrist, or The Catalogue: is an encyclopedic bibliographical work completed in 987 C.E. by the Baghdadi bookseller Ibn al-Nadim (d. 990 C.E.).

like the card catalogue for the combined libraries of tenth-century.

the single most important source on the translation of Greek, Persian, and Sanskrit scientific legacies into Arabic in the 8th-10th centuries Conrad Gesner is known as the "father of bibliography". In 1545 he published Bibliotheca universalis (Universal Bibliography 1545–1549) (Twenty-one Books of Encyclopedias or Universal Divisions).





The Roles of Otlet & La Fontaine in Universal Bibliographic Vision

- Both played key roles in advancing the concept of universal bibliographic control through their works
 - The Universal Decimal Classification.
 - The Universal Bibliographic Repertory.
 - The Mundaneum: as a central repository for the world's information.
 - The World City or Cité Mondiale.



The First Intellectual Contributions in the Modern Era to Reach The Standard Bibliographic Catalogue backs to the efforts of Antonio Panizzi in 1841, where this effort was Clear in his book "Ninety-One Cataloguing Rules", where he established the Bibliographic Rules for Describe the Collections of British Museum. Charles Cutter's efforts in 1876 are considered one of the greatest efforts in this context, that led to the identification of three basic objectives for bibliographic Catalogs, namely:

The ability to enable users to find books by author – title - subject.
Showing the library's information resources about: a specific author. on a specific topic.
Assistance in choosing information resources based on edition. or subject.

Ithort

U3



in 1908 a Radical and Important changes has been made in Bibliographic Control Efforts, Where efforts have become an institutional rather than an individual, where the American Library Association and the British Library issued a set of joint bibliographic rules, that became the core of the Anglo-American Cataloging Rules.

The Catalogue Has to Tell You More Than What You Ask For The Answer Of a Good Catalogue Is Not To Say Yes Or Not, But To Tell The User That The Library Has The Item In So Many Editions And Translations, And You Have Your Choice.

Seymour Lubetzky (1898 – 2003).

Henriette Avram, who developed the MARC format (Machine Readable Cataloging), that opened a new horizons for Data representation to Machine.

Tillett & Bibliographic Models





THE DAILY NEWS

www.dallynews.com

THE WORLD'S FAVOURITE NEWSPAPER

- Since 1879

عبقرية التأليف العربي، مؤلف يكشف عظمة المؤلفين



Corem Ipsum In libris graecis appetere mea. At vim odio lorem omnes, pri id iuvaret partiendo. Vivendo menandri et sed. Lorem volumus blandit cu has.Sit cu alia porro fuisset.

Ea pro natum invidunt repudiandae, his et facilisis vituperatoribus. Mei eu ubique altera senserit, In libris graecis appetere mea. At vim odio lorem omnes, pri id iuvaret partiendo. Vivendo menandri et sed. Lorem volumus blandit cu has.Sit cu alia porro fuisset.

Ea pro natum invidunt repudiandae, his et facilisis vituperatoribus. Mei eu ubique altera senserit,

ARAFAT'S EFFORTS

Bibliographic Control Based on Relationships



We are ROTTERDAM

IFLA.org

IFLA was established in 1927 at Edinburgh, Scotland. In 1971, the Head Quarter was shifted at Netherland, the Hague.z

IFLA is a Non-Governmental, Not-For-Profit, an International Organization.

IFLA- provides, Information Specialists to the World, with a forum for exchange Ideas,

IFLA Promoting International Co-Operation, Research, and development in all fields of library and its services

#WeAreIFLA

IFLA & Universal Bibliographic Control

Universal Bibliographic Control (UBC) is a concept has been Coined by Herman Liebaers (the president of IFLA).

According to the UBC "... Any Document would only be cataloged once in its country of origin, and that record would then be available for the use of any library in the world..."

In the currently era, IFLA has played a central role, stimulating National Bibliographic Agencies to promote standards and collaborations that go beyond the national sphere, leading to multicenter and even more cooperative bibliographic control.

Universal Bibliographic Control

• The Aim of UBC is Promotion of a worldwide system for the control and exchange of bibliographic information.

• Requirements of UBS are Need an Authoritative Organization-UNESCO, NBA- Create the Authoritative, Produce and Distribute the records in a standard physical form.





Library Catalogs

- A comprehensive list of bibliographic data of sources of information found in a given institution, arranged in an orderly manner for ease of retrieval (usually alphabetically by author, title, or subject).
- It is available in the form of recordings that reflect the formal and implicit descriptions of these sources.
- the main objectives from catalogues are:
 - Enable anyone to find any source of information.
 - Shows what the library contains.
 - To help choose the source of information

A systematic list of works written by a specific author or on a specific topic, or that share one or more common characteristics (language, form, period, place of publication, etc.).

WRIGHTS

LIFE OF

PLUTARCHYS

LIVES.

LIVES

31 bliograph

WRIGHTS

LIFE OF

GRAYS

POETICAL

WORKS.

WORKS

RKS

Search Engines

A search engine is a software program that helps people find the information they are looking for online using keywords or phrases. Search engines are able to return results quickly—even with millions of websites online—by scanning the Internet continuously and indexing every page they find.



How Can Create Bibliographic Control?










