Information Technology for Digital Humanities

Lecture 7

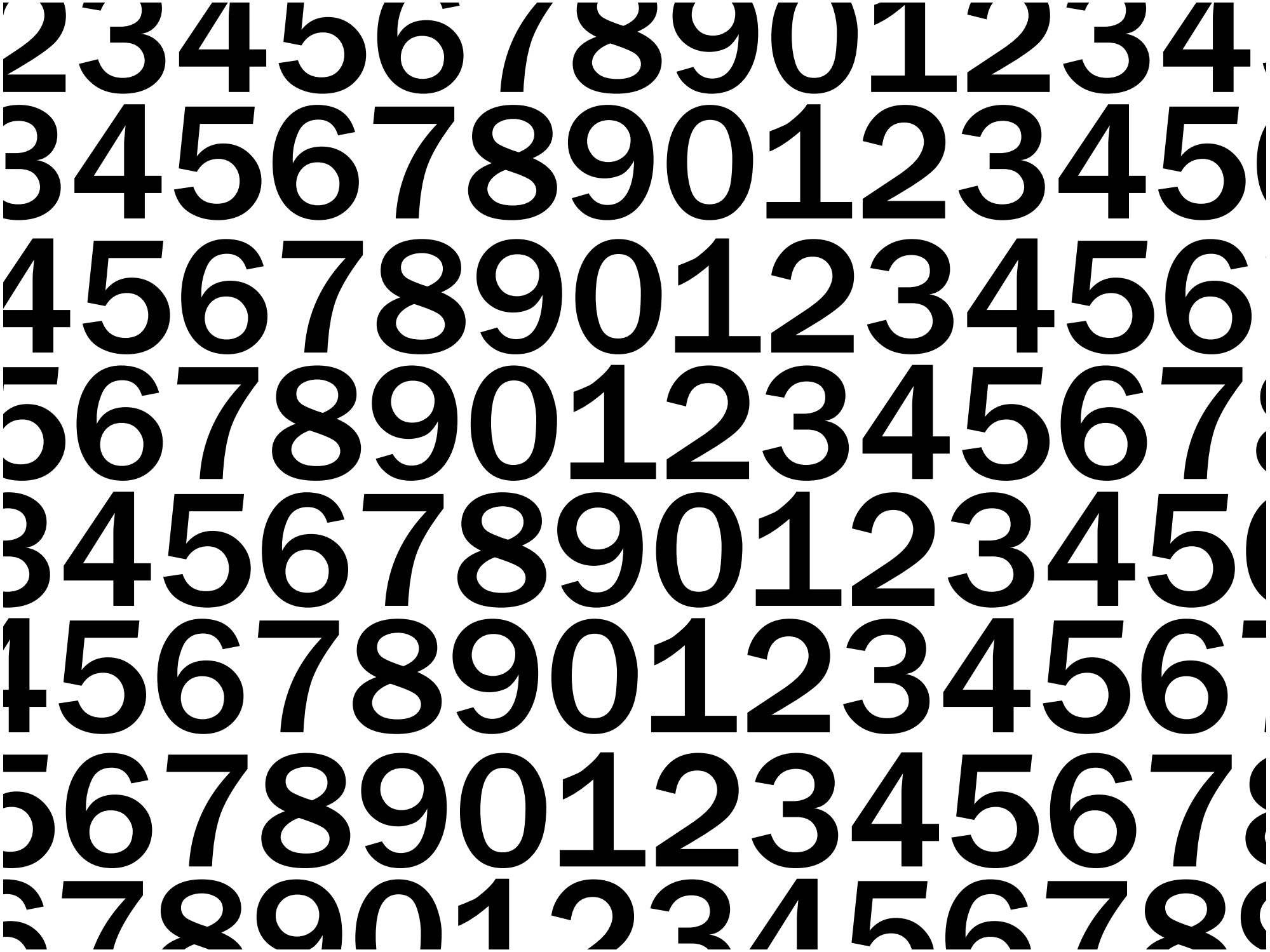
Mario Verdicchio

Università degli Studi di Bergamo

Academic Year 2023-2024

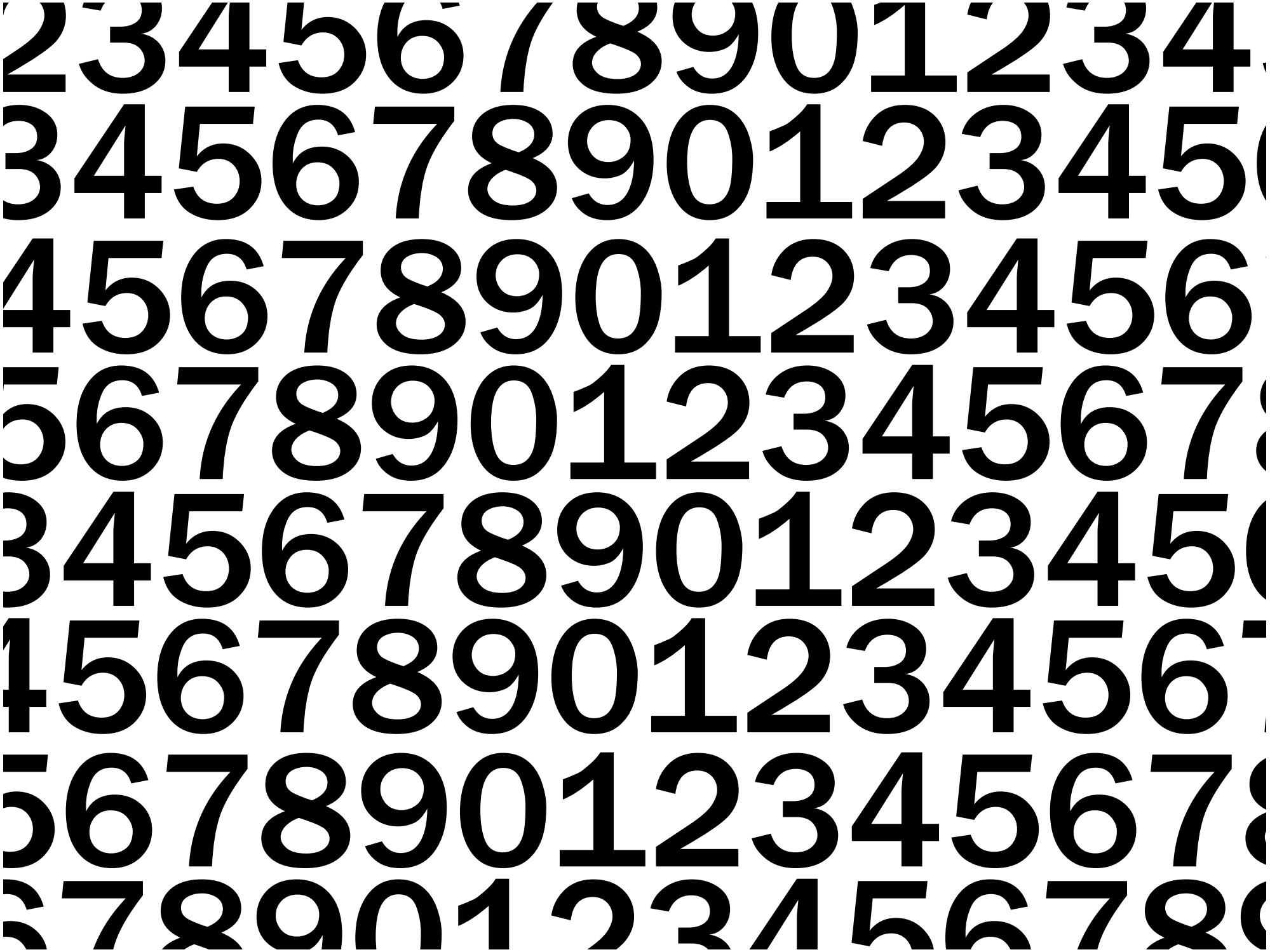
Lecture 7 (October 17 2023)

- Digital memories


23456789012345
34567890123456
45678901234567
56789012345678
34567890123456
45678901234567
56789012345678
78901234567890

We have seen that, in the context of Information Technology, everything is a number. We have seen how digital images are images mapped onto numbers.


x y r g b
(9,14,245,133,167)

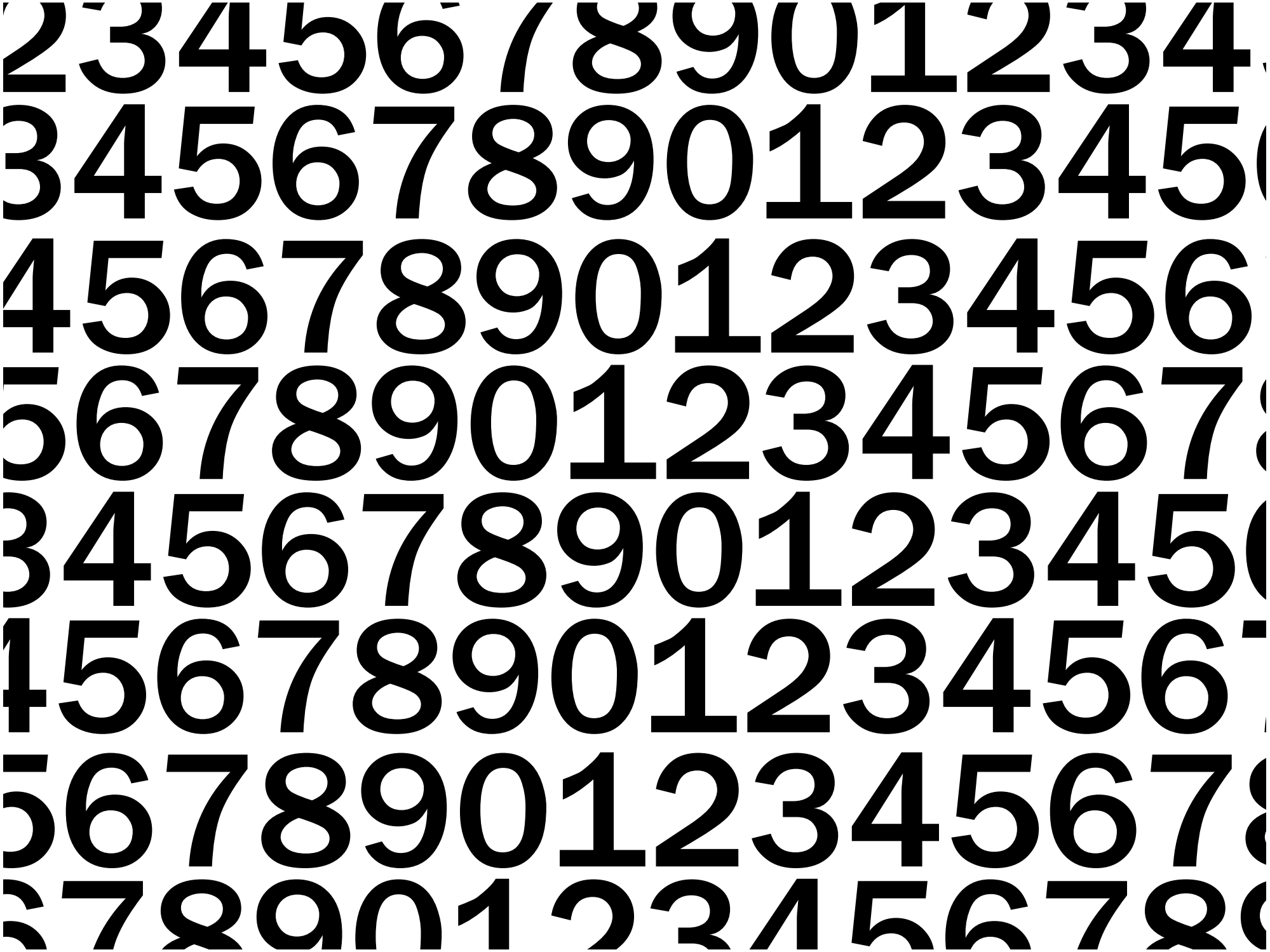

23456789012345
34567890123456
45678901234567
56789012345678
34567890123456
45678901234567
56789012345678
78901234567890

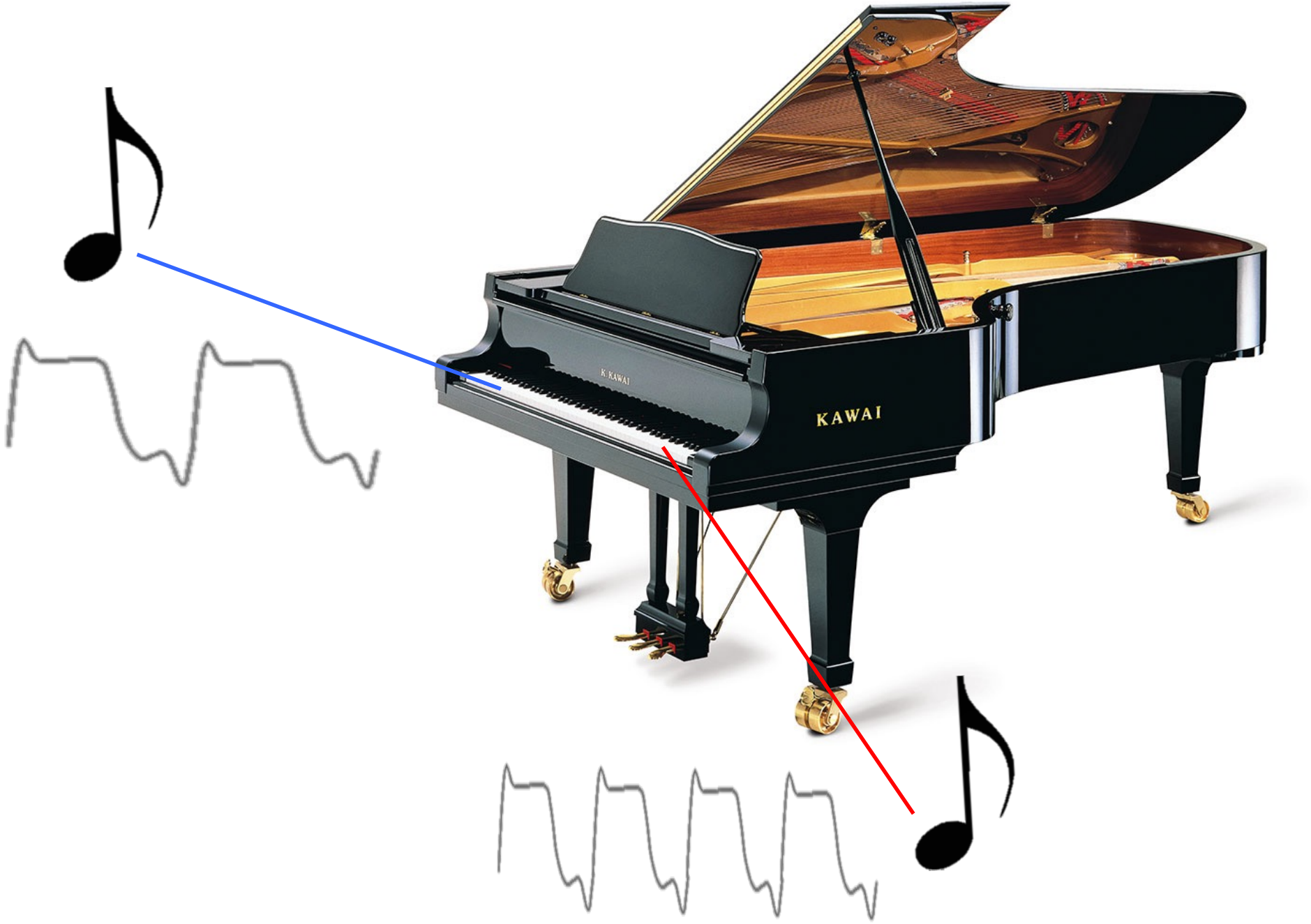
US-ASCII Code Chart. Scanner copied from the material delivered with TermiNet 300 impact type printer with Keyboard, February 1972, General Electric Data communication Product Dept., Waynesboro, Virginia.

