A brief history of data and databases

Technological School Instituto Técnico Centrals

ETITC - 2024 - 2

Slides created and hosted by University of Rochester History Site

Record Keeping – How long?

Xerxes I inscription at Van

Al-Hasakah, 3300–3100 BC, Uruk culture

The Kish tablet, 3500 BC, Kish period

Jemdet Nasr period, c. 3100–2900 BC

Sumerian, 3300 B.C.

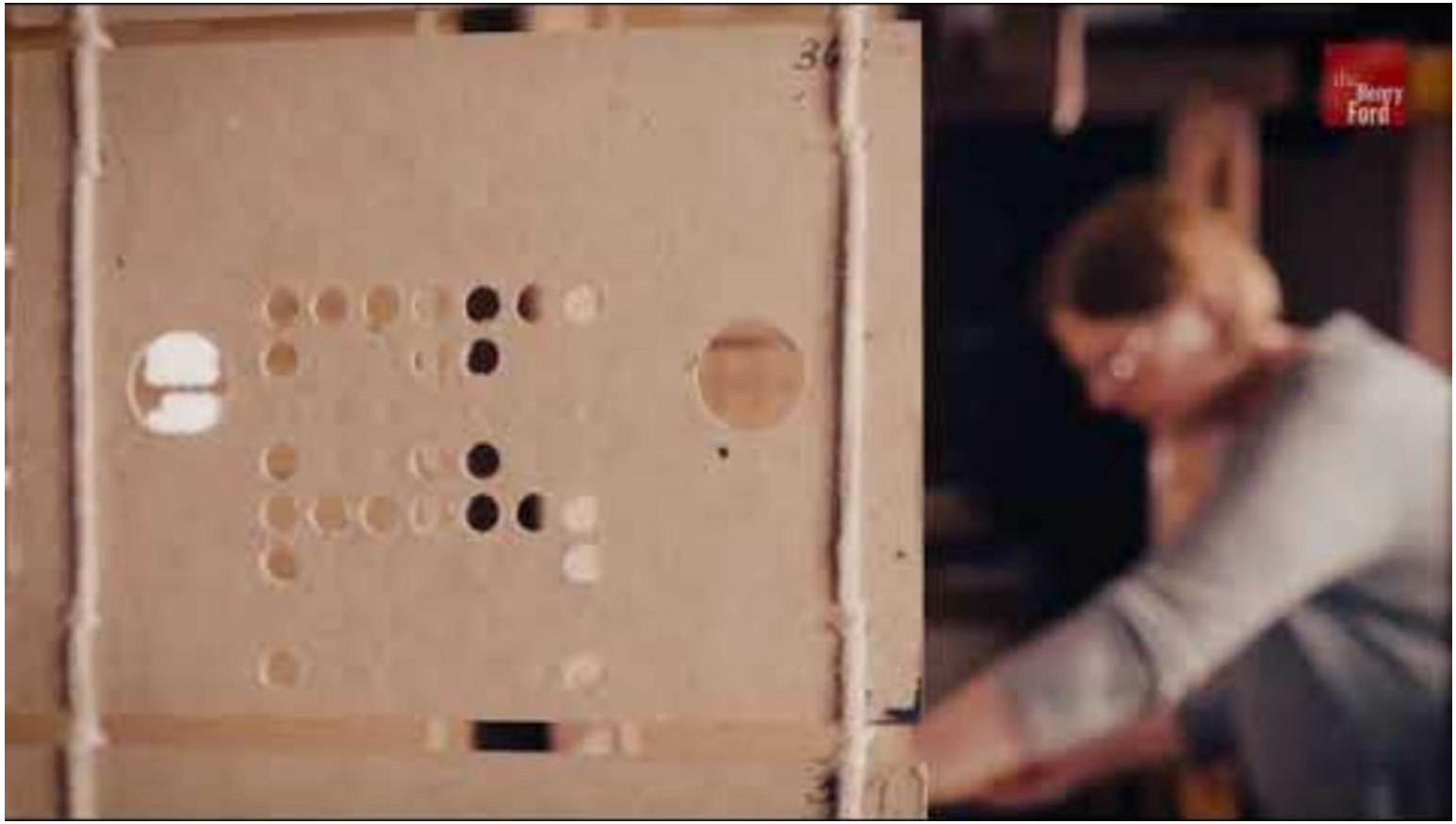
https://en.wikipedia.org/wiki/Cuneiform

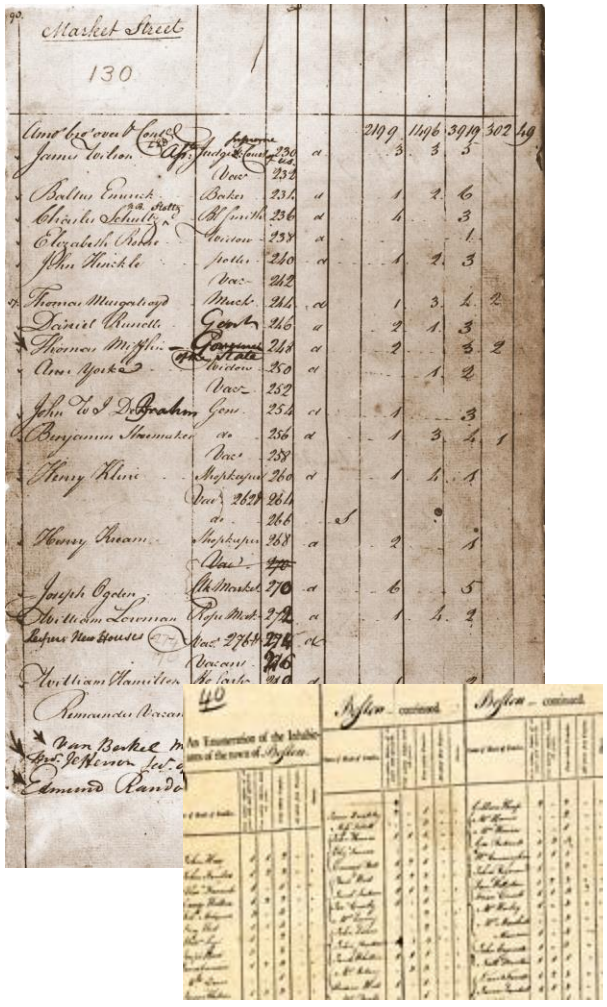
Why?