RESOURCE DESCRIPTION AND ACCESS



001		26132811
008		920608s1992 nyu bf 001 0 eng
010		\$a92021087
020		\$a1555701205
040		\$aDLC\$cDLC\$dAGL
049		\$aHNKA
050	00	\$aZ678.9\$b.C58 1992
082	00	\$a025/.00285\$220
100	1	\$aCohn, John M
245	10	<pre>\$aPlanning for automation :\$ba how-to-do it manual for</pre>
		librarians /\$cJohn M. Cohn, Ann L. Kelsey, Keith Michael
		Fiels
260		\$aNew York :\$bNeal Schuman Publishers,\$cc1992
300		\$avi, 116 p. ;\$c28 cm
440	0	\$aHow-to-do it manuals for libraries ;\$vno. 25
504		\$aIncludes bibliographical references and index
650	0	<pre>\$aLibraries\$xAutomation\$xManagement\$xHandbooks, manuals, etc</pre>
650	Θ	<pre>\$aLibrary planning\$xHandbooks, manuals, etc</pre>
700	10	\$aKelsey, Ann L
700	10	\$aFiels, Keith Michael



Based on Text Analysis

< >	** * *					Q Search
	target			Do	cuments	4
Image: Second state sta	Image Image <th>Best Three S Character : Character : Character :</th> <th>Arabic Character Recognition Arabic Character Recognition Clear Recognise Loading Done cores : alefui With Score : 1.0 meem-,</th> <th></th> <th>suments del.data</th> <th>Q. Search 4</th>	Best Three S Character : Character : Character :	Arabic Character Recognition Arabic Character Recognition Clear Recognise Loading Done cores : alefui With Score : 1.0 meem-,		suments del.data	Q. Search 4
O Remote Disc Tags	🗮 Macintosh HD 🔉 📷 Users 🔉 🏠	De Emaraic > 📄 NetBean	veloped by : Taha Emara Email: taha@emaraic.com sProjects > 🛅 ArabicCharsRecognition > 🛅 target > 🎅 A	n		
			1 of 4 selected, 28.32 GB available			-0

To be Result Like this!



The Q: Are BC Tools Meeting Users Needs Currently?

The Current State of Bibliographic Control Tools

A survey conducted by the OCLC Network in 2010 about The usage of research tools in academic societies in Europe and USA revealed that library catalogs has a last rank between other research tools in conducting scientific research.







Bibliographic Control Tools Don't Integrate with others





Despite it is online

🥑 Brief Record (Library of Congress Online Catalog) - Mozilla Firefox	3
<u>File Edit V</u> iew Hi <u>s</u> tory <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp	
😝 Beyond MARC: 🧐 Models for decis 🧐 Models for decis 👺 Brief Record (× 🛛 🔀 LC control no: 2 🖓	+
🗲 🎯 catalog.loc.gov/cgi-bin/Pwebrecon.cgi 🔤 🏠 🤜 😋 🔯 🔯 - Bing 🖉 🐥 🧍	3
The Library of Congress Authorities >> Go to Library of Congress Authorities	-
LIBRARY OF CONGRESS ONLINE CATALOG	
Help New Search Headings Titles Request Account Start Over	
Try the new version: Library of Congress Online Catalog	
DATABASE: Library of Congress Online Catalog YOU SEARCHED: Title Keyword = semantic web SEARCH RESULTS: Displaying 1 of 1.	E
Brief Record Subjects/Content Full Record MARC Tags	
International journal on Semantic Web and information systems.	
Relevance: 🧶 🗶 🗶 🧶	
LC control no.: 2004214933	
Type of material: Serial (Periodical Newspaper etc.)	
Main title: International journal on Semantic Web and information systems.	
Published/Created: Hershey, PA : Idea Group	
Publication history: Began with v. 1, no. 1 (Jan./Mar. 2005).	
Description: volumes ; 26 cm	
ISSN: 1552-6283	
Linking ISSN: 1552-6283	
CALL NUMBER: TK5105.88815.158 Set 1	-

It Doesn't Be A Result In Search Engines Results





The Reasons Are:

our data are not in Google Why? Google does not understand: Marc, ISBD, OAI-PMH, RDA, Z39.50, Onix

So, let's developing many Initiatives













<!DOCTYPE motd [<!!
<motd>
<!-- created: 2003-12-12-->
<sentence>Do not throw
out the <keep>baby</>with the
<refuse>dirty</>, <refuse>stinky</><refuse>bathwater </>.
</><!-- finish this later-->
</motd>
SGML





Finally, we did it !!!





BIBFRAME

In November 2012, the Library of Congress launched The Bibliographic Framework initiative,

the goal is providing a complete and rich bibliographic description format that integrates indexes with other search tools on the Web.



Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services

Library of Congress Washington, DC November 21, 2012



BIBFRAME Architecture





BIBFRAME Vocabulary

$\langle \rangle$	file:///Users/momen/PhD/Ph.D%20Thesis/Ph.D/مصححة20% الرسالة/Bibliographic%20Framewo	• ± +
لار الببليوجرافي	محرر الإط	موقع مبادرة الإطار الببليوجرافي »
		الصفحة الرئيسية / المحرر
+ + +	عناصر وصف العمل وفقا لقواعد وام - السلاسل والدوريات (RDA 19.2) منثن العل (RDA 62.2) منثن العل (RDA 62.3) النوان المثمل للعل (RDA 62.3) النوان البيل للعل (RDA 6.2) مثل العل (RDA 6.2) مثل العل	السفحة الرئيسية / المحرر فشاء تسجيلة لمصدر المنفردات النوت الموسيقية السلامل والدوريات المرائط الخرائط الأفلام ذات خاصية 35 ملم جالأفلام ذات خاصية 35 ملم
+	(RDA 7.3) التغطية الزمنية أو الجغرافية للمحتوى (RDA 7.3) التغطية الزمنية أو الجغرافية للمحتوى (RDA 7.3)	
+	(RDA 7.7) الجمهور المقصود (RDA 7.7) الجمهور المقصود (RDA 7.7)	
+	 الأشخاص، الأسر، الهيئات الأخرى المرتبطة بالعمل الأشخاص، الأسر، الهيئات الأخرى المرتبطة بالعمل الأسر (قائمة رووس موضوعات الأخرى ذات الصلة بالعمل الأسر (قائمة رووس موضوعات مكتبة الكونجرس) الهيئات الأشغاص المكان الموضوعات الأشخاص (RDA Chapter 23) موضوعات العملي بيانات تمنيف ديوي العشري بيانات تمنيف ديوي العشري الأعمال المرتبطة الأعمال المرتبطة الأعمال المرتبطة الأسرة المحتويات الأسرة المحتويات الأسرة المحتويات المحتويات المحتويات الأمر (RDA Chapter 25) الأمر (RDA Chapter 25) الأعمال المرتبطة الأعمال المرتبطة الأعمال المرتبطة الأمرة المحتويات المراح المحتويات الأمرة المحتويات المراح المحتويات الأمرة المحتو	