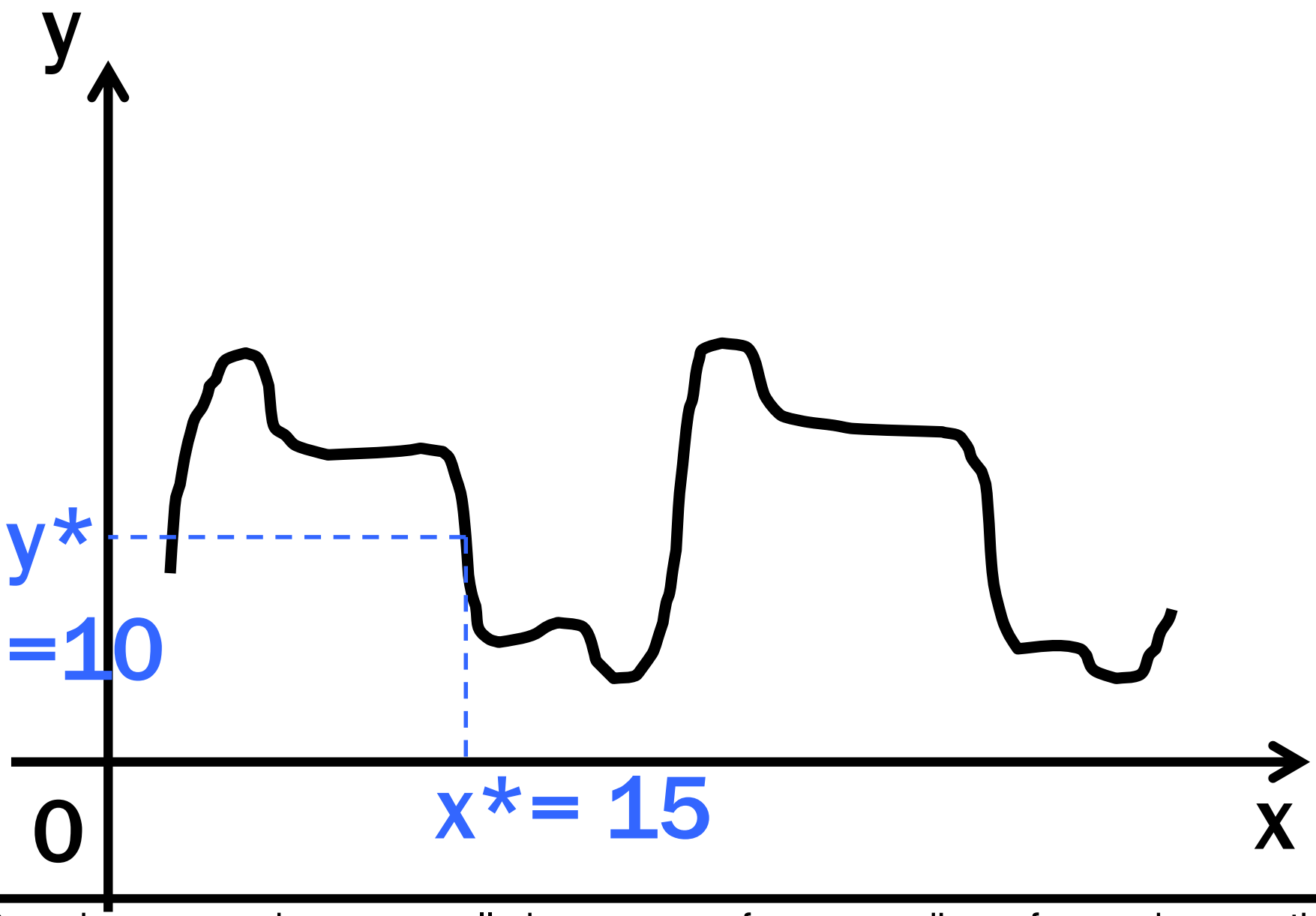
USASCII code chart

$b_7 b_6 b_5$ → / Bits					0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
b_4	b_3	b_2	b_1	Row ↓ Column →	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	|
1	1	0	1	13	CR	GS	-	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	SI	US	/	?	O	_	o	DEL

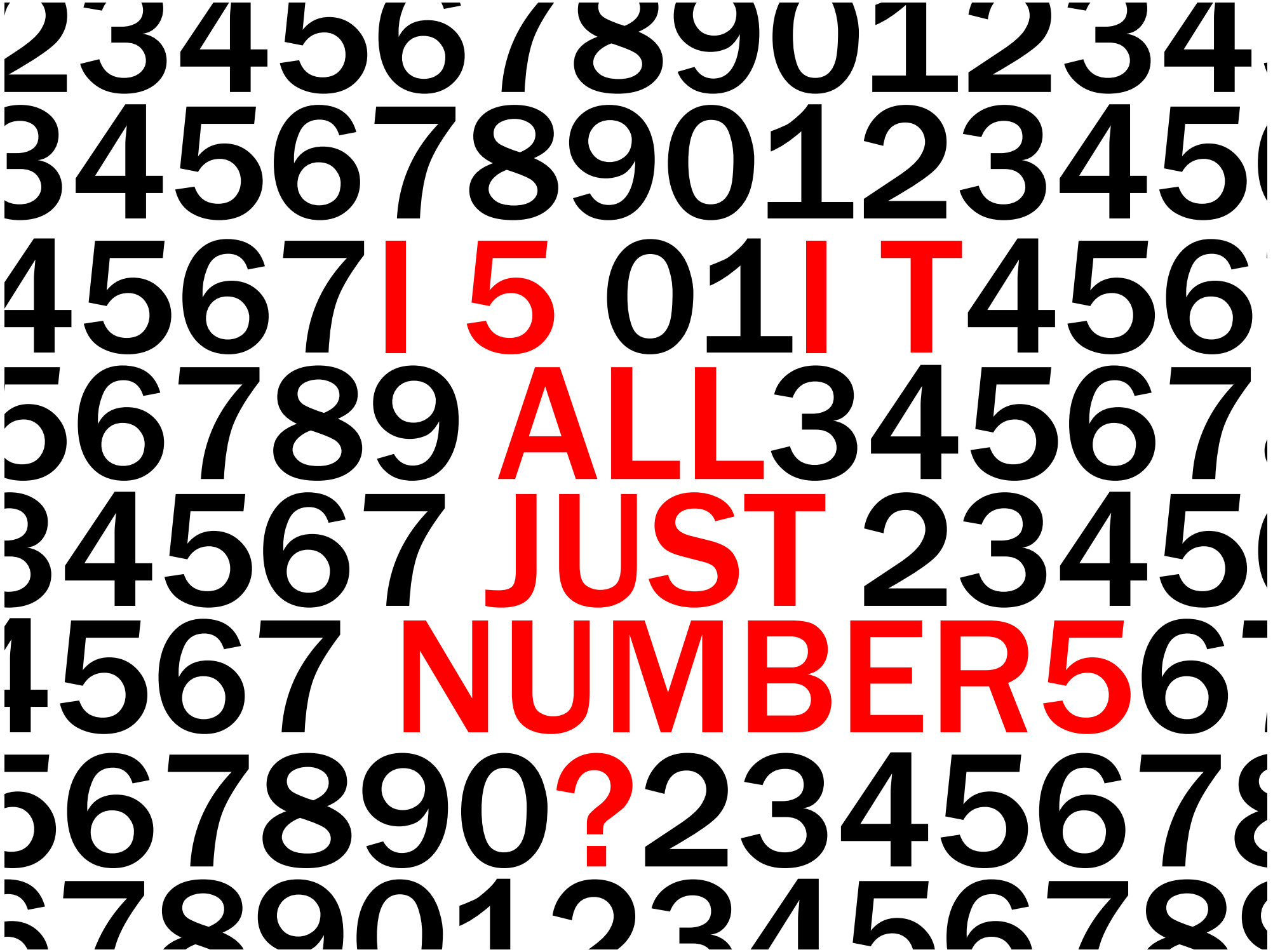
Texts are numbers as well, based the encoding of the characters the texts are comprised of. This is the earliest example of character encoding, which has been and still is expanded throughout the years.

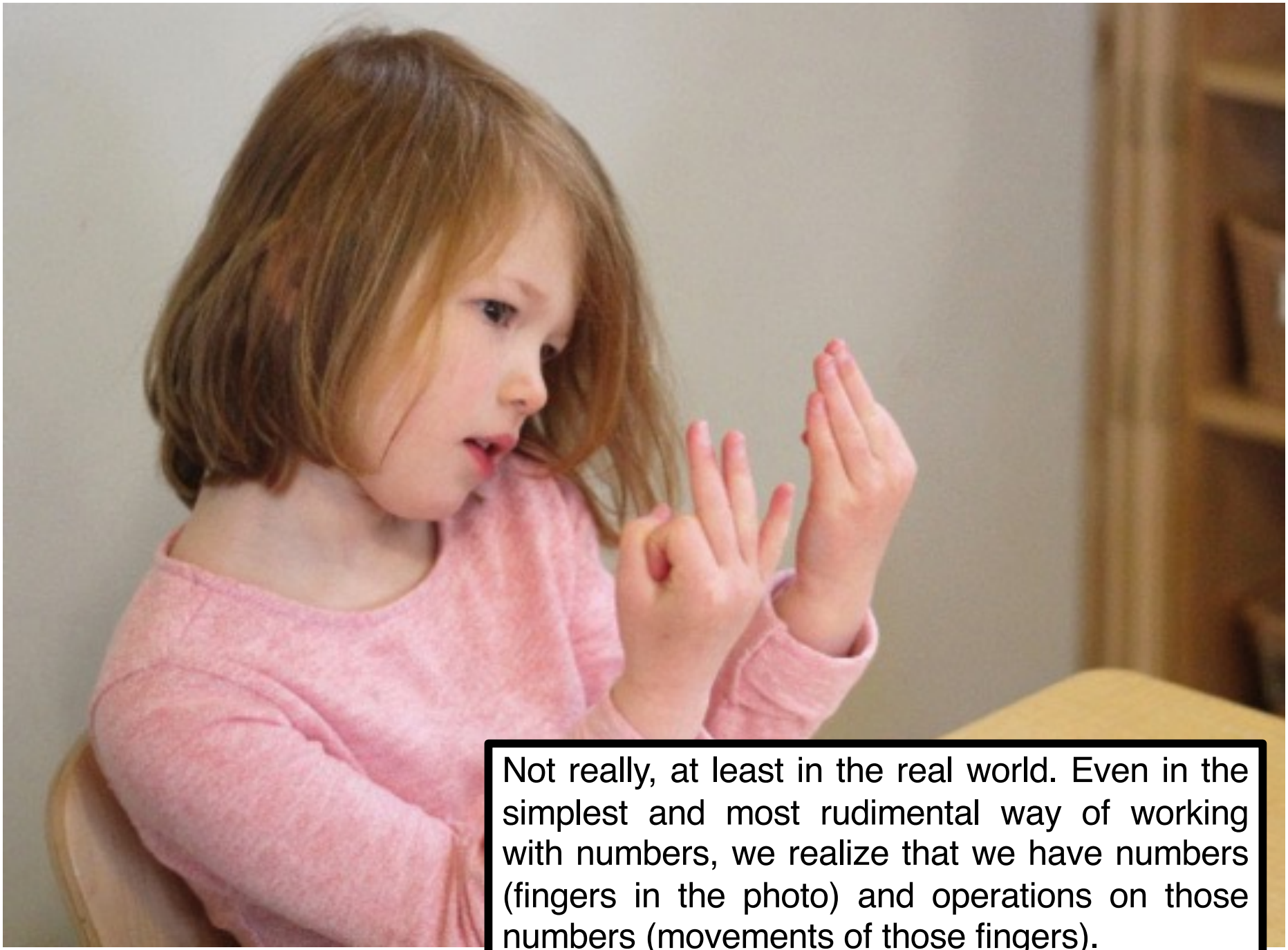

23456789012345
34567890123456
45678901234567
56789012345678
34567890123456
45678901234567
56789012345678
78901234567890


KAWAI



Sounds are numbers as well, by means of an encoding of soundwaves that enables the mapping of samples onto pairs of numbers.


I 5 01I T
ALL
JUST
NUMBER5
?



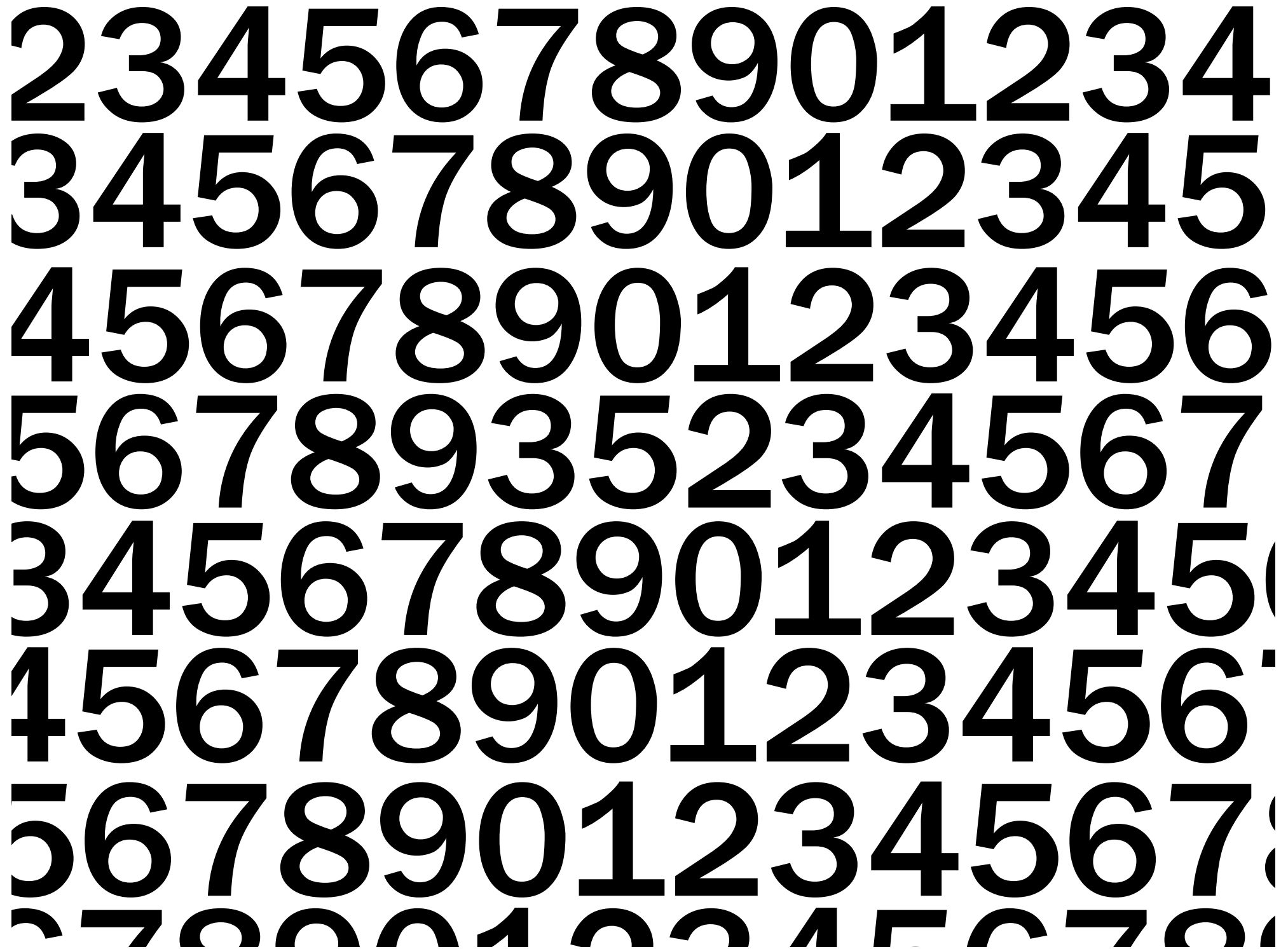
Not really, at least in the real world. Even in the simplest and most rudimental way of working with numbers, we realize that we have numbers (fingers in the photo) and operations on those numbers (movements of those fingers).