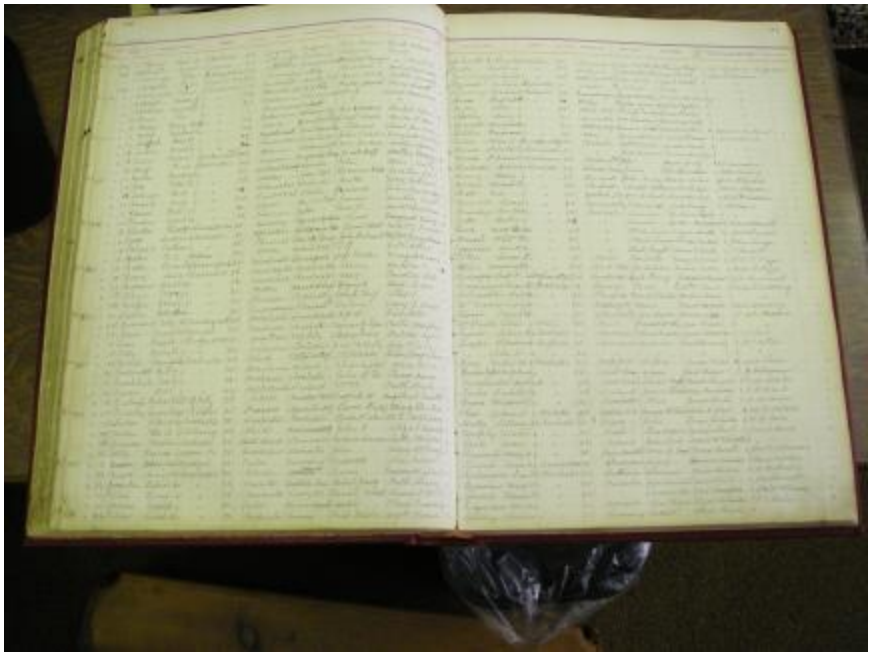
- We use records to never forget.
- We use records to measure “stuff”.
- And most of these records are not digital.

Jacquard Loom (1804)

Jacquard Loom

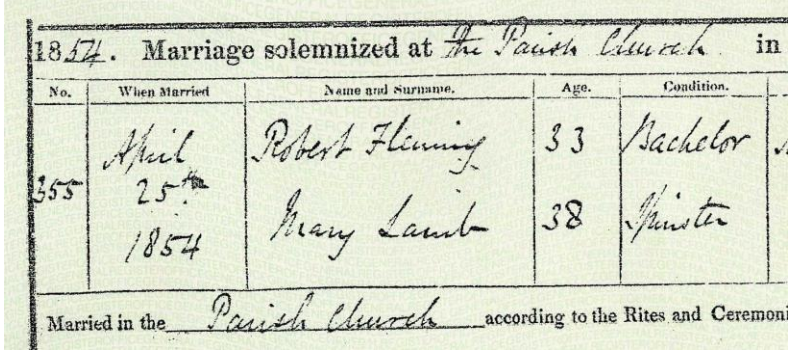


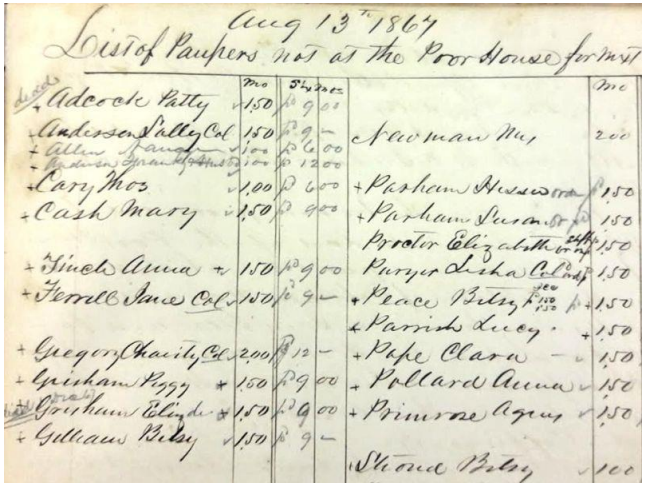




Marriage book, Rochester, NY, chronological filing of marriage licenses.