Now we have SW OPACS

a			0		2	-	
🐨 libris.kb.se/bib/5060570		G⊽C	0		م	+	
	HELP PÅ SVEN	SKA PREFI	ERENCE	S MY LIBRARI	ES CLEAR	HIST	DRY
IBRI3							
atart Extended search Index A-Ö Boolean	Subdatabases				Search	i histo	ry
farance angine	Sear	ch 🗌 🗠	libearie				0
			inprairie	20			
arch: difference engine ~ the difference engine							
and and check engine and and and and engine				fic.			_
2 of 21				► Search fur	ther		
	Overview	Details			200		
			0	Related tit	les		
The difference engine / William Gibson and	Bruce			- Search out	tside LIBRI:		
Sterling		LINN! THE PACE		Tisle to Co.	and a Dearth	3	0
Gibson, William, 1948- (author)	9971	MAN DIBNON	NICE .	Search:	ogle book		
Sterling, Bruce, 1954- (author)	201			* The differe	ence engi		
London : Vista, 1996		CN.		Extend you	ur search t	0:	
English 383 s.	28.0	THE		Google Soogle Bo	ok Search		
Book	E	IFFERENC	E	Google Sc	holar		
			-	Scirus			
			Q	- LibraryThi	ng		
			©	Wikipedia	about the a	author	
				 William Gil 	bson		
Subject headings							

The Semantic Web OPACs

🥑 William Gibson - Wikipedia,	the free encyclopedia - Mozilla Firefox	ARRENT CONTRACTOR STOCKED	
<u>File Edit View History Boo</u>	okmarks <u>T</u> ools <u>H</u> elp		
🖏 Google Translate	× IIBRIS - The difference engine /	W William Gibson - Wikipedia, the free e × +	and the second sec
(en.wikipedia.org/wik	i/William_Gibson	🔚 🏫 🔻 😋 🔯 🧕 📴 - Bing	P 🖡 🏠 🔇
	Article Talk	Read Edit View history Se	Create account & Log in
WIKIPEDIA The Free Encyclopedia	William Gibson		*
Main page Contents Featured content Current events Random article Donate to Wikipedia	For other people named William Gibson, se William Ford Gibson (born March 17, 1948) i novelist who has been called the "noir prophet" coined the term "cyberspace" in his short story popularized the concept in his debut novel, Nei Gibson created an iconography for the informat the 1990s. ^[29] He is also credited with predicting	ee William Gibson (disambiguation). s an American-Canadian speculative fiction of the cyberpunk subgenre. ^[2] Gibson "Burning Chrome" (1982) and later <i>irromancer</i> (1984). In envisaging cyberspace, ion age before the ubiquity of the Internet in the rise of reality television and with	William Gibson
 Interaction Help About Wikipedia Community portal Recent changes Contact page 	establishing the conceptual foundations for the video games and the World Wide Web. Having changed residence frequently with his fa ungainly teenager who often read science fictic private boarding school in Arizona, Gibson eval emigration to Canada in 1968, where he becan	rapid growth of virtual environments such as amily as a child, Gibson became a shy, n. After spending his adolescence at a led the draft during the Vietnam War by the immersed in the counterculture. After	
▶ Tools	settling in Vancouver he eventually became a fi	Ill-time writer. He retains dual citizenship. ^[4]	
Print/export	Gibson's early works are bleak, noir near-future computer networks on humans—a "combination	n of lowlife and high tech". ^[5] The short	
 Languages العربية Bân-lâm-gú Бепаруская (тарашкевіца) Български Вългарски 	stories were published in popular science fictio characters developed in these stories culminat garnered critical and commercial success, virtu Although much of Gibson's reputation has rem work has continued to evolve. After expanding complete the dystopic Sprawl trilogy, Gibson b	n magazines. The themes, settings and ed in his first novel, <i>Neuromancer</i> , which hally initiating the cyberpunk literary genre. ained associated with <i>Neuromancer</i> , his on <i>Neuromancer</i> with two more novels to ecame an important author of another	In his 60th birthday in Paris during a al interview for the French release of <i>ook Country</i> (March 17, 2008) William Ford Gibson March 17, 1948 (age 65) Conway, South Carolina, US





About 10,800 results (0.46 seconds)

The Unsinkable Molly Brown Cookbook: May Bennett and Caroline ...

https://www.amazon.com/Unsinkable-Molly-Brown-Cookbook/dp/B000FCH4CK **The Unsinkable Molly Brown Cookbook** [May Bennett and Caroline Bancroft Wills] on Amazon.com. *FREE* shipping on qualifying offers.

Unsinkable Molly Brown Cook Book: May Bennett Wills, Caroline ...

https://www.amazon.com/Unsinkable-Molly-Brown-Cook-Book/dp/080400305X **Unsinkable Molly Brown Cook Book** [May Bennett Wills, Caroline Bancroft] on Amazon.com. *FREE* shipping on qualifying offers. Short stories of Molly Brown and her times with prints. Chapters include: Come at five-ish The Miracle of the Fishes Come and Get It One Man's Meat is Another Man's Poison Chicken on the Run ...

The Unsinkable Molly Brown Cookbook 1989 paperback reprint: May ...

https://www.amazon.com/Unsinkable-Molly-Cookbook-paperback.../B005GQA60C **The Unsinkable Molly Brown Cookbook** 1989 paperback reprint [May Bennett Wills and Caroline Bancroft] on Amazon.com. *FREE* shipping on qualifying offers. cookbook.