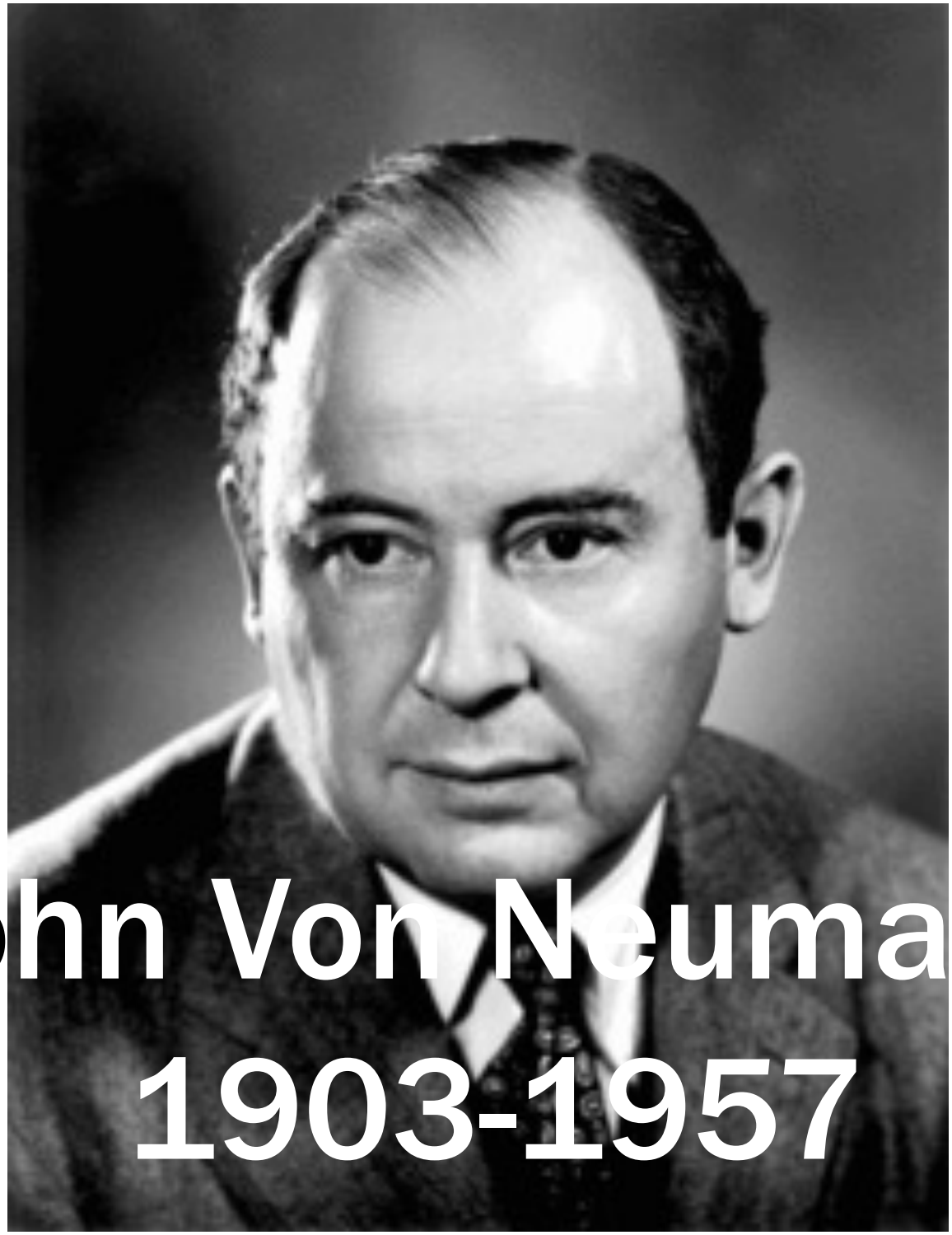
3+2

When using digits for numbers and symbols for operations, the distinction is clear.

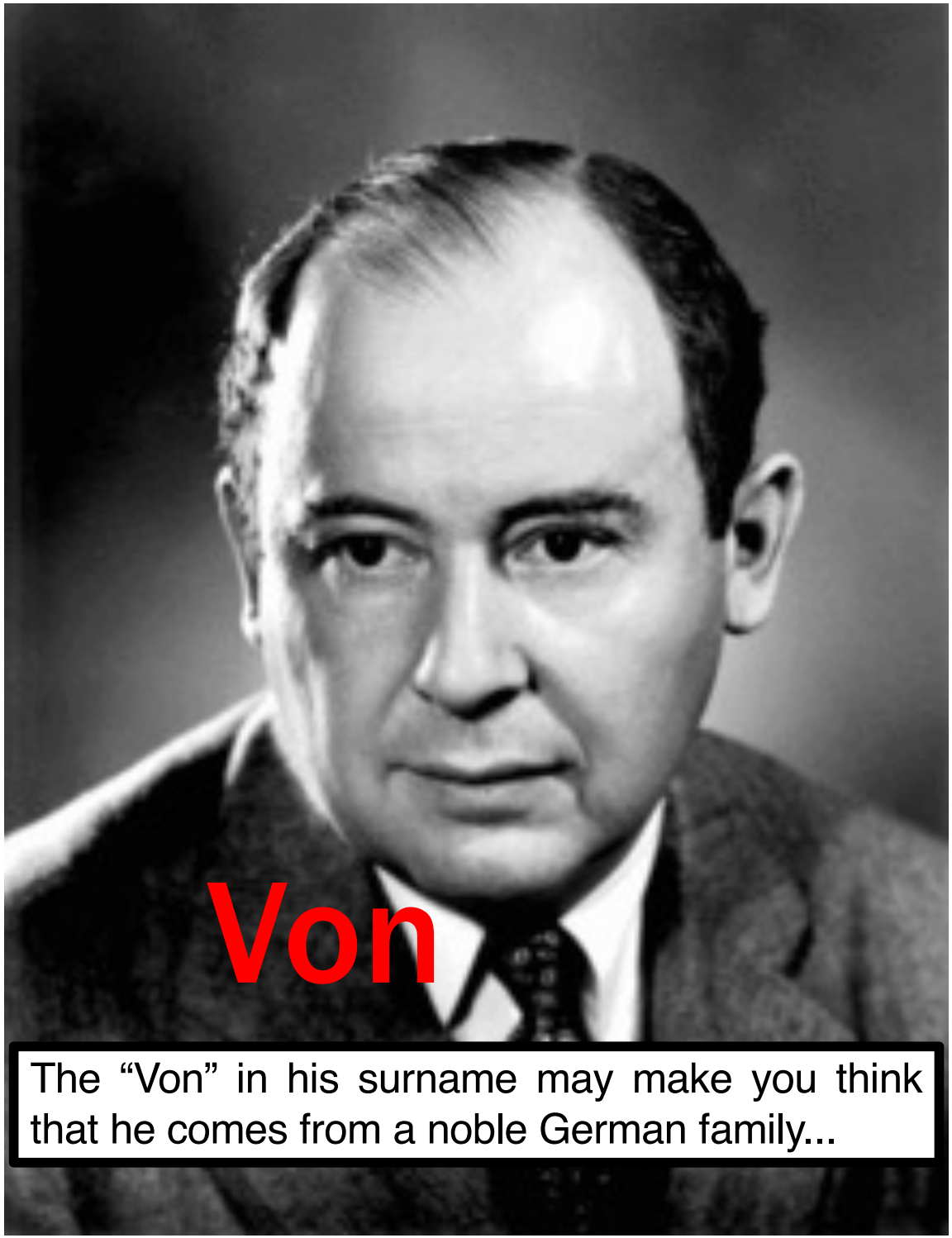
352

However, if operations get encoded and hence expressed with numbers, it may get confusing. Is this “threehundredfiftytwo”? We need a clear criterion to deal with numbers that are data and numbers that are operations on such data.

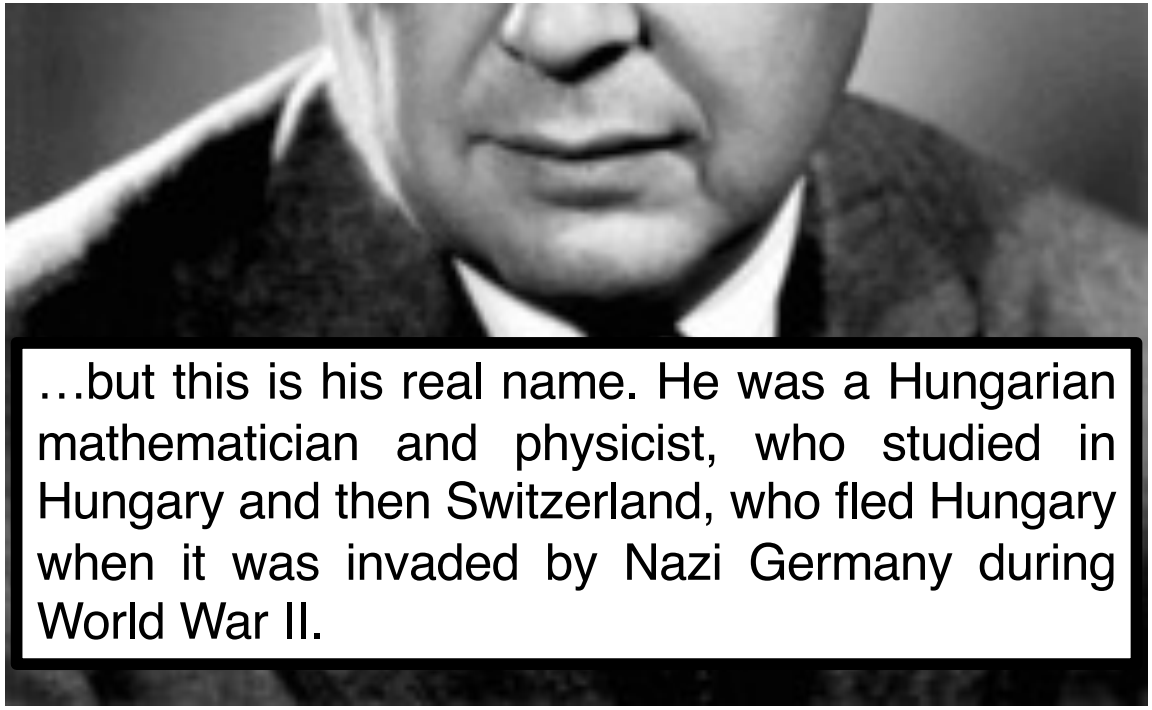

2345678901234
3456789012345
4567890123456
5678935234567
3456789012345
4567890123456
5678901234567



John Von Neumann
1903-1957


Von
The “Von” in his surname may make you think that he comes from a noble German family...

János Lajos Neumann



…but this is his real name. He was a Hungarian mathematician and physicist, who studied in Hungary and then Switzerland, who fled Hungary when it was invaded by Nazi Germany during World War II.

Nuclear bomb test (Bikini Atoll, Micronesia, 1946)

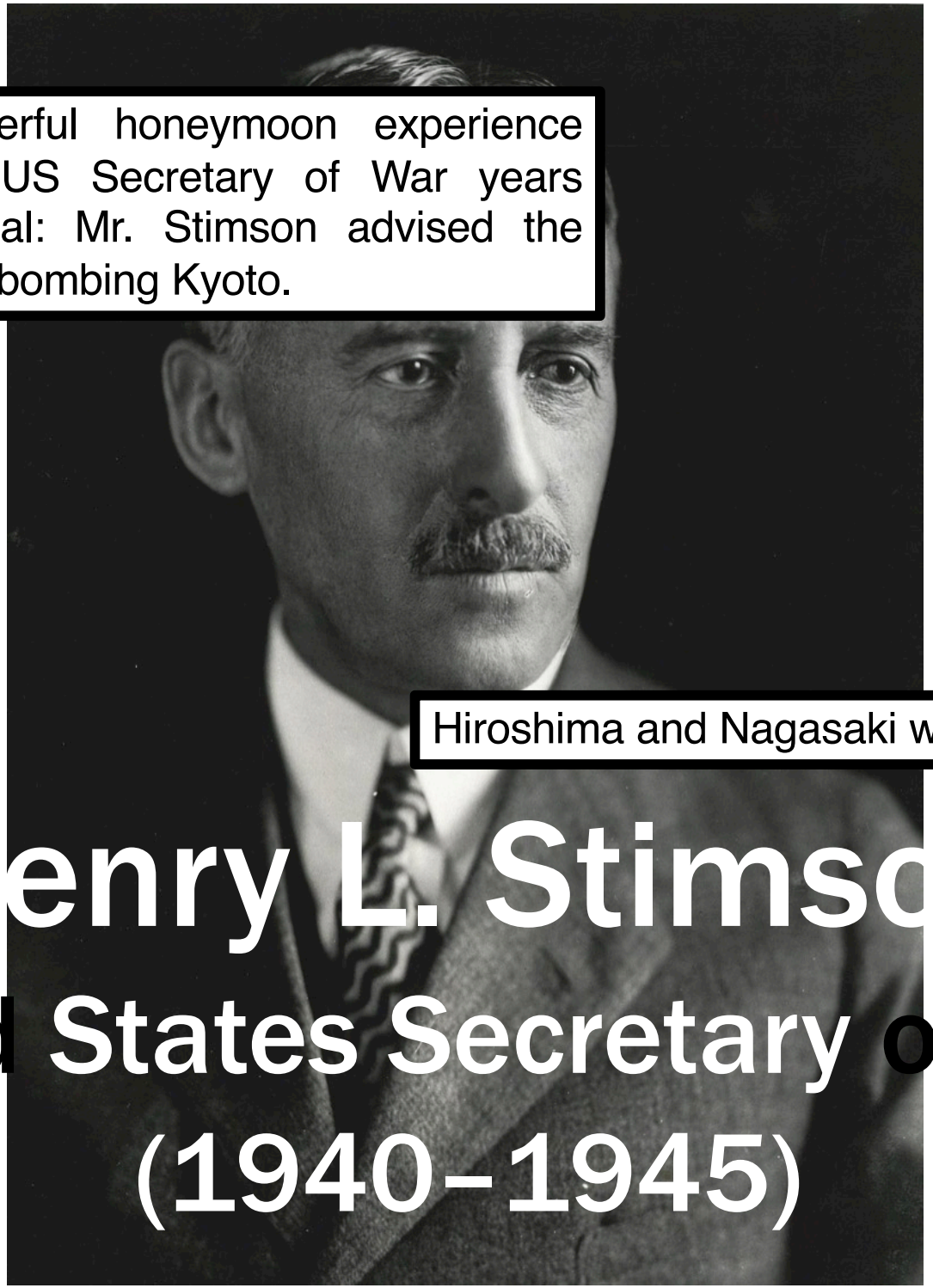

He was extremely keen on obtaining the most devastation possible from his experiments, calculating the most effective profile shapes and trajectories of the launched bombs. He proposed to bomb Kyoto, the cultural capital of Japan.

insidekyoto.com/honeymoon-in-kyoto

Honeymoon in Kyoto

Kyoto is an incredibly romantic city. With intimate restaurants, atmospheric lanes, superb accommodations and a thousand quiet gardens and temples, it's the perfect place to spend time with someone you love. Here's our full guide to honeymooning in Kyoto.