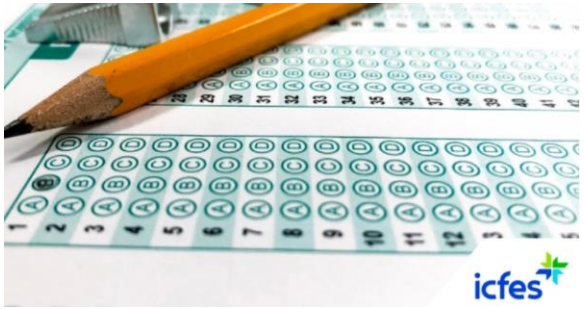
1790 US census





Other non-electronic records

- SS cards – 35 million hand typed between 1937-1938
- Motor vehicle licenses and registrations
- Financial records for companies
- School records



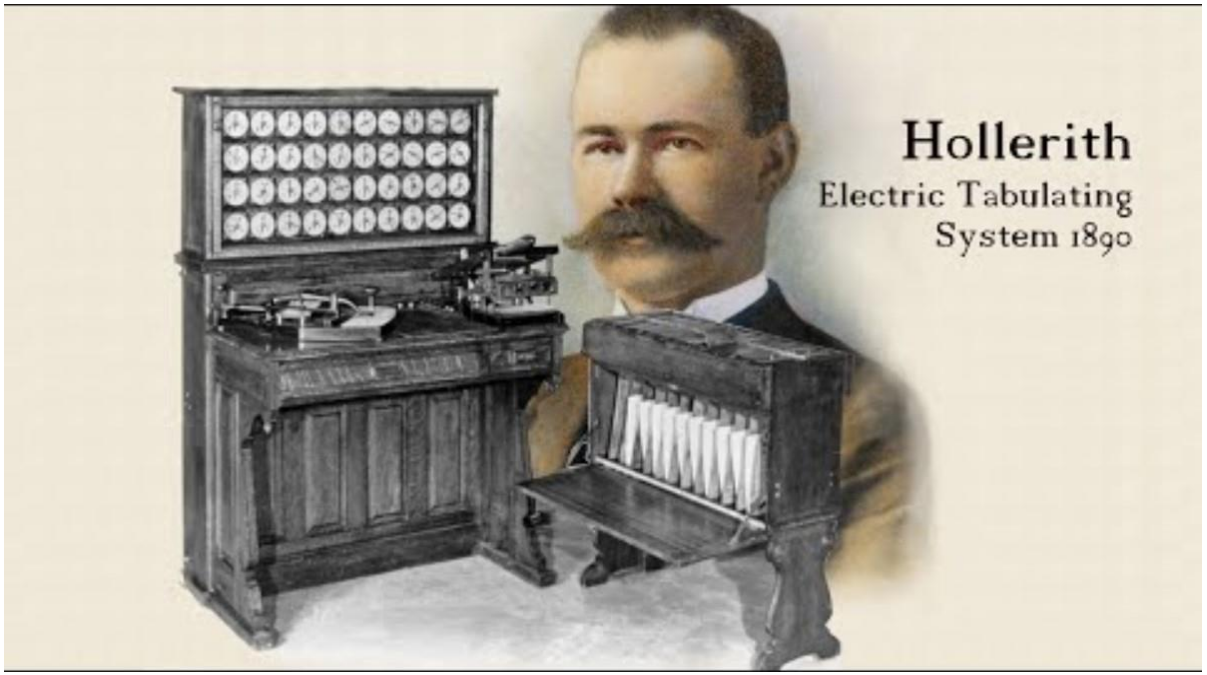
Card Catalogs – An ingeneous indexing system

•It didn’t indicate whether the book was available, just where it should be found.(example cards)
•Creating the cards required the expertise of librarians.

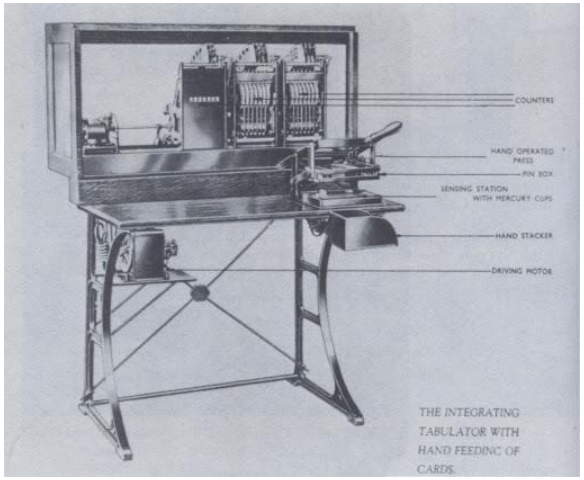
Problem – The 1890 census

- Enter Herman Hollerith.