< >

Let's Try Again!!!

Edgewear, Rubbing, And Back Has Some Staining All books shipped with trac.

the Unsinkable Molly Brown Cookbook - Google Search

The Unsinkable Molly Brown cookbook in SearchWorks catalog https://searchworks.stanford.edu/view/9734480 ▼

The Unsinkable Molly Brown cookbook. Responsibility: by May Bennett Wills and Car Imprint: Denver : Sage Books. [1966]: Develop description: 118 p. : ill., port. ; 22 cm.

Denver Public Library Catalog

https://catalog.denverlibrary.org/search/title.aspx?cn=32231 -

The unsinkable Molly Brown cookbook / by May Bennett Wills and Caroline Bancroft. May Bennett. Call Number: 641.5 W685un. Publisher, Date: Denver : Sage Books, 1966. 118 p. : ill., port. ; 22 cm. Subjects: Cooking. Colorado -- Imprints -- Denver -- 1966. Other Brown, Margaret Tobin ...

The Unsinkable Molly Brown cookbook / by May Bennett Wills a nla.gov.au/anbd.bib-an658314 -

The Unsinkable Molly Brown cookbook / by May Bennett Wills and Caroline Bancroft. May Bennett. Other Authors. Bancroft. Caroline. (ioint author.) Brown. Margaret (Tobin).



Now We Have Smart OPACS

Language -								
TA DENVER PUBLIC LIBRARY								
DPL Home C	Catalog S	Search -	My Ac	ccount -				
	DISPLAYIN	IG 1 OF 1						
	Title: Th	e unsinkable Mo	Ily Brown coo	kbook / by Bancroft				
	Author: Wi	lls, May Bennett.	s and ouronn					
1966	C II Number: 64	1.5 W685un						
	Pus inher, De	nver : Sage Book	s, 1966.					
	Date:		2					
	Subjects: Co	\mathbf{p}	z cm.					
	Subjects: Cooking. Colorado Imprints Denver 1966.							
	Other Author: Brown, Margaret Tobin, 1867-1932.							
Notes: Includes index.								
EXPAND ALL COLLAPSE ALL								
▼ Where Is It?								
Call Number	Shelf Location	Status	Circ Date	Туре				
Central Library (0 of 1 available)								



Library Catalogues I don't Think so! The Main Target of Bibliographic Control Tools is leeting the Users needs

By helping Users to:

- Searching,
- Find,
- Identify,
- Select,
- Acquire or Obtain Access,
- Navigate
- and Explore Information
- Resources



Few Years Later! Many Surveys has been Conducted, Conjirming that Usage Stats of Al Search Tools will take a Charge of Meeting Users Need

Users going to Al Search Tools Stats

- 37% of people in advertising and marketing.

- 35% of people in technology.
- 30% of people in consulting.
- 19% of people in accounting.
- 15% of people in healthcare.

Types of Content Where Consumers are Concerned with Artificial Intelligence Being Used



Generative Al Market Growth Projections



Year

According to Forbes

Despite having concerns about application of AI in various businesses and verticals, 65% of consumers have trust in the businesses which use AI technology.





More than quarter of people in the UK (26%) Moving to use Generative AI Search Tools,

According to new findings from Deloitte's 2023 Digital Consumer Trends research

Why Users are going to Machines (Generative AI) Tools rather than OPACs

So

Despite it maybe untrustful like Bibliographic tools?




THE BIBLIOGRAPHIC CONTROL FOCUS ON ORGANIZING PHASE





Clustering Sentiment Artificial Intelligence

To Meet Currently Users Needs Which Are Be...



Descriptive Needs \rightarrow "What has happened in the present and past?"



Diagnostic Needs → "Why it Happened?"



Predictive Needs \rightarrow "What could Happen in future?" (ML)



Prescriptive Needs \rightarrow "What should be Done?" (Gen. AI)

How Can Make UBC Meets Users Needs? Or Backing to the Main Question that has been Asked in IFLA WLIC 2014 !!!



Let's Exploring !!

As We had Known Previously,

the Human being was Responsible to Produce Inter and Control it ...





Is it good a Machine to be a Human Partner in Data Production?

But!

IDC Announcement



- The International Data Corporation Annual Report (IDC) In 2011 Indicated, The Data Volume that Generated have Reached to 1021 Trillion Exabytes Of Data..
- The Data Volume has Doubled About 9 times than before,
- These Numbers can be considered as Alarm, and as Declaration to Great Phenomenon, that will be





So,

What is Big Data

What is Big Data

"... Refers to Data that is so Large, Fast, and Unstructured or Complex that it's Difficult or Impossible to Process Using Classic Methods (as Bibliographic Control) and Techniques ..."

Size of Data



Internet Users in the world

1,728,794,416

Total number of Websites



Emails sent today



6,403,934,886 Google searches today



6,107,136 Blog posts written today



Tweets sent today



6,633,032,954

Videos viewed today on YouTube



77,615,781 Photos uploaded today on Instagram



Tumblr posts today



Facebook active users



Google+ active users





Big Data Features

VOLUME	VARIETY	VELOCITY	VERACITY	VALUE
The amount of data from myriad sources.	The types of data: structured, semi-structured, unstructured.	The speed at which big data is generated.	The degree to which big data can be trusted.	The business value of the data collected.
0000 0000	€ `} *	$\textcircled{\begin{tabular}{ c c c c } \hline \hline$		



Unstructured Data

Unstructured Data

• Unstructured data refers to data that lacks a predefined structure or organized format.

• This type of data does not fit neatly into traditional rows and columns like structured data found in databases.

• Unstructured data is typically more flexible in terms of content and format, making it challenging to manage and analyze using conventional data management tools and techniques.