This beautiful photo of Kyoto is taken from a website promoting honeymoon trips to the city.



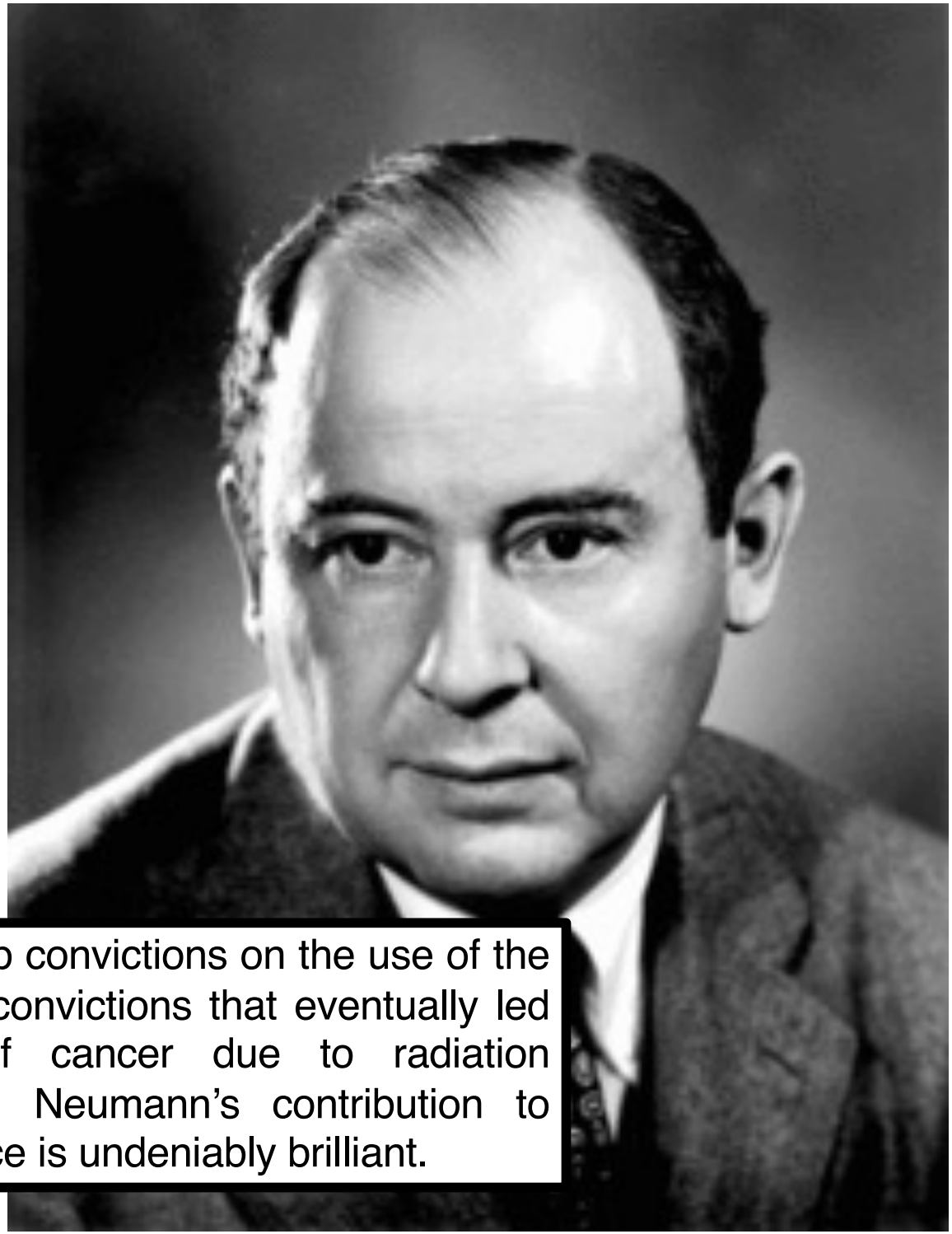
Indeed, a wonderful honeymoon experience enjoyed by the US Secretary of War years before was crucial: Mr. Stimson advised the President against bombing Kyoto.

Hiroshima and Nagasaki were chosen instead.

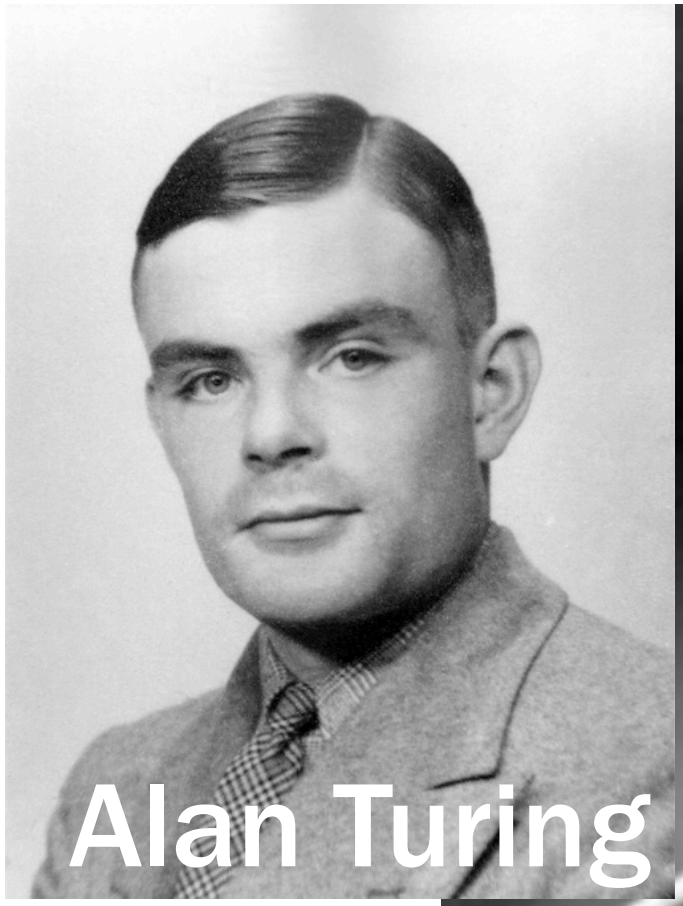
Henry L. Stimson

United States Secretary of War

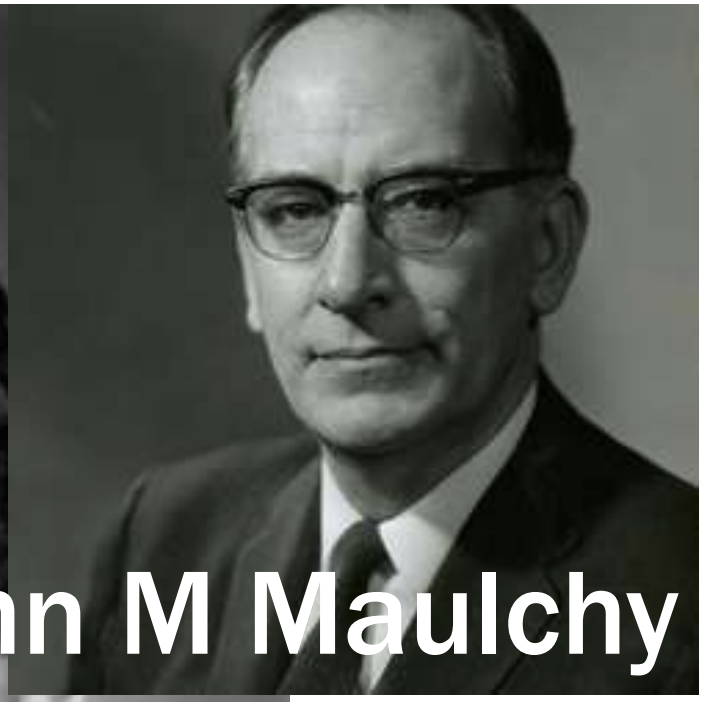
(1940–1945)



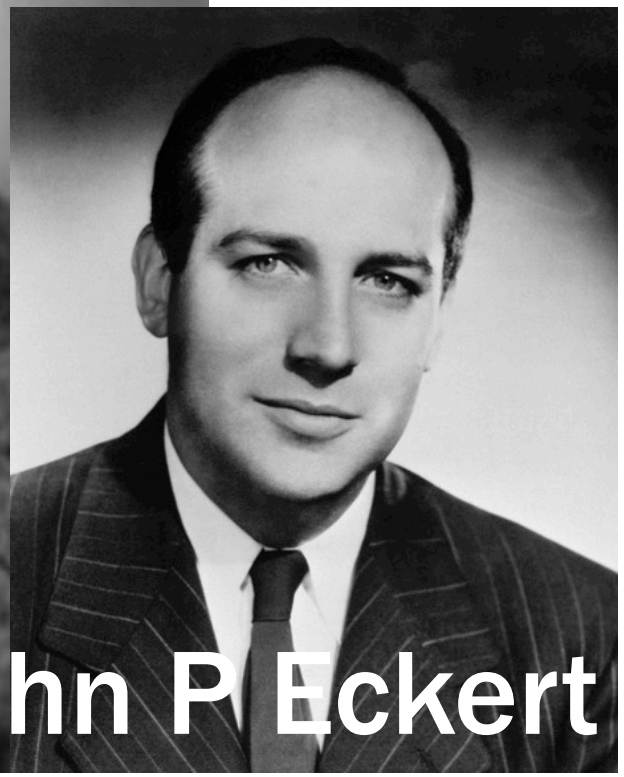
Despite his deep convictions on the use of the nuclear bomb (convictions that eventually led him to die of cancer due to radiation exposure), Von Neumann's contribution to computer science is undeniably brilliant.



Alan Turing



John M Maulchy



John P Eckert

The merit of the paradigm of data organization inside a computer he proposed needs to be shared with some of his contemporaries, who were working and collaborating with him on the topic.


THE
STORED
PROGRAM

THE STORED PROGRAM

Both operations and operands can be stored in the same place.

THE STORED PROGRAM

Both operations and operands can be stored in the same place.

We are going to color-code the different ways in which data is used inside a computer's memory. Red for operations, blue for operands (i.e. the data to which we apply the operations), green for indicating the locations in the memory where data is stored. These colors are only useful to humans, obviously. No such colorful distinction is possible inside a computer.

operations

operands

3+2

operations

operands

By the use of colors, we humans know that the ‘5’ in the sequence is to be interpreted in a different way than ‘3’ and ‘2’. How does a computer draw such a distinction?