Hollerith's device

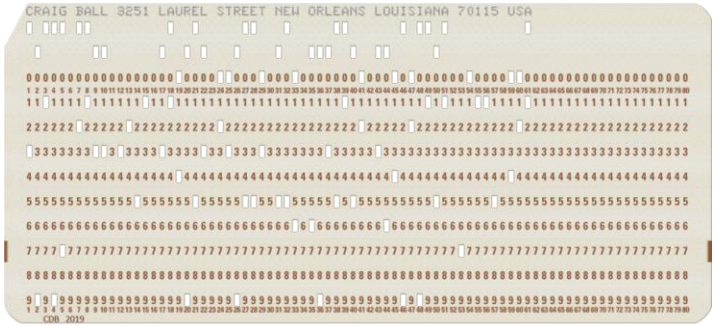


Integrating machine

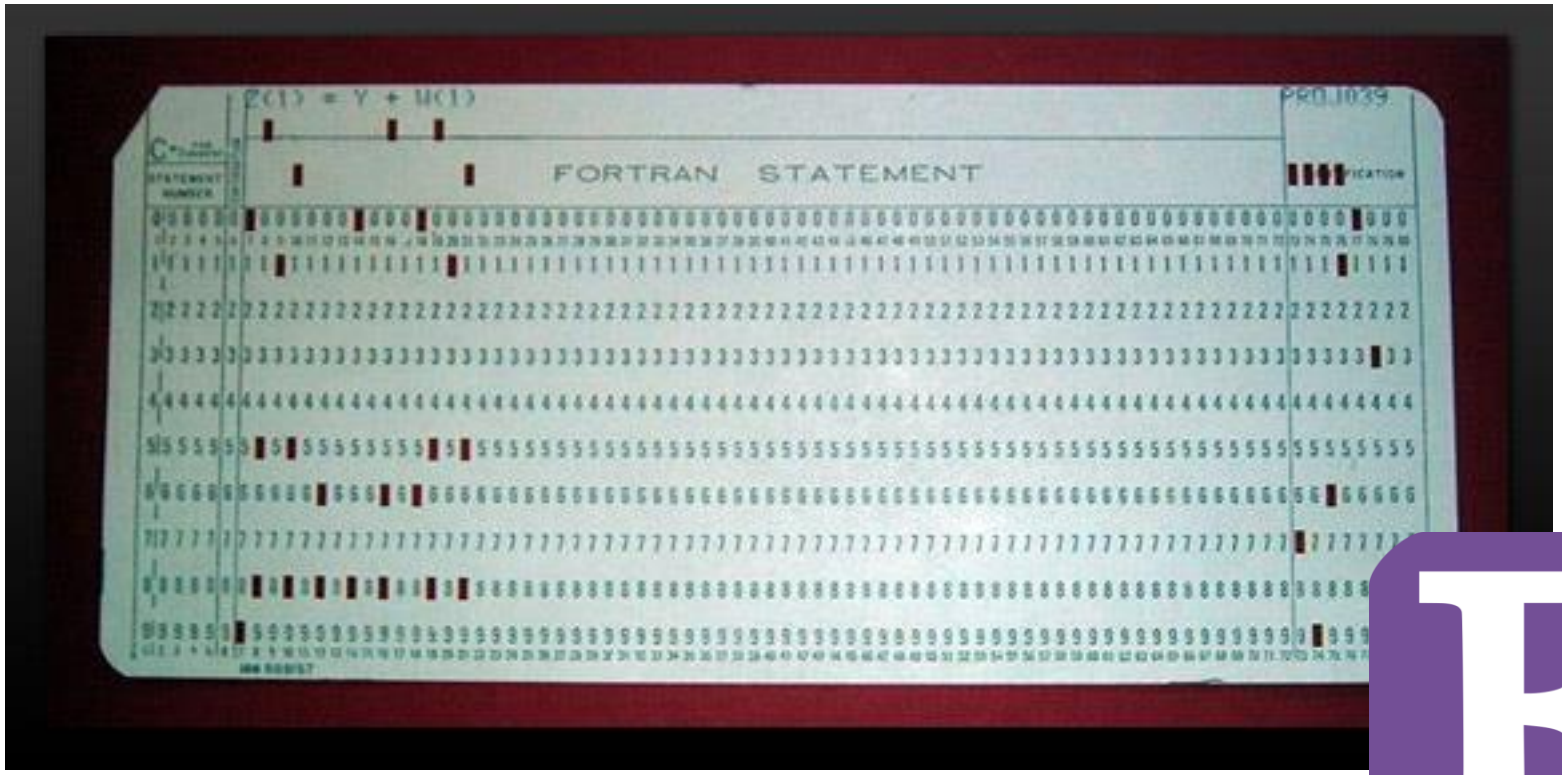


pantograph

Hollerith card



First computers



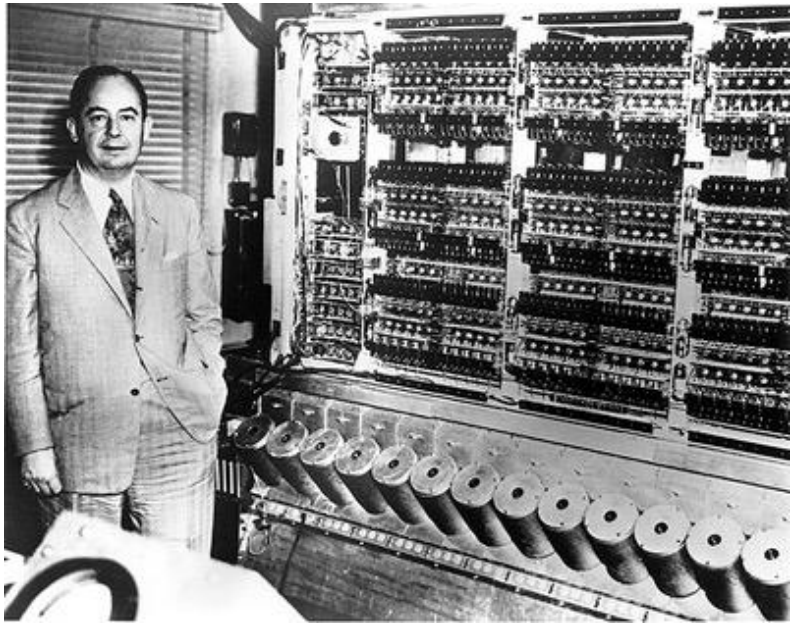
The program, the data, the JCL – all done with punchcards