Unstructured Data Types



Unstructured Data Size



A DAY IN DATA



From Three Resources !!!



The Common Sources of Big Data:

- Social Media.
- IoT Devices.
- Online Transactions.
- Digital Media.
- Log Files.
- Sensor Networks.
- Genomic Data.
- Scientific Research.
- Government and Public Data.
- Text and Documents.
- Financial Data.
- Geospatial Data.





Based on Three phases

- Knowledge Representation.
- Knowledge Understanding.
- Knowledge Generation.

1 st phase: Knowledge Representation

Knowledge representation is a fundamental concept in artificial intelligence (AI).
It refers to the process of capturing and structuring information in a way that can be understood, processed, and manipulated by computers or intelligent agents.
The goal of knowledge representation is to create a formal system or framework that allows computers to reason, make inferences, and draw conclusions from the stored information.

Representation Tools

• Semantic Networks.

• Frames:

• Ontologies:

• Neural Network Embeddings:

Semantic Networks

Semantic Networks: These represent knowledge as nodes (objects or concepts) connected by edges (relationships). They are easy to understand but can become complex for large knowledge bases.

Francis Like (BIBFRAME – RDF)

Frames organize knowledge into structured units, with attributes and slots that define properties and relationships of objects or concepts. They are used for representing structured information about objects or events.



Ontologies are formal models that define concepts, their relationships, and properties within a specific domain. They provide a common vocabulary for various AI systems to communicate and reason about shared knowledge.

Ontologies

Probabilistic Representations

In situations involving uncertainty, probabilistic models can be used to represent knowledge using probabilities and statistical methods.

Neural Network Endeddings

In more recent years, techniques like word embeddings and neural network-based models have been used to capture semantic relationships between words and concepts, enabling AI systems to understand and represent knowledge in a distributed and continuous manner.
2nd phase: Knowledge Understanding !

Knowledge Understanding Is the Machines Ability to Comprehend and Make Sense of Information.



It is the process of extracting meaning from data and using that meaning to answer questions, solve problems, and make decisions.

Knowledge Understanding

Knowledge understanding is a complex process that involves multiple cognitive skills, such as:

- Machine Attention: The ability to focus on relevant information and ignore irrelevant information.
- Machine Memory: The ability to store and retrieve information.
- **Comprehension:** The ability to understand the meaning of information.
- **Inference:** The ability to draw conclusions from information.
- **Evaluation:** The ability to assess the quality of information.





Q	The Result !!!	
Barack Hussein Obama II is an American politician, law United States from 2009 to 2017. A member of the Dem president of the United States.	yer , and author who served as the 44th president of the ocratic Party , Obama was the first African-American	
Entities (Out of the Box Model) Edit Text		
Extraction Classification Linguisti	cs Custom	
Entities Keywords Concepts Relations		
Name	Туре	Confidence
Obama	Person	0.99
lawyer	JobTitle	0.91
author	JobTitle	0.70
president	JobTitle	0.97
United States	Location	0.81
office offices		

С

Based on Watson Discovery

 \mathcal{O}

\equiv IBM Cloud		Q	Catalog	Docs	Support	Manage \vee	Po
්ර Watson	Select Watson services below to get	started with yo	our proj	ect			
Starter Kits	Simply pick one or more APIs from the catalog, click Add Serv	vices, and get your API cre	dentials.				
Watson Services	¢	Q					
Browse Services Existing Services Developer Resources ~	Watson Assistant Watson Assistant lets you build conversational interfaces into any application, device, or channel.	Watson Discovery Add a cognitive search	and content a	inalytics e	ngine to app	lications.	
Apps	Lite - IBM	Lite - IBM					
	t Bal	¢1					
	Natural Language Understanding Analyze text to extract meta-data from content such as concepts, entities, emotion, relations, sentiment and more.	Speech to Text Low-latency, streaming	transcription	R			
	Lite - IBM	Lite - IBM					
		2_2					
	Text to Speech	Natural Language Clas	sifier				
	Synthesizes natural-sounding speech from text.	Natural Language Class	ifier uses adv	anced na	tural languag	e processing	
						113	

DETECTION !!!

 \mathbf{O}

6

Q

 \bigcirc

C

	information science 🙇			Configure data 🚳 🔤
品	Overview Errors and warnings (0)	Search settings		
∠ ∠	1 document	 ♥ 0 documents failed Created on View details Last updated 	3/18/2021 8:18:59 am EDT 3/18/2021 8:18:59 am EDT	documents T
	Identified 1 field from your data	Added 4 enrichments to your data		Now you're ready to query !
	Text Need to identify more fields? Add fields	Entity Extraction AI (1) Amazon (1) Apple (1) Bell (1) Bell Tower (1)	Sentiment Analysis 100% 0% 0% positive neutral negative	Entities of type Company which have positive sentiment Run
		Concept Tagging	Category Classification	Most common entity types and their top entities Run
		Academic library (1) Edward Feigenbaum (1) Intelligence (1) Learning (1)	education → homework and study	Documents that contain Academic library, but not

≡ ІВМ V	Vatson Discovery		Instance: ma:	ster 🗸 🗄
	dentify fields Manage fields	Enrich fields		Apply changes to collection ③
	10-1108_ILS-02-2018-0011.pdf 1	<text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text>		Field labels Identify document elements using the labels below. + Create new Upgrade answer author footer header question subtitle table_of_contents text title upgrade upgrade upgrade table