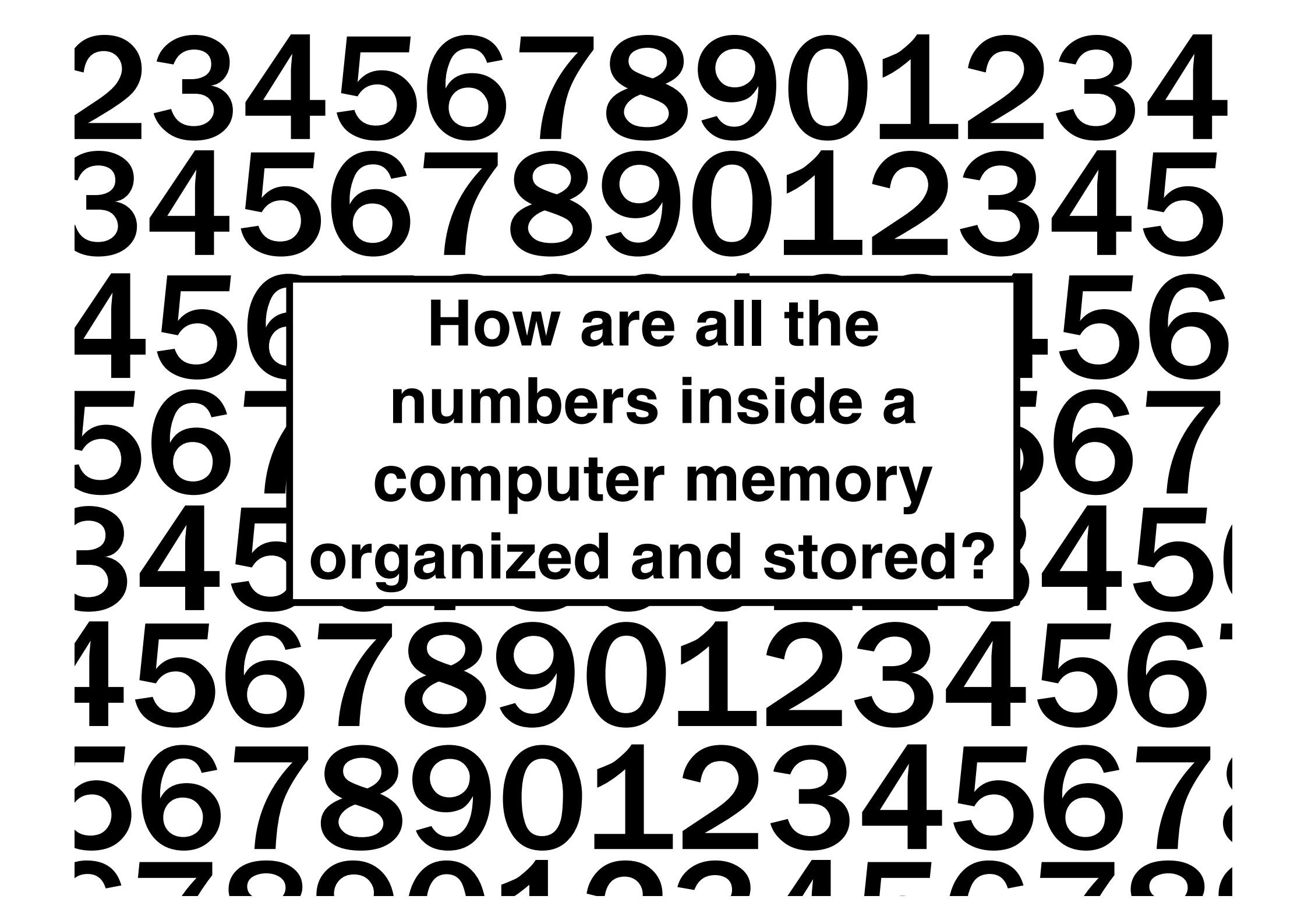
352

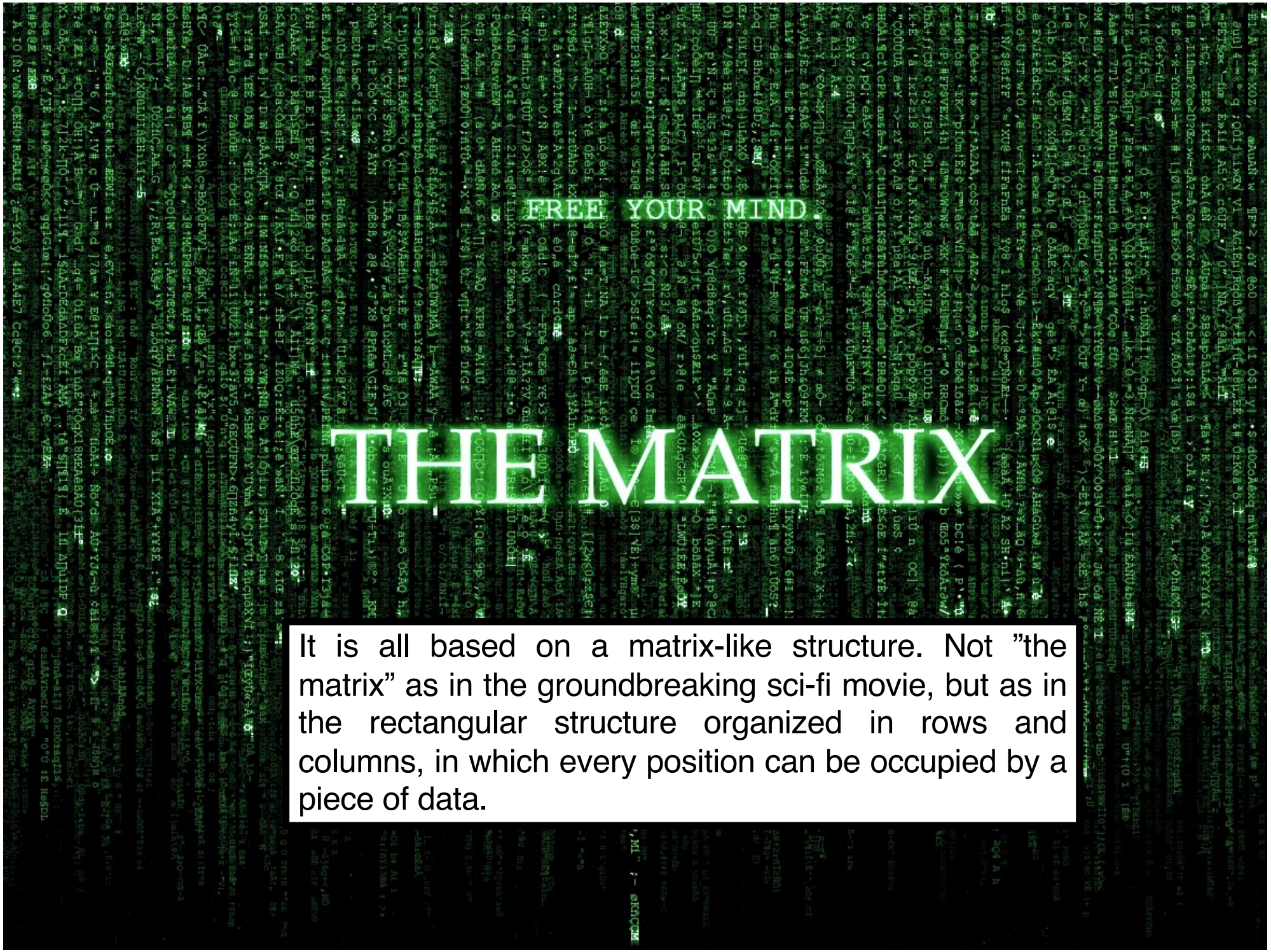
operations

operands

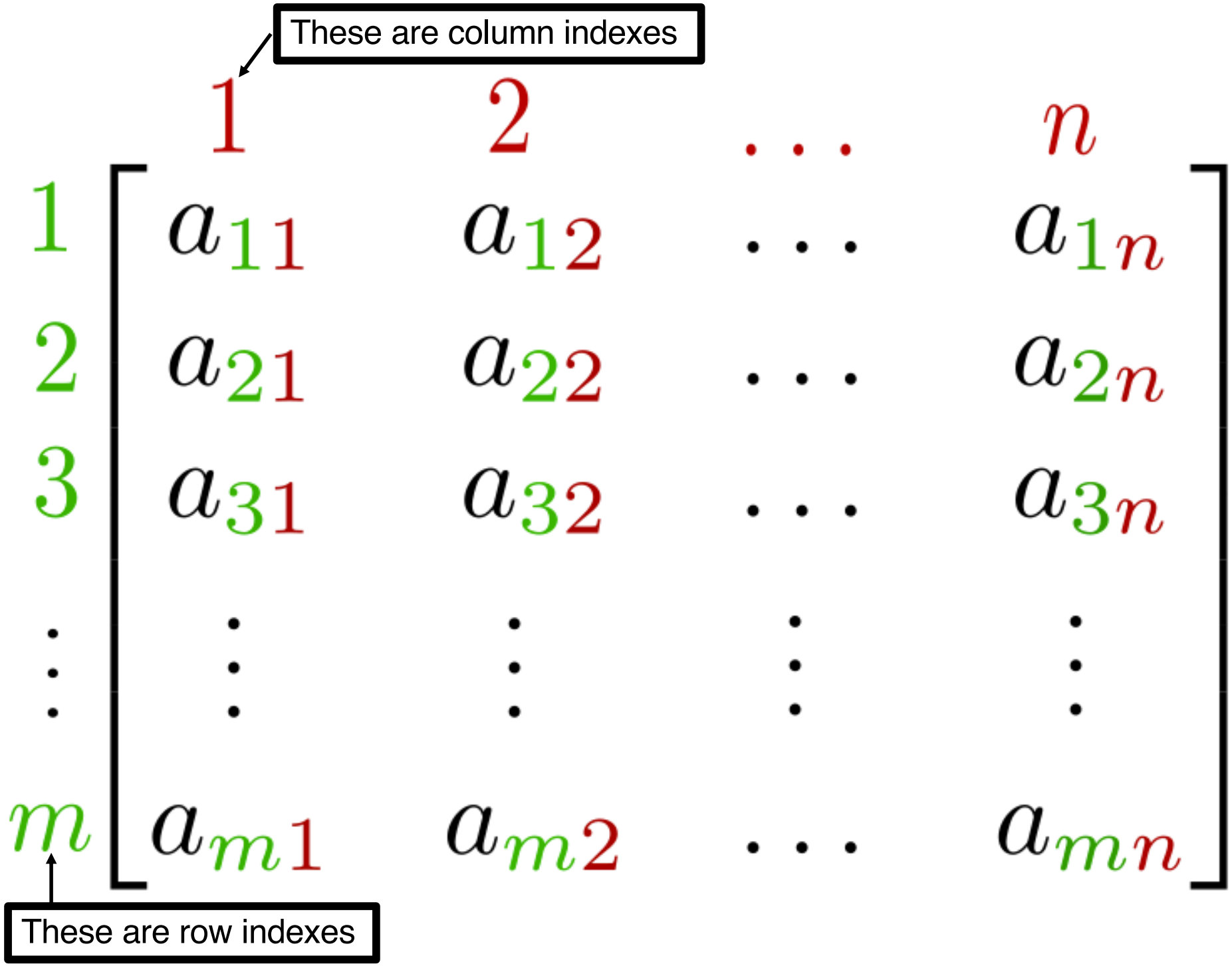
Such distinction is location-based, that is, we need to deal with where the data are stored.

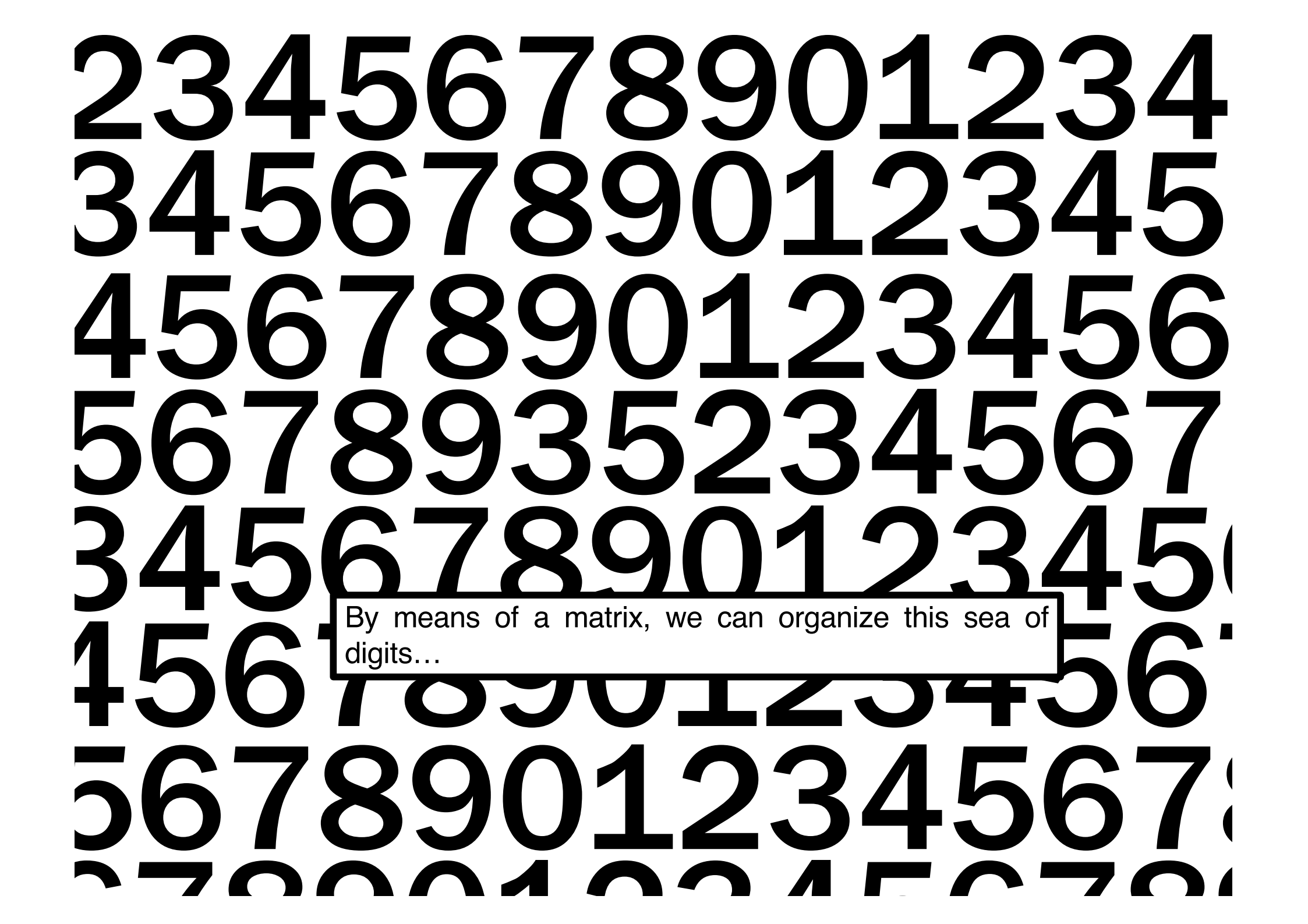
place


How are all the numbers inside a computer memory organized and stored?

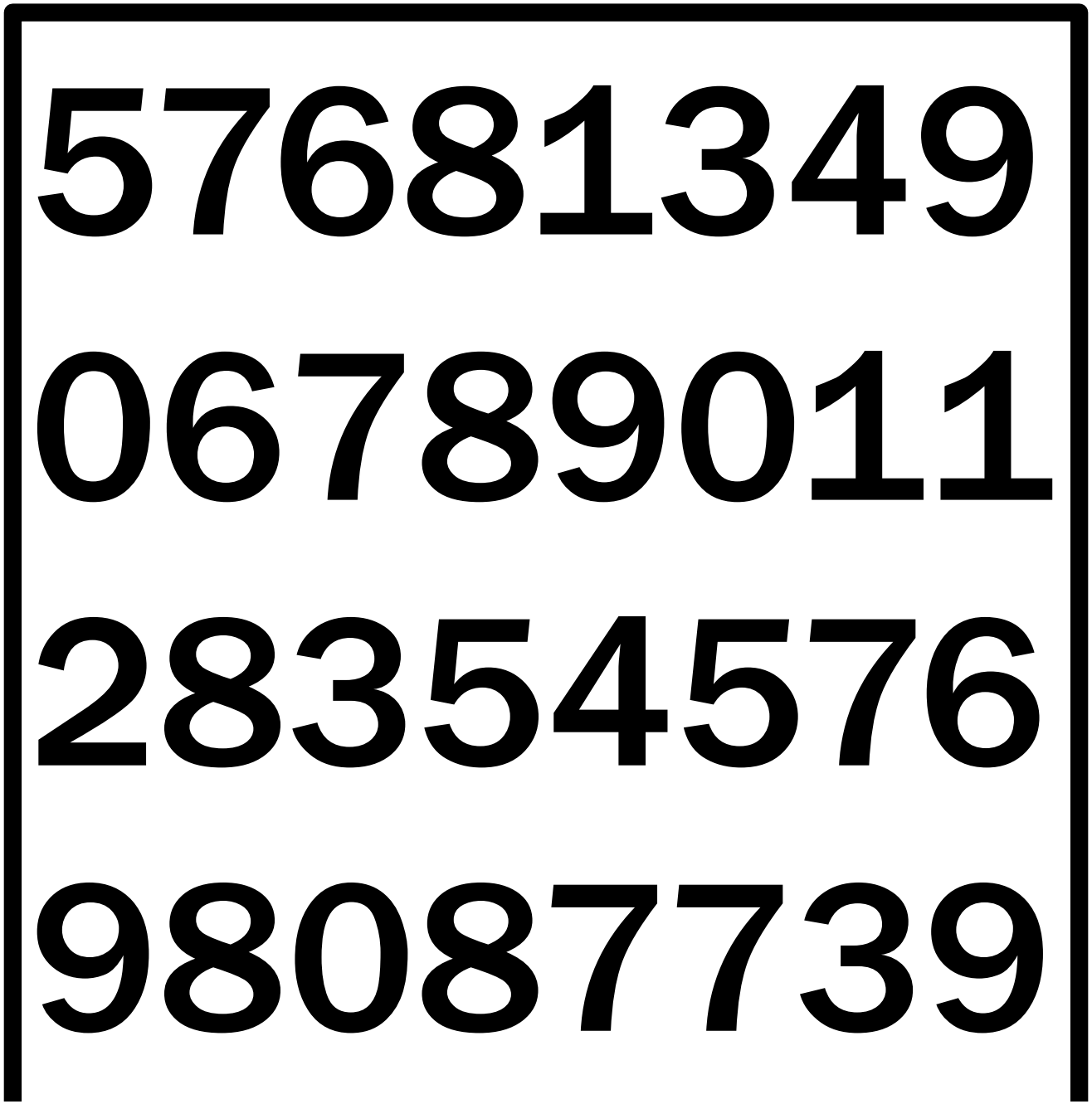


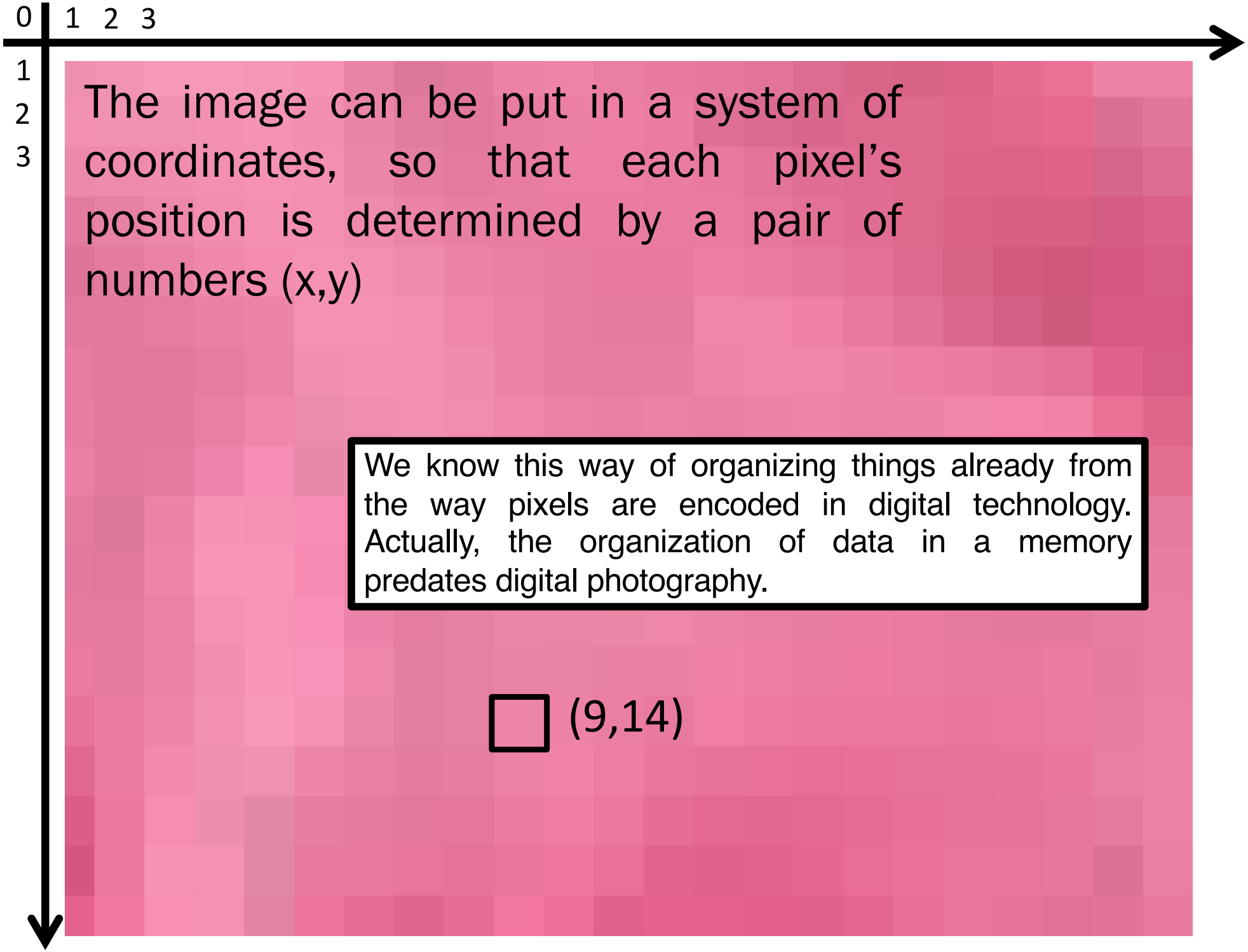
It is all based on a matrix-like structure. Not ”the matrix” as in the groundbreaking sci-fi movie, but as in the rectangular structure organized in rows and columns, in which every position can be occupied by a piece of data.