First computers

1964 IBM 029 Keypunch Card Punching Demonstration

Electronic files – Early computing 1950s

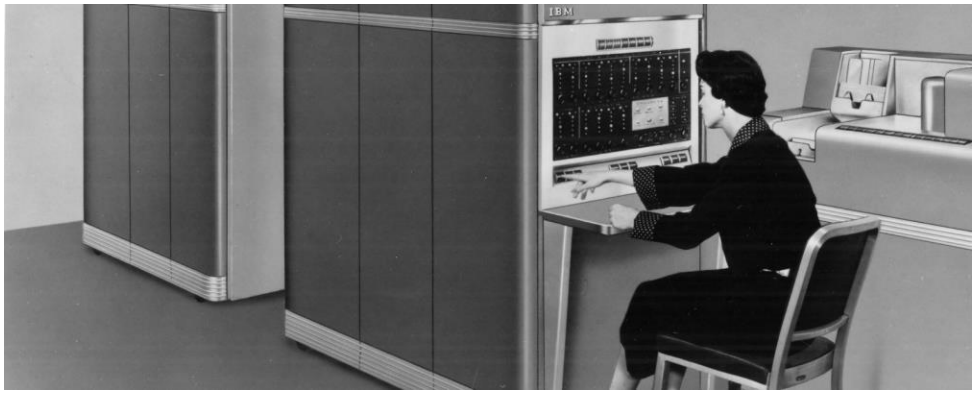
Universal Automatic Computer (UNIVAC) 1951



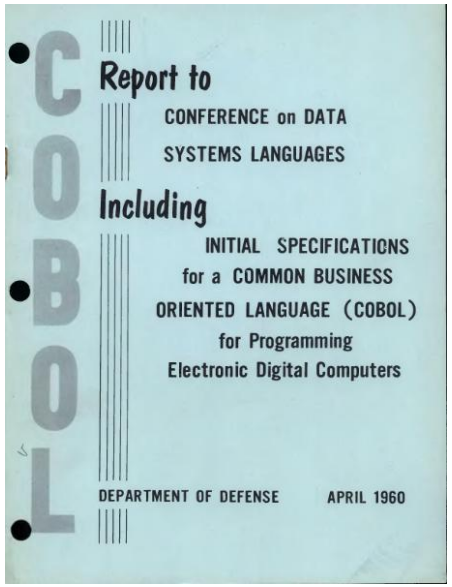
von Neumann, precursor innegable de la físicamoderna, y la EDVAC (Electronic Discrete Variable Automatic Computer) (1952-1957)

Electronic files – Early computing 1950s

The IBM 650 (1954)



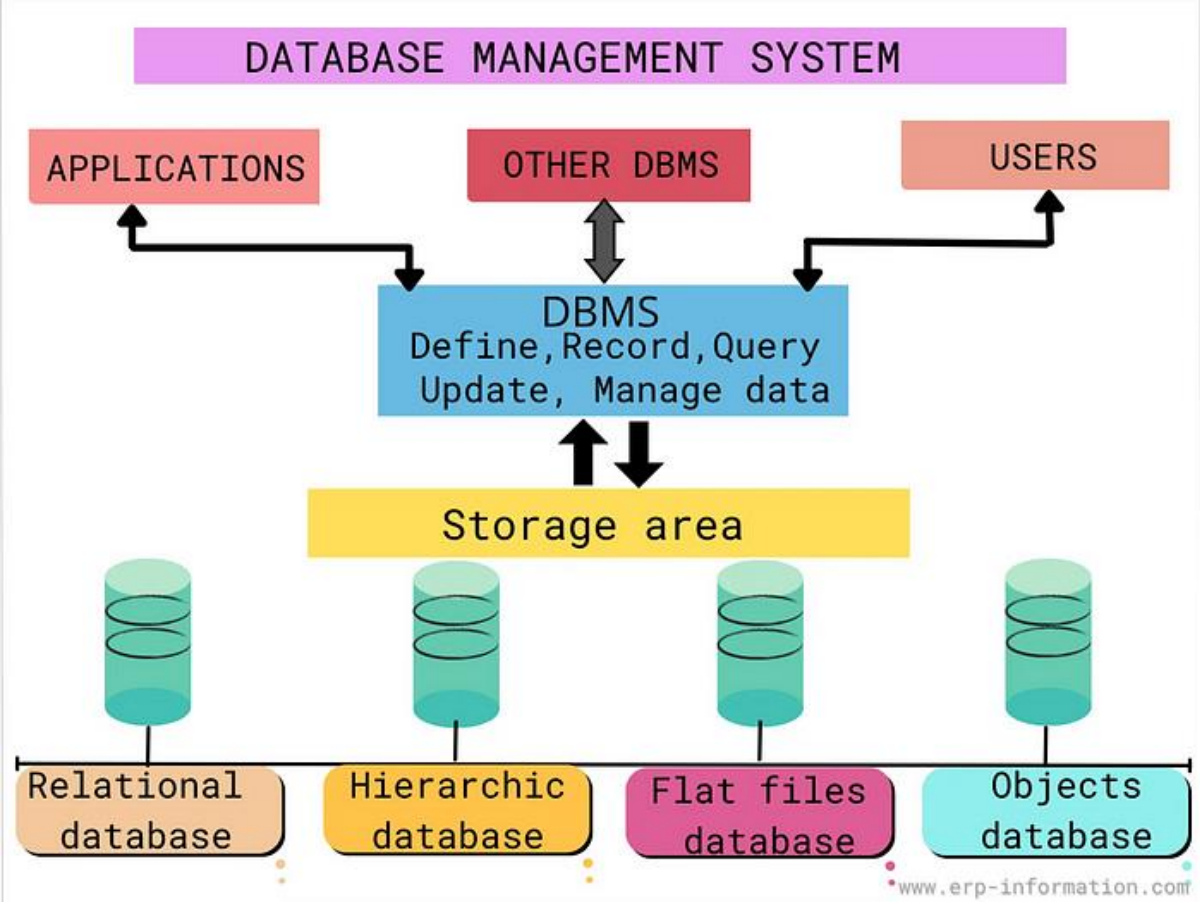
COBOL (Common Business Oriented Language) (1959)
CODASYL



The IBM 700 Series (1953)

Enter the database – Early 1960s

- Objects in a database can be related to one another.
- Hierarchical – One record leads to the related record. (Like a tree)
- Network – Allowed for multiple relationships (like a network)
- The databases used pointers to relate one record to another.