	TEST	ď
	IBM Watson Discovery information science / Build queries	Instance: master
Q	Build a query using one or more of these components. Learn more.	ery
	Search for documents Use natural language Use the Discovery Query Language When did the artificial intelligence appear? + Include analysis of your results + Filter which documents you query	Passages "cial intelligence lab at the University of Rhode Island cial intelligence to in" "cial intelligence lab at the University of Rhode Island cial intelligence to in" "cial intelligence lab at the University of Rhode Island cial intelligence to in" "Arti fi cial intelligence arrives in the library" "Abstract Purpose – The purpose of this column is to examine arti fi cial intelligence (AI) and its potential relationship to the library. Design/methodology/approach – Literature review and commentary on this topic have been addressed by professionals, researchers and practitioners." "Keywords Services, Libraries, Librarians, Arti fi cial intelligence, Virtual Paper type Viewpoint Introduction "Arti fi cial Intelligence (AI), " at er m fi ret of fi cially coined in the 1950s, but perhaps imagined by some at an even earlier point in time"
2	> More options	Results Showing 1 of 1 matching documents Arti fi cial intelligence arrives in the library Sentiment positive
	Run query-	Text "/blogs/digital-library-blog/2017/11/arti fi cial- intelligendkan@jibrary-blog/2017/11/arti fi cial- intelligendkan@jibrary-blog/2017/11/arti fi cial- (2001), " The age of intelligent machines/knowledge processing Officin fi file karvars to knowledge eswersict/OWS. Kurzwell Accelerated Intelligence. available at:

The Results !!!

O

6

Q

 \bigcirc

Sur			Trai	in Watson to imor	ove resul	te
501	mmary JSON					
-	100					
-				instance (42 of c	000 10	in the second se
Quei	yokr mups.//api	.us-south.discovery.wats	on.cloud.tom.com/	instances/45060	1098-16	
HC					•	
Pas	sades					
	.54205					
"cia	l intelligence lab at	the University of Rhode Is	land cial intelligence	e to in"		
-cial	l intelligence lab at	the University of Rhode Is	land cial intelligence	e to in"		
Art	i fi cial intelligence	arrives in the library"				
Abs	stract Purpose – Th	e purpose of this column i	ia to examine arti fi c	ial intelligence (A	1)	
and	its potential relation	onship to the library. Desig	n/methodology/app	roach – Literature		
and	ew and commentar practitioners."	ry on this topic have been i	addressed by protes	sionals, researche	ers	
	ovorde Services, Li	braries Librarians Artific	ial intelligence Virtu	al Paper type		
Viev	/words Services, Li vpoint Introduction	braries, Librarians, Arti fi o n " Arti fi cial Intelligence (ial intelligence, Virtu AI), " at e r m fi rat o	al Paper type f fi cially coined in	the	
Viev 195	/words Services, Li vpoint Introduction 0s, but perhaps im	braries, Librarians, Arti fi o n " Arti fi cial Intelligence (agined by some at an ever	ial intelligence, Virtu AD, " at e r m fi rst o n earlier point in time	ual Paper type f fi cially coined in e ^e	the	
Viev 195	words Services, Li vpoint Introductior Os, but perhaps im	braries, Librarians, Arti fi o n " Arti fi cial Intelligence (agined by some at an ever	ial intelligence, Virtu AI), " at e r m fi rat o 1 earlier point in time	ual Paper type f fi cially coined in *	the	4
Viev 195 Res	words Services, Li vpoint Introduction Os, but perhaps im sults	braries, Librarians, Arti fi o n " Arti fi oial Intelligence (agined by some at an ever	ial intelligence, Virtu AI), " at e r m fi rst o r earlier point in time	ual Paper type f fi cially coined in *	i the	
Viev 195 Res	words Services, Li vpoint Introduction Os, but perhaps im sults wing 1 of 1 matchin	braries, Librarians, Arti fi o n " Arti fi cial Intelligence (agined by some at an ever	ial intelligence, Virtu AI), " at e r m fi rst o 1 earlier point in time	ual Paper type f fi cially coined in *) the	
Viev 195 Res	words Services, Li vpoint Introductior Os, but perhaps im sults wing 1 of 1 matchin	braries, Librarians, Arti fi o n " Arti fi oial Intelligence (agined by some at an ever ng documents	ial intelligence, Virtu AI), " at e r m fi rst o 1 earlier point in time	ual Paper type f fi cially coined in *	the	
Res Sho	words Services, Li vpoint Introduction Os, but perhaps im sults wing 1 of 1 matchin Arti fi cial intellig	braries, Librarians, Arti fi o n " Arti fi oial Intelligence (agined by some at an ever ng documents (ence arrives in the library	iat intelligence, Virtu AD, " at e r m fi rst o n earlier point in time	ial Paper type f fi cially coined in *	i the	
Res Sho	words Services, Li vpoint Introduction Os, but perhaps im sults wing 1 of 1 matchin Arti fi cial intellig Sentiment	braries, Librarians, Arti fi o n " Arti fi oial Intelligence (agined by some at an even ng documents (ence arrives in the library positive	ial intelligence, Virtu AI), " at e r m fi rat o 1 earlier point in time	ual Paper type f fi cially coined in ="	the	
Res Sho	words Services, Li vpoint Introduction Os, but perhaps im sults wing 1 of 1 matchin Arti fi cial intellig Sentiment Text	braries, Librarians, Arti fi o n " Arti fi olal Intelligence (agined by some at an even ng documents (ence arrives in the library positive	ial intelligence, Virtu AI), " at e r m fi rat o i earlier point in time / / brary-blog/2017/11	ial Paper type f fi cially coined in * *	, the	
Res Sho	words Services, Li vpoint Introduction Os, but perhaps im sults wing 1 of 1 matchin Arti fi cial intellig Sentiment Text	braries, Librarians, Arti fi o " Arti fi olal Intelligence (agined by some at an ever ng documents gence arrives in the library positive "/blogs/digital-li intelligenceand jui	iat intelligence, Virtu AI), " at e r m fi rat o r earlier point in time brary-blog/2017/11 prary-blog/2017/11	Jai Paper type f fi cially coined in " " Arti fi cial- d Fieigenbaum, E.	the	

3rd phase:

Knowledge Generation.

Knowledge Generation

Generative knowledge is a type of artificial intelligence (AI) that can automatically create new knowledge from existing data.