These are column indexes
1 2 . . . n
1 a11 a12 . . . a1n
2 a21 a22 . . . a2n
3 a31 a32 . . . a3n
m am1 am2 . . . amn
These are row indexes


By means of a matrix, we can organize this sea of digits…

…into a much neater structure with rows and columns.

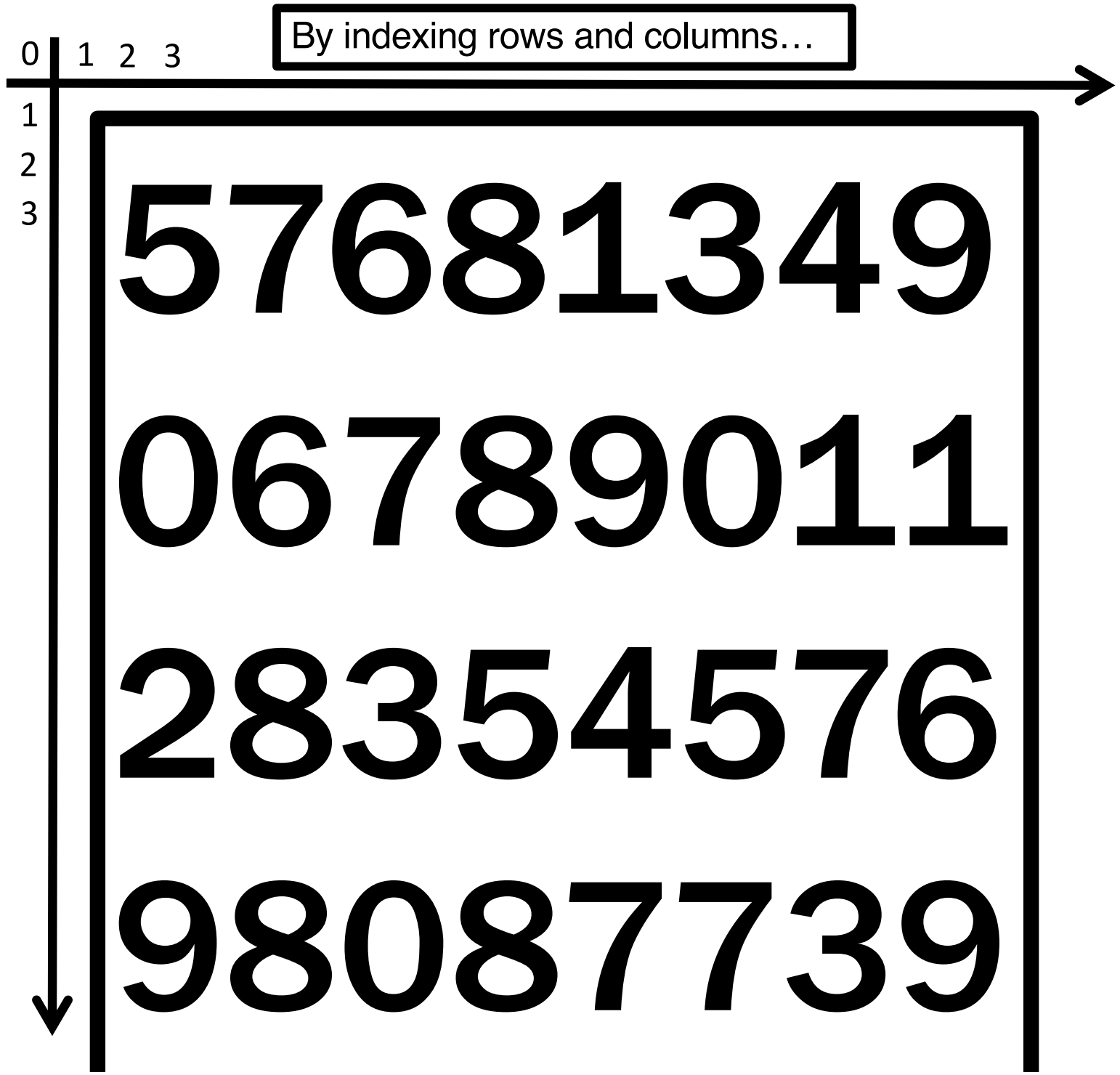




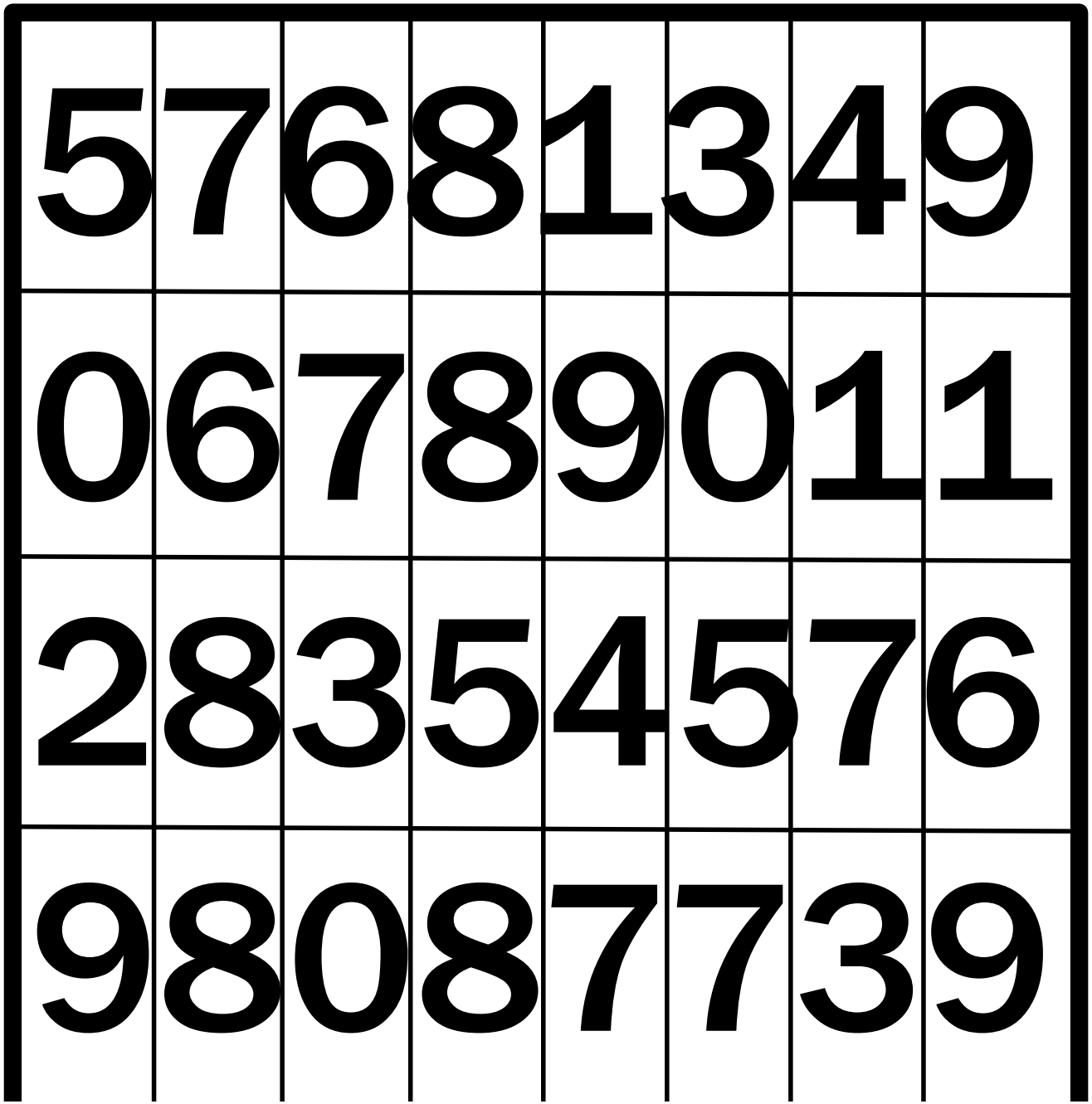
The image can be put in a system of coordinates, so that each pixel's position is determined by a pair of numbers (x,y)

We know this way of organizing things already from the way pixels are encoded in digital technology. Actually, the organization of data in a memory predates digital photography.

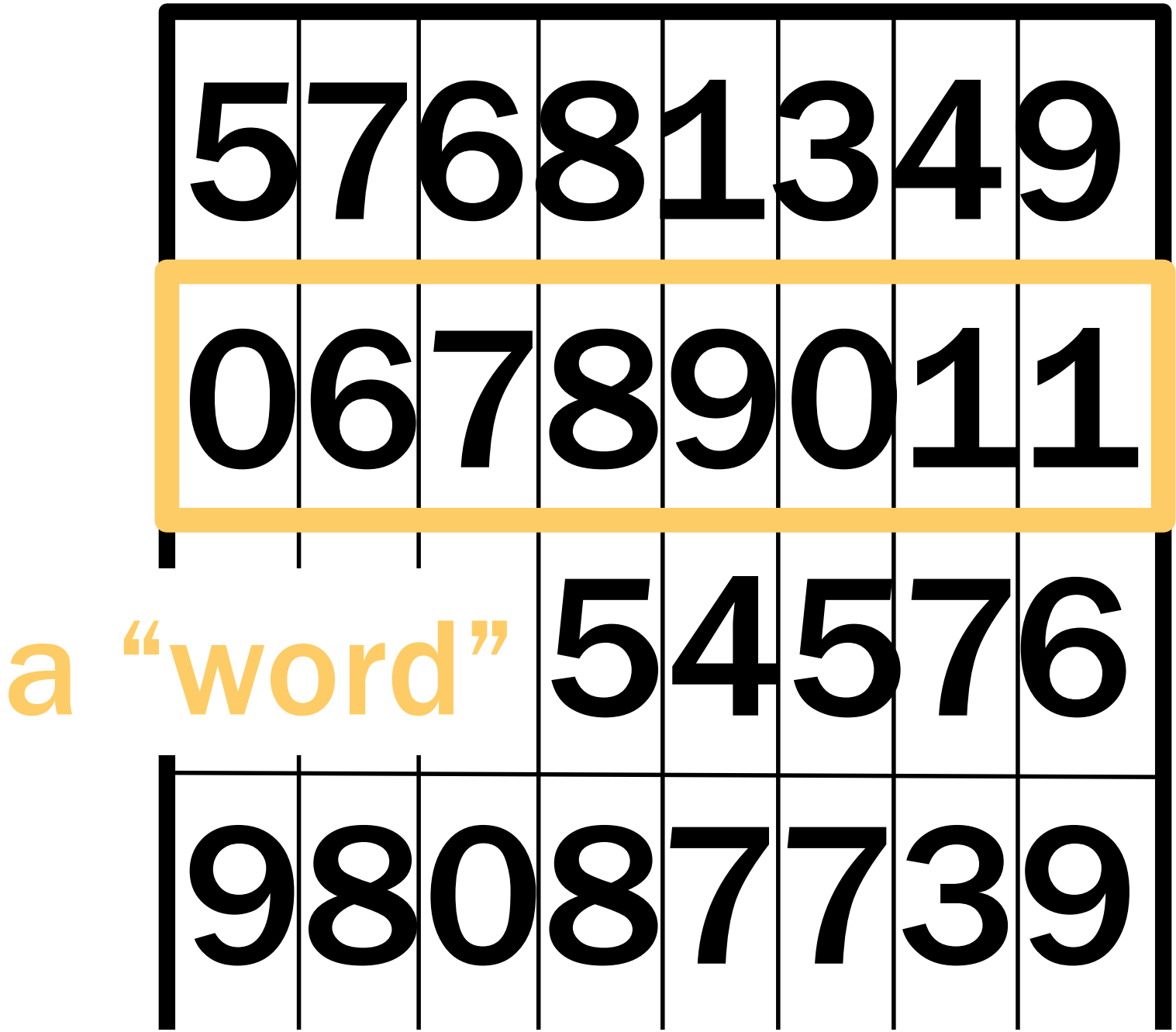
(9,14)