Electronic files – Early computing 1960s

CODASYL (Conference on Data Systems Languages) (1959)

ASCII (American Standard Code for Information Interchange) (1963)

$b_7b_6b_5$ → / $b_4b_3b_2b_1$ ↓ (Bits / Row)	000	001	010	011	100	101	110	111
	0	1	2	3	4	5	6	7
0000	0	NUL	DLE	SP	0	@	P	`	p
0001	1	SOH	DC1	!	1	A	Q	a	q
0010	2	STX	DC2	"	2	B	R	b	r
0011	3	ETX	DC3	#	3	C	S	c	s
0100	4	EOT	DC4	$	4	D	T	d	t
0101	5	ENQ	NAK	%	5	E	U	e	u
0110	6	ACK	SYN	&	6	F	V	f	v
0111	7	BEL	ETB	'	7	G	W	g	w
1000	8	BS	CAN	(	8	H	X	h	x
1001	9	HT	EM	)	9	I	Y	i	y
1010	10	LF	SUB	*	:	J	Z	j	z
1011	11	VT	ESC	+	;	K	[	k	{
1100	12	FF	FS	,	<	L	\	l	|
1101	13	CR	GS	—	=	M	]	m	}
1110	14	SO	RS	.	>	N	^	n	~
1111	15	SI	US	/	?	O	_	o	DEL

Electronic files – Early computing 1960s, Charles Bachman

Integrated data store (IDS) – Dow Chemical
CASE products (Computer Aided Engineering)

Electronic files – Early computing 1960s

Information Management System (IMS)
Hierarchical data model.
Hard drives

"My professor brought in a 10MB hard disk from the 1960's" xD

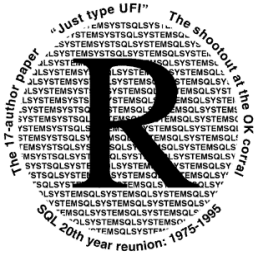
Some Issues

- While an improvement over file-based systems, these systems required knowledge of the structures to use them. No built-in search mechanism.
- Very few users understood the structures, access limited to an elite few.
- Queries were complex. Took time to get new information and expensive programmer time to produce.

Enter the relational DBMS
1970, Edgar Codd

Relational DBMS
- Mathematician at IBM
- Based on Relational Calculus and set theory

U of Michigam
- MicroDBMS



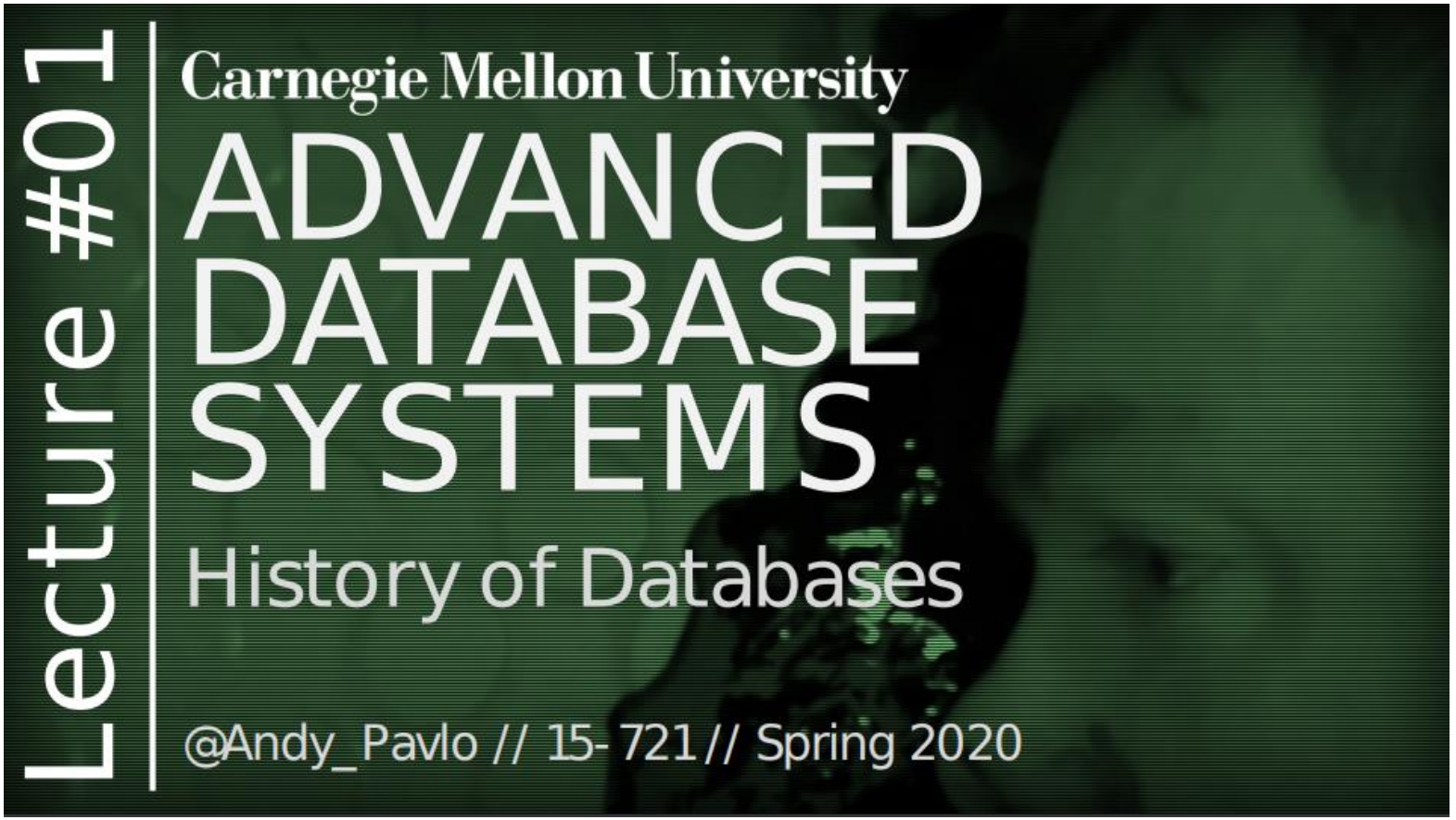
IBM
- System R (1975)
- First implementation of SQL