How Generate Knowledge?

- There are several different ways to implement generative knowledge, including:
- Machine learning: Machine learning algorithms can be used to learn from existing data and generate new knowledge.
- Natural language processing: can be used to extract information from text and code, and then generate new knowledge in a human-readable format.
- Graph neural networks: can be used to model the relationships between different pieces of knowledge, and then generate new knowledge that is consistent with these relationships.





Supervised learning

In supervised learning, machine learning algorithms are trained on a set of labeled data. This data includes both the input data and the desired output data. The machine learning algorithm learns to map the input data to the desired output data. Once the machine learning algorithm has been trained, it can be used to generate new knowledge by predicting the output data for new input data.



Machine Learneng Can generate Knowledge like This



Unsupervised Machine Learning

In unsupervised learning, machine learning algorithms are trained on a set of unlabeled data. This data does not include any desired output data. The machine learning algorithm learns to identify patterns in the data. Once the machine learning algorithm has been trained, it can be used to generate new knowledge by predicting the relationships between different pieces of data.



Output

Ex: Unsupervised Machine Learning

Case Study – Book Cover Classification, Analysis, and Recognition

In this chapter, we will work on classifying an image cover into one of 30 categories. We will start by finding a publicly available book cover dataset, continue with analyzing the dataset, and then work on building a custom classifier using a combination of features extracted from a pretrained neural network and custom architecture. By the end of this chapter, you will also see how you can expose your solution to a REST web service.

The code will first output shapes for the images that are being processed. As you can see, the shapes are fixed in height but are different in width:

(Any[(500, 381), (500, 333), (500, 324), (500, 360), (500, 339)]...)

Finally, you will see a preview of the images. The books are very different—from the **National Geographic** to **French Grammar**:



Now that we have our first understanding of the images dataset, we are ready to have a quick look at categories.

Differences between ML Sup & Unsup



Reinforcement learning

In reinforcement learning, machine learning algorithms are trained to take actions in an environment in order to maximize a reward. The machine learning algorithm learns to associate certain actions with certain rewards. Once the machine learning algorithm has been trained, it can be used to generate new knowledge by exploring the environment and finding new ways to maximize the reward.





NLP

Natural language processing (NLP) is a branch of computer science and artificial intelligence that deals with the interaction between computers and human (natural) languages. It is concerned with the understanding, interpretation, and generation of human language.



NLP for Documents



By Extracting:

- Entites.
- Keywords.
- Concepts.
- Other LS.

NLP for Resouces Indexing





The Results will be Like This



Neural Networks



Nemal herborks

A neural network is a type of AI algorithms that is inspired by the human brain. It is a collection of interconnected nodes, or artificial neurons, that learn to perform a task by analyzing data. Neural networks are typically used for tasks that are difficult or impossible for traditional machine learning algorithms to solve, such as image recognition, natural language processing, and speech recognition.

- Here is a simple overview of how a neural network works:
 - The neural network is presented with a set of data.
 - The data is passed through the network, layer by layer.
 - At each layer, the nodes in the layer process the data and pass it on to the next layer.
 - The final layer of the network outputs a prediction or classification.

Multiple Hidden Layers



What's Going on Inside It?

To Make an Artificial Neural Network, We Need to Use the Most Universal Language Mathematics.





By Convolution Neural Networks



Ajter all these, How Can Control These Generated Data? 3 11010 10101010 0

Backing to the Main Question that has been Asked in IFLA WLIC 2014!!!

Or

Ų.







Bibliographic Control Still and will be Remains

• Bibliographic control can play a role in generative AI in several ways:

• It can help to identify and select relevant information resources. Generative AI systems need access to a large corpus of text in order to generate new text that is both coherent and meaningful. Bibliographic control can help to identify and select relevant information resources, such as books, articles, and websites.

It can help to describe and organize information resources. Bibliographic control systems provide a standardized way to describe information resources, such as their title, author, publication date, and subject matter. This can be helpful for generative AI systems to understand the content of information resources and to generate new text that is consistent with the style and tone of the resources.

It can help to track and manage information resources. Bibliographic control systems can help to track and manage information resources, such as their current location, ownership, and circulation status. This can be helpful for generative AI systems to ensure that they are using the most up-to-date information resources and to prevent plagiarism.



- Here are some specific examples of how bibliographic control can be used in generative AI:
- A generative AI system could be used to write a news article about a recent scientific discovery.
- The system could use bibliographic control to identify and select relevant scientific papers that have been published on the discovery.
- The system could then use the information from these papers to generate a new article that is both accurate and informative.
- A generative AI system could be used to create a new poem.
- The system could use bibliographic control to identify and select relevant poems from different cultures and time periods.
- The system could then use the information from these poems to generate a new poem that is both creative and original.
- A generative AI system could be used to write a new business proposal. The system could use bibliographic control to identify and select relevant business reports and case studies.
- The system could then use the information from these resources to generate a new proposal that is both persuasive and feasible.


becafne

The Digital Tide Won't Knock UBC Out !!!



But!!! To Achieve that we need review our Bibliographic Tools to to include the jollowing:



Finally, We Can Say

"Now We Should Creating Bibliographic Control not for Human Users, but for Machine, Generative AI, and Smart Agents"



Based on Bibliographic Control, the Generative AI Will Do ...

0

 \bigcirc



In the 2010 World Cup of Soccer, the team from the <u>Netherlands</u> distinguished themselves well, losing to <u>Spain</u> 1-0 in the Final. Early in the second half, <u>Dutch</u> striker <u>Arjen Robben</u> almost changed the tide of the game on a breakaway, only to have the ball deflected by <u>Spanish</u> keeper, <u>Iker Casillas</u>. Near the end of regulation time, winger <u>Andres Iniesta</u> scored, winning <u>Spain</u> the World Cup.