By indexing rows and columns…
0 1 2 3
1
2
3
57681349
06789011
28354576
98087739

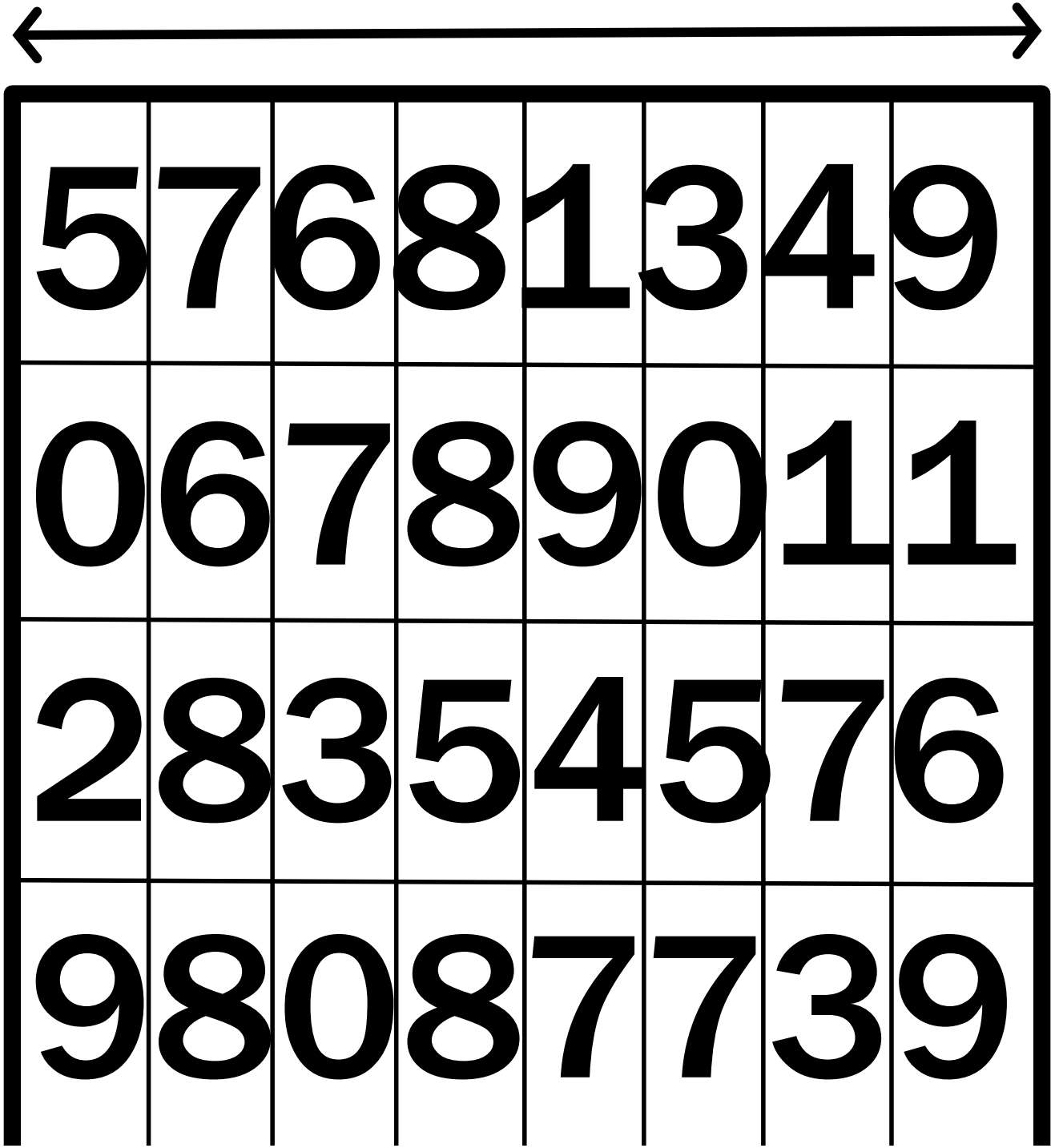
…we introduce a matrix of places. Each place can contain one piece of data.



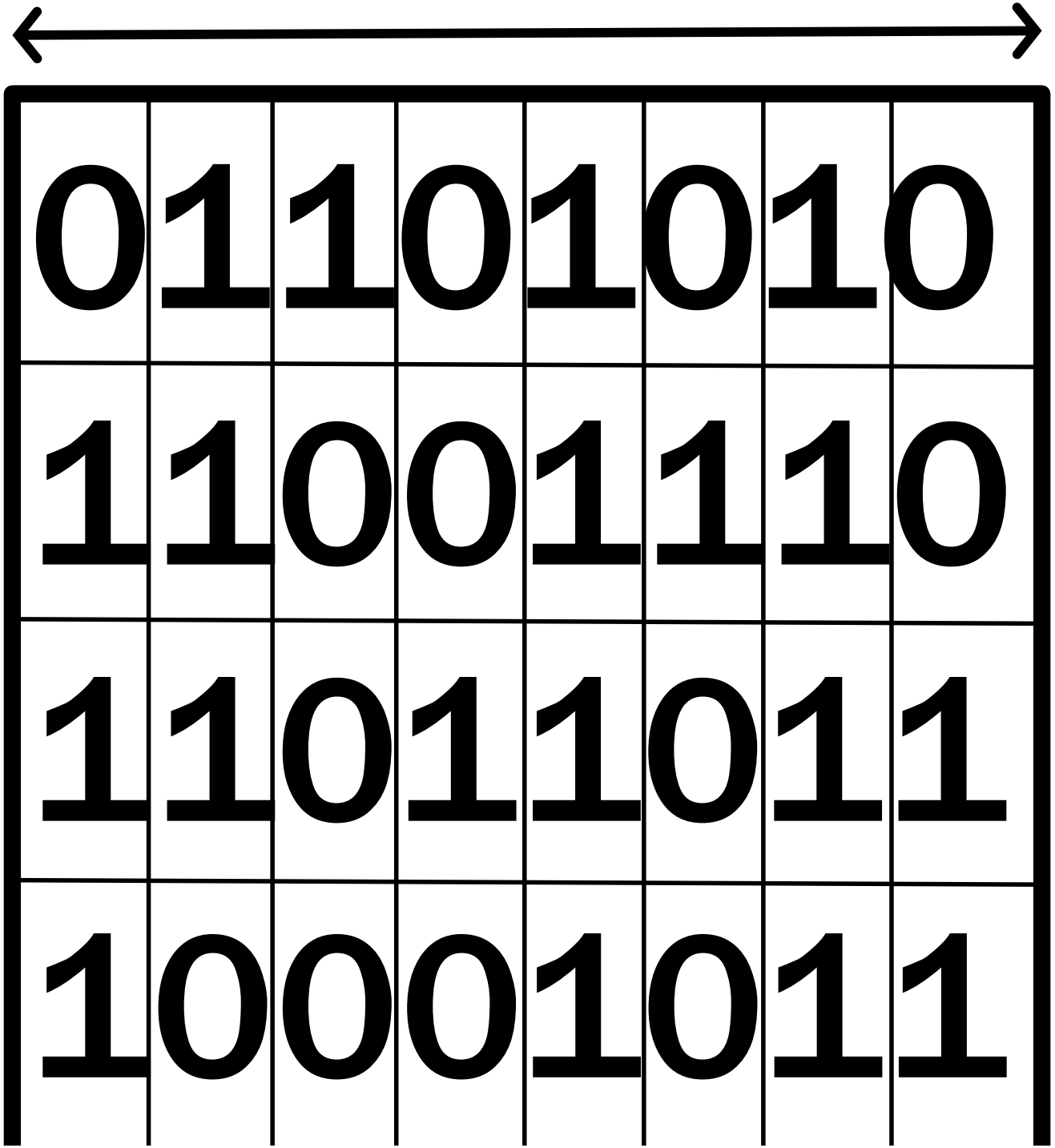
Computers usually access data row by row, meaning that they do not transfer one single data, but an entire row at a time. Rows inside a computer memory are called “words”.



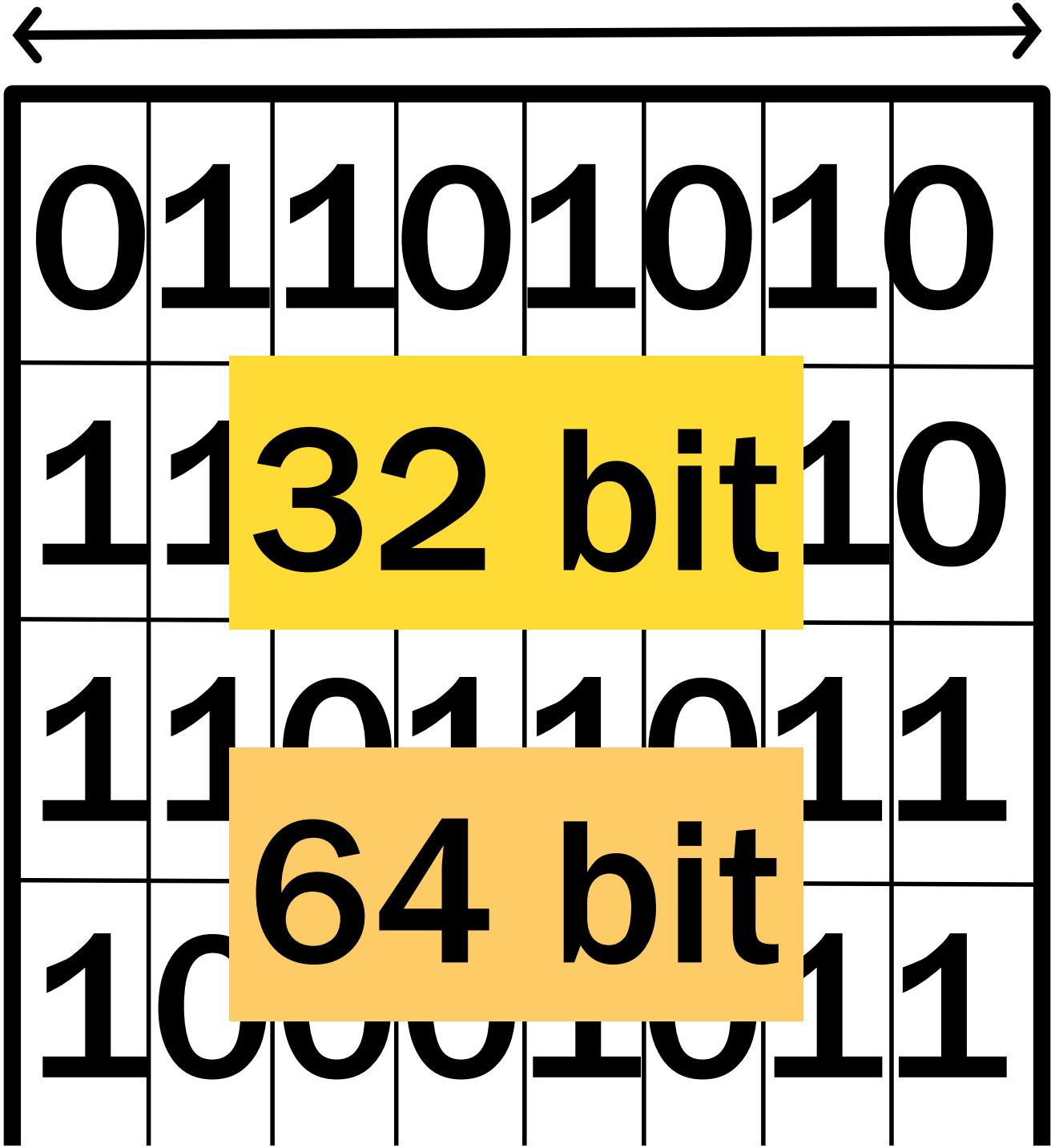
The width of the memory coincides with the length of a row.
In real computers, each position is not occupied by a digit in the decimal system…



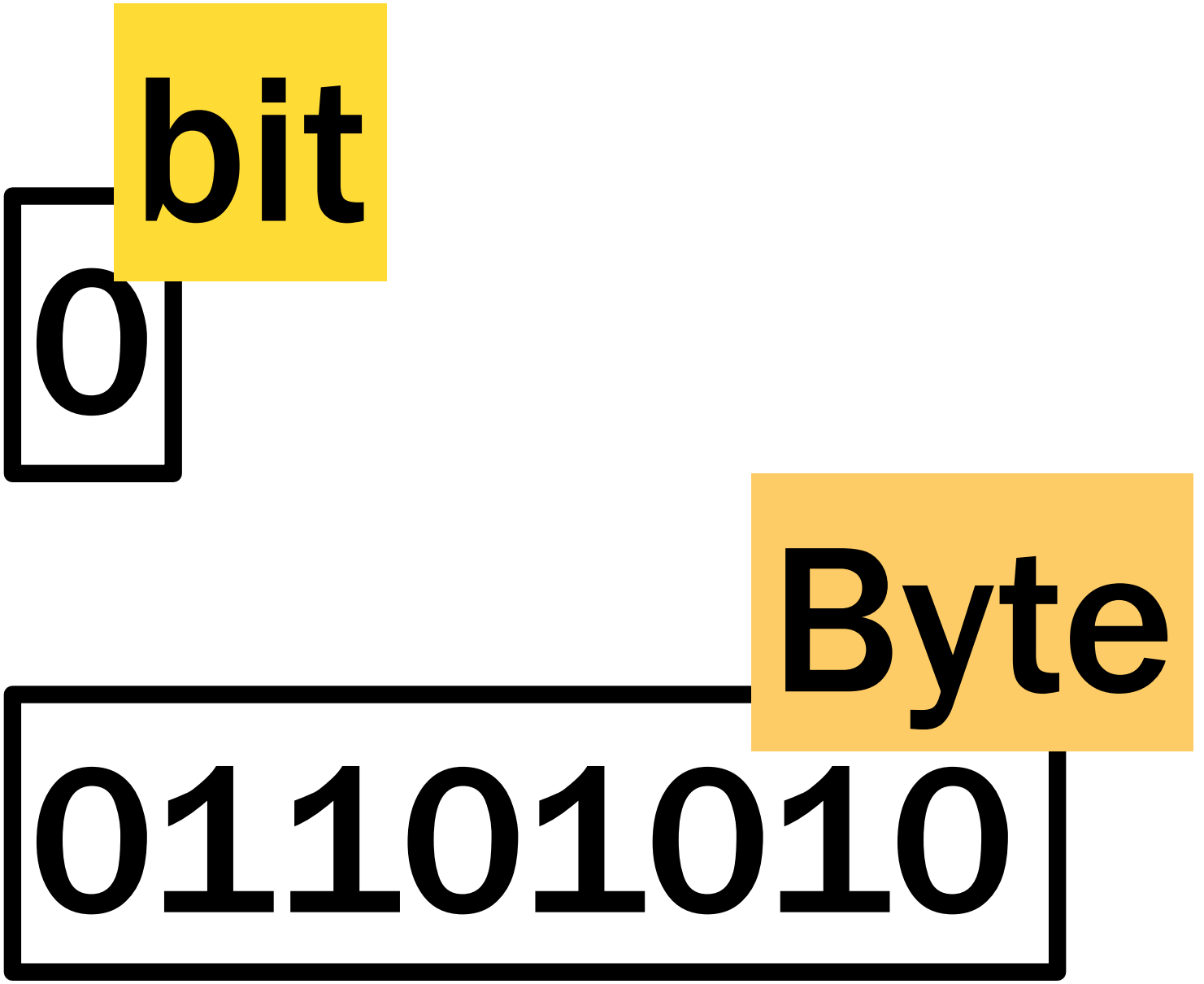
…but by a bit, that is, a binary digit (0 or 1).
Hence, the length of a word can be measured in bits.



Most common computers today contain memories that are 32 or 64 bit in width.
Such width determines the size of a single piece of data that the computer can transfer from and to its memory.



Don’t forget that a bit is the smallest data size: it is a 0 or a 1. A sequence of 8 bits forms a byte. Bits are indicated with ‘b’, bytes with ‘B’.