Led to
- Oracle
- IBM DB2
- INGRES
- Informix
- Sybase
- MS SQL Server (based on Sybase)

Sperry Univac computer system (1978)

DBMS Timeline

https://15721.courses.cs.cmu.edu/spring2020/slides/01-history.pdf Pag. 42-59



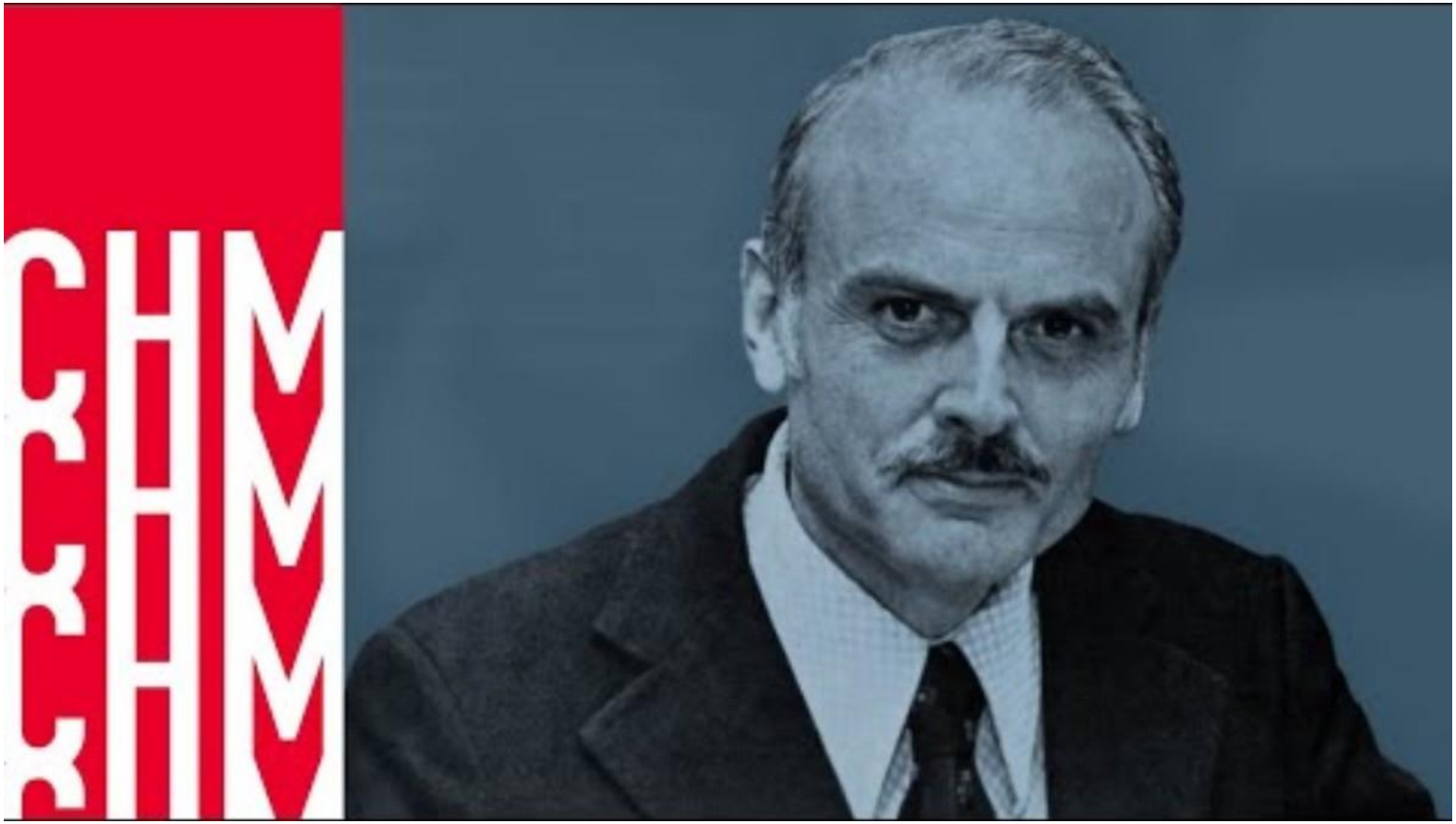
Relational Ideas

- Data is represented as a series of tables.
- The tables are Related to one another through a series of keys and foreign keys.
- A standard language is used to define the database (DDL) and to query the database (DML).
- Tables within the database contain the data about the database (meta data).