Name	Position	Country
Arjen Robben	Striker	Netherlands
Iker Casillas	Goalkeeper	Spain
Andres Iniesta	Winger	Spain



WolframAlpha computational intelligence.

6 oictober 1973					
ATURAL LANGUAGE	MATH INPUT	IIII EXTENDED KEYBOARD	EXAMPLES	6 🛨 UPLOAD	🔀 RANDOM
Input interpretation					
Saturday, October 6, 19	973				
Date formats		More formats/calendars			
06/10/1973 (day/month	n/year)				
9 Ramadan, 1393 (unti	l sunset)				
Time difference from today	y (Sunday, August 6, 2023)				
49 years 10 months ag	0				
2600 weeks 1 day ago					
18201 days ago					
49.83 years ago					
Time in 1973		🔍 Enlarge 🛛 🕻	🛃 Data 🍄 C	Customize A	Plain Text
279 th day		▶			
40 th week					
Observance for October 6,	1973 (Egypt)				
Armed Forces Day (Eg	ypt)				
Events on October 6, 1973				Show and	niversaries
start of Yom Kippur Wa	ır				
birth of Ioan Gruffudd					
birth of Jeff B. Davis (a	ctor, etc.)				
Daylight information for Oc	tober 6, 1973 in Cairo				More
sunrise	5:51 am EET				
sunset	5:34 pm EET				

Smart OCR









а

b

συνδημοντας αθηνησιν τον αυτον ορκον. τον δε πολεμον τον προς αλεξανδρον τον μη εξειναι καταλυσασθαι τους θετταλοις ανευ αθηναι ων τοις αθηναιοις απο του αρχοντος και του κοινου του θετταλων. επαινεσαι δε αγελαον τον αρχοντα περι και περι των θετταλων οτι ευ και προθυμώς επιμεμελησθαι περι ων αυτοις η πολις επηγγειλατο επαινεσαι δε και τος πρεσβεις των θετταλων τος ηκοντας και καλεσαι αυτος επι ξενια εις το πρυτανειον εις αυριον. την δε στηλην την προς αλεξανδρον καθελειν τος ταμιας της θεο και περι της συμμαχιας. τοις δε πρεσβεσι δοναι τον ταμιαν του δημο εις εφοδια Ο δραχμας εκαστωι. την δε συμμαχιαν την δε αναγραψαι τον γραμματεα της βολης εν στηληι λιθινηι και στησαι εν ακροπολεί εις δε την αναγραφην της στηλης δοναί τον ταμιαν το δημο Ο δραχμας. ειναι δε και θεαιτητον τον ερχιεα ως λεγοντα αριστα και πραττοντα ο τι αν δυνηται αγαθον τω δημωί

©nature

Detecting Text, Images..etc



quid of Youns may more que en ten miller butenten en enunstatie - er poeien of our day way moder- or tobs cat be doublet que per finances to employie plant the LOUND CONTINUE OF LA DOIS the bob tenents be fin emutit-Car Laboure her chann lace EA CORPORATE CRASHE BACE one sony qualitat militan 000 to basing an encountry of their the descent sill a concernance lie chile elinesge-er ison fothens Ross operation on the sol of the In chanadianty bull-quar now or fourt at bacas - Besterr ICan Tater & distantianty fast or fit. companying has been poor or tansatar er bunna fer fre the or langene part- under a ter be term side of h builting this four creations out the Rows, contain growthing of pr dunnet ?ts ont quest goest on encode lie and comment 00.0,07-Commenter bue demotive

Comments brief decision priver all a court on your mance dust in polification costs of commentant in user assesses fails to points space demote the assessment of demote the assessment of demote the assessment of the assessment of the costs provide courts failed out matched courtoof

man of the



to do national or

OCANTC Glamos

tor fait country and of Assess In diam the dissources of the badance of profit for the sector will also bennetfele que elas effere bunesfile puse que surre mere ne pust cibe sues tomo er nono en alembe mon-outside -or feed manue these quark wy former manue be. Linconstant - Linesson in a most commissionly distant quat premil behave que al guilter of four Asper purcha protection in terms and the protection of the process of the standard stand of first stoking print is south plan mongser jum or å han grittert companymet of distantices of the territor is de demotification (hotsee con-ticelle, frighte from hysterause is for clas tour mannesd? que il cum cuin falle-s fers suder deudere is my mit arour is on or is calastery unio is uner con falastery

to our brand and the base to contail can que si ne for top TOXAGED gas for the fore first concernent links material or of the spin over all because or of the spin over all because or of controls for four-or of terms ten torio per vil consister e ta sue ecolometeri fo suscenti comme uno que ta la fazi ciporante uno que ta la fazi ciporante en fazi de la fazither que the toget of Balls to fmultere bob Counter-er ler ישמונהלם שבול לגיל קרארוו crue ic control pric net cruste quar set forta oc quar me burn-million cit mensionet eccolie monaucie-fibultan distantion_shee dension" qua unus consos dension" Dere fine st vere desates tes Cunton ca befor tend te trass manudad qu'ils fa La borndile chartes k dynambler elbertrope-some fante teadlows selle da pourti Deca-randzió ch se fennani lo OLCO IN ALC CH. AND ALL R M. ten fambon mut-

These of traditionates transt-ore boarding or the network of the traditional network of the traditional parts of the sport data before the traditional traditional for sports the sport and the sports the sport data before the traditional traditional for the traditional traditional for the traditional traditional for the traditional traditional for the traditional traditional traditional the traditional traditional traditional the traditional tradity



Generate Images Doesn't Exist based on Metadata.





✓ 5s cc

Executing (7s) <cell line: 5> > decorate_context() > _call_impl() > forward() > _call_impl() > _call_i



So, We HaberTopertidifity; Fintegretion Machines to..



The Final Advice



The Faster We Adapt, The Faster We Escape Extinction

The Half of Science, Is Organizing !!!

Thank you for Listening