Why Relational?

- It is easy for most people to “see” and “get it”.
- Makes the data accessible for a wider number of users through user friendly query tools.
- Through good database design, space usage is efficient (although this has become less of an issue of late).

In a nutshell



Electronic files – Early computing 1950s – 1970s

Database Paradigms

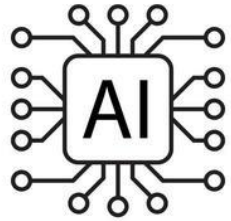
Databases today

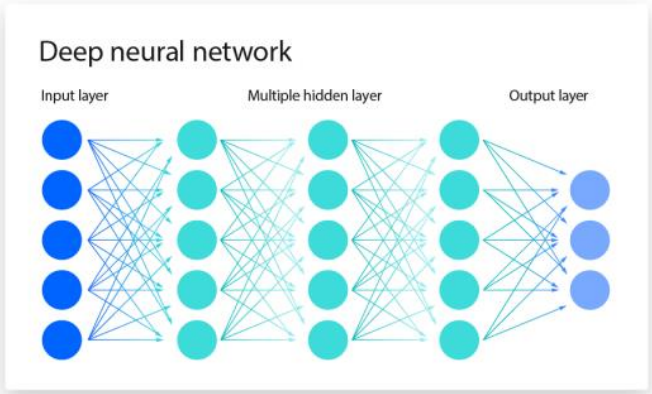
10 popular database management systems (DBMS)

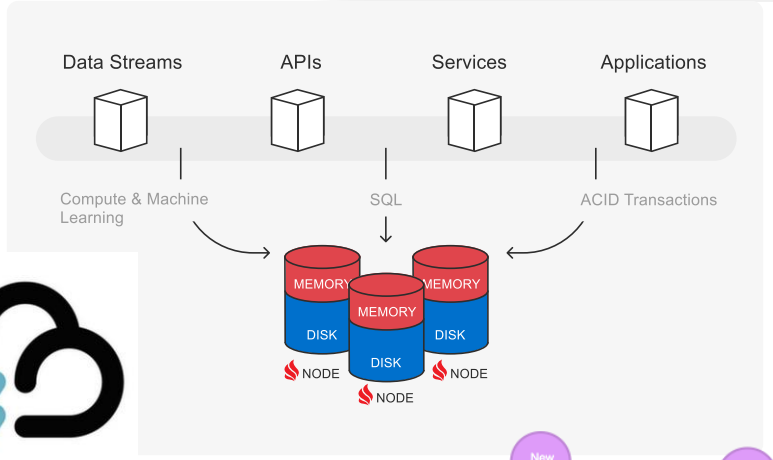
The Future?

- Object Oriented Databases
  - Combine data and operations on those data
  - Allows for inheritance
  - Oracle (Object-Relational Database)
  - Postgre(open source object-relational DBMS)
    - http://www.postgresql.org/about/
- XML and XML DBMS
  - XML designed to transport and store data initially envisioned as moving data across the web (w3schools.com)
  - XML Database Management System manages that data

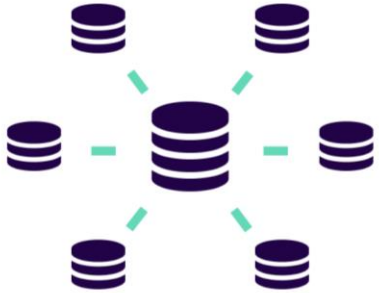
The Future?

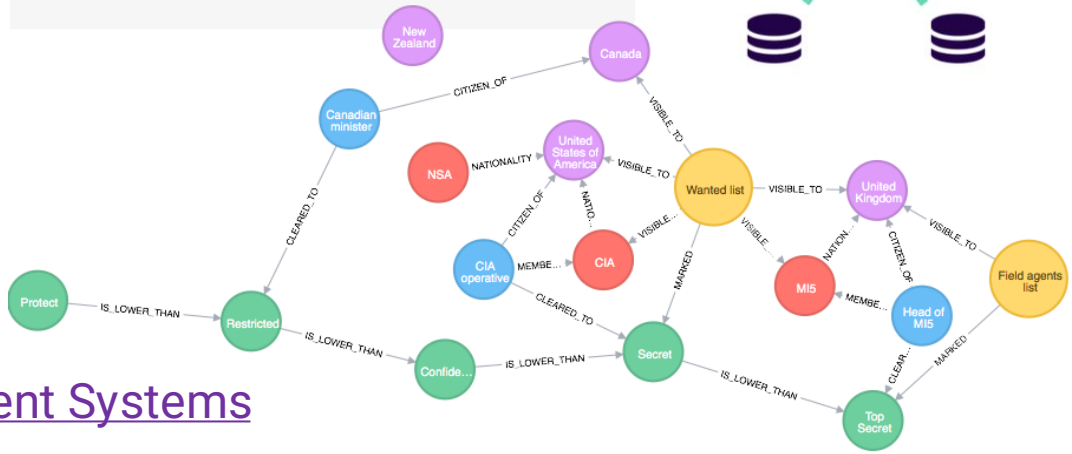






CYBERSECURITY
SLOGAN GOES HERE





[The Future of Database Management Systems]
[Exploring the Future of Databases]

Thanks!

Sebastián Aguilera Novoa
saguileran@itc.edu.co