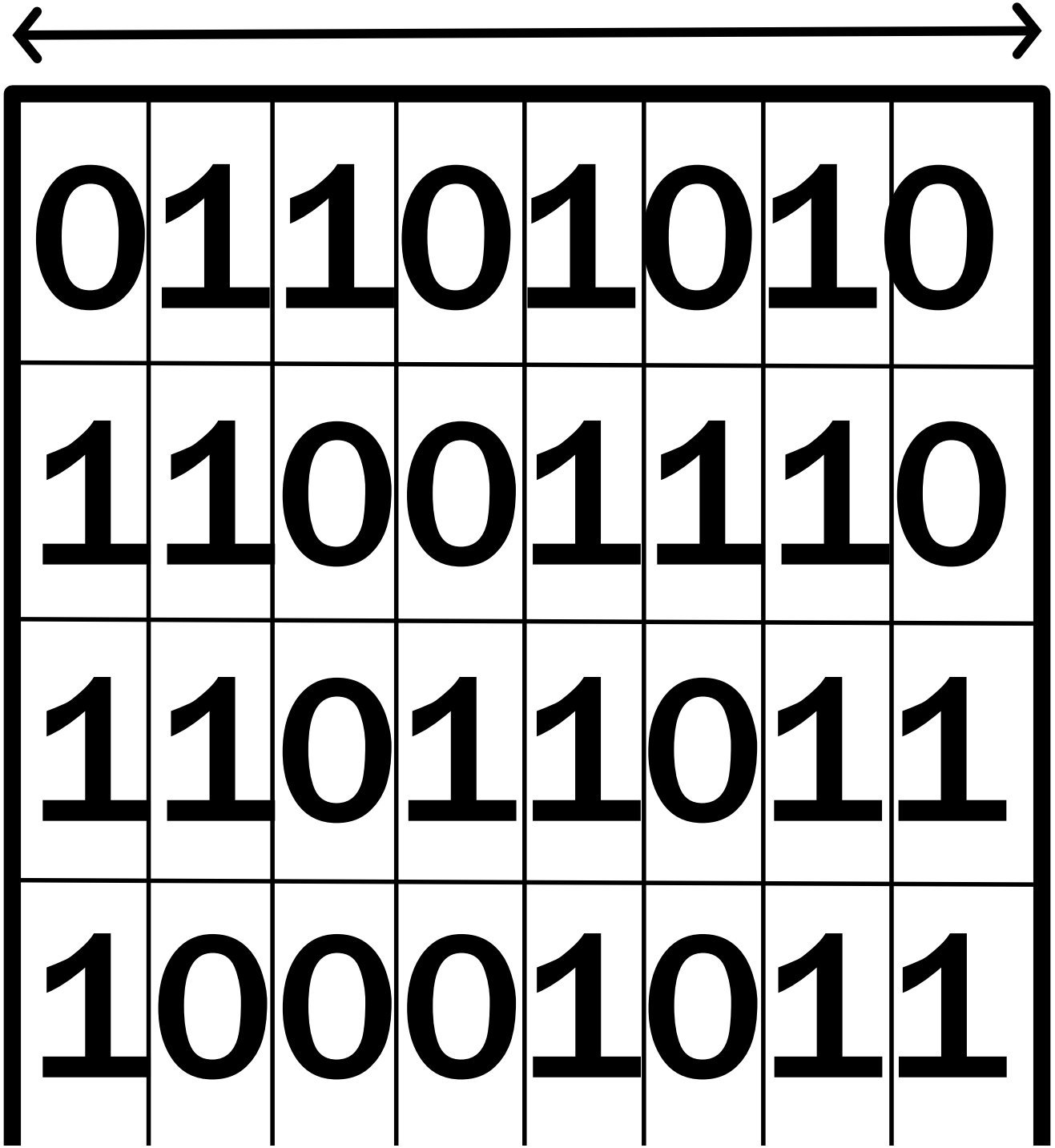


Why is 1 Byte made of 8 bits?

Choices.

Federico Faggin, Marcian “Ted” Hoff Jr., and Stanley Mazor
with the pioneering microprocessor they created in the early 1970s, the Intel 4004

Now that we have organized the memory as a matrix of places for bits…



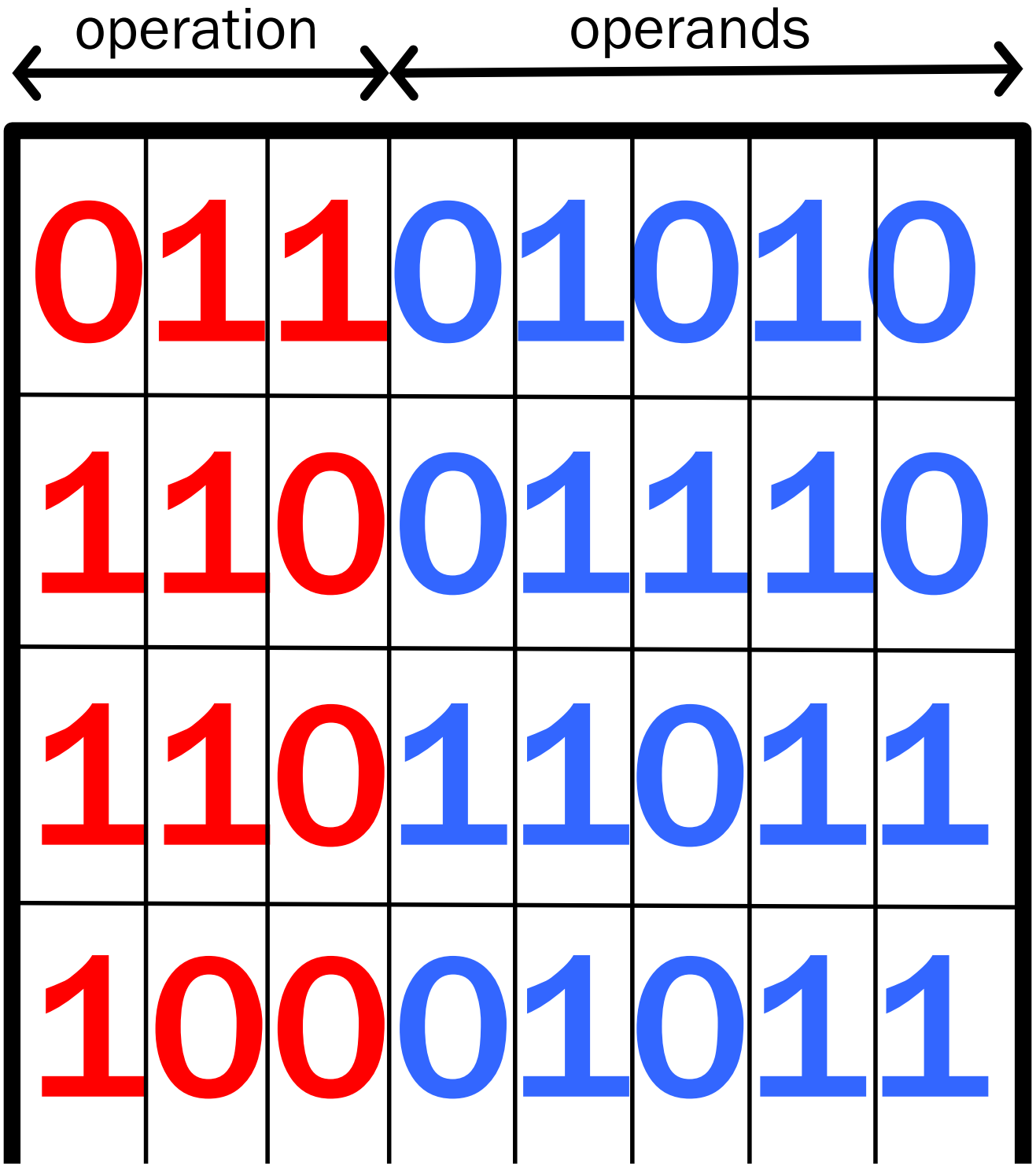
…how do we distinguish bits representing operations from bits representing operands?

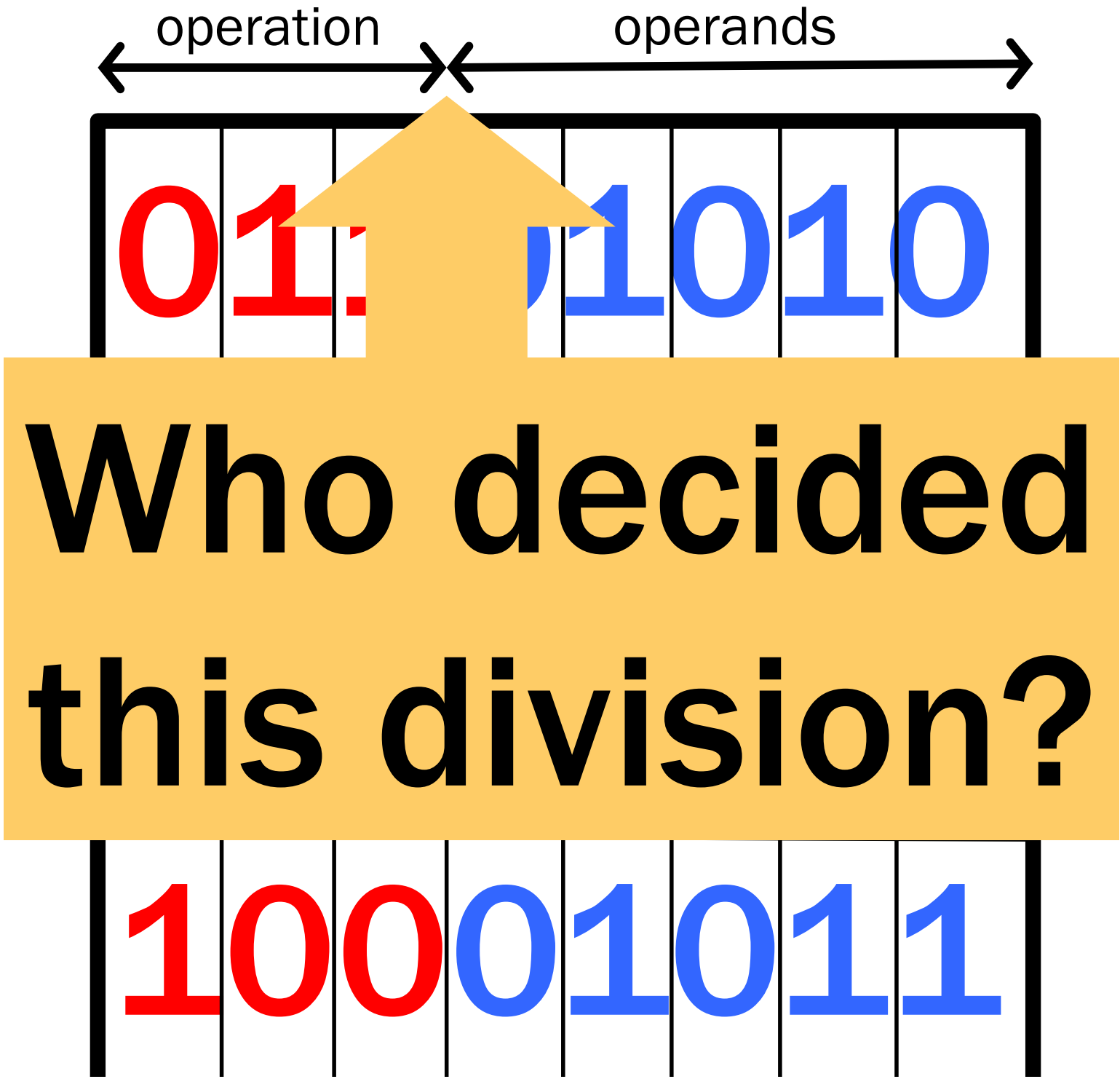
operations

operands

352

By dividing words into an operation section and an operands section.

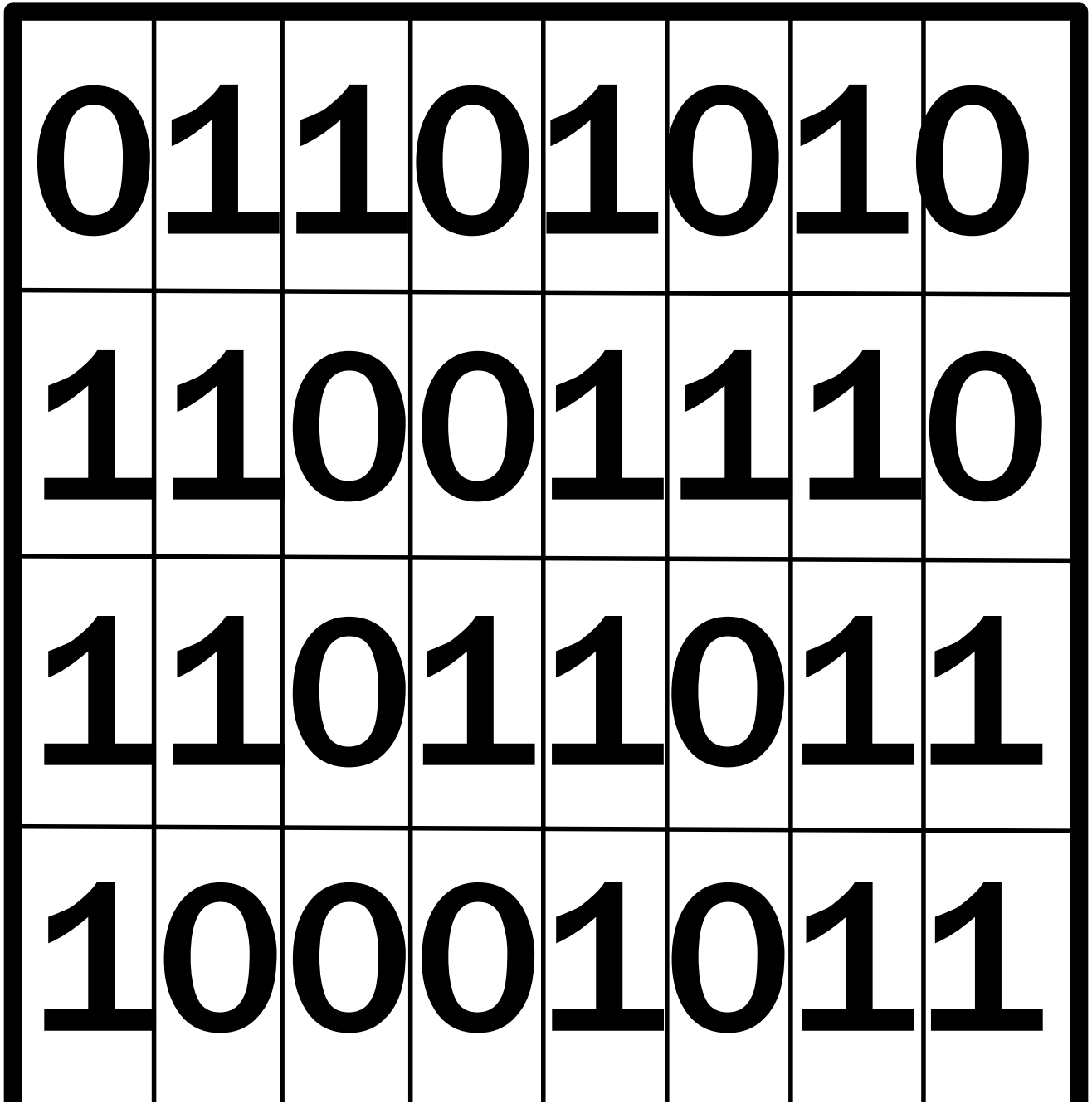



operation
operands
01
1010
Who decided this division?
10001011

Again, choices.

The position of the division is arbitrary. It is all about programming the computer in a way

All bits inside the memory are the same from a physical persepctive, but depending on where they are, they are used in different ways.



THE STORED PROGRAM

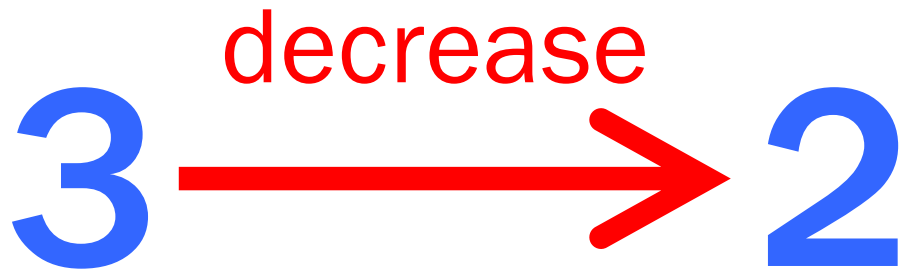
Both operations and operands can be stored in the same place.

THE STORED PROGRAM

Both operations and operands are bits stored inside words.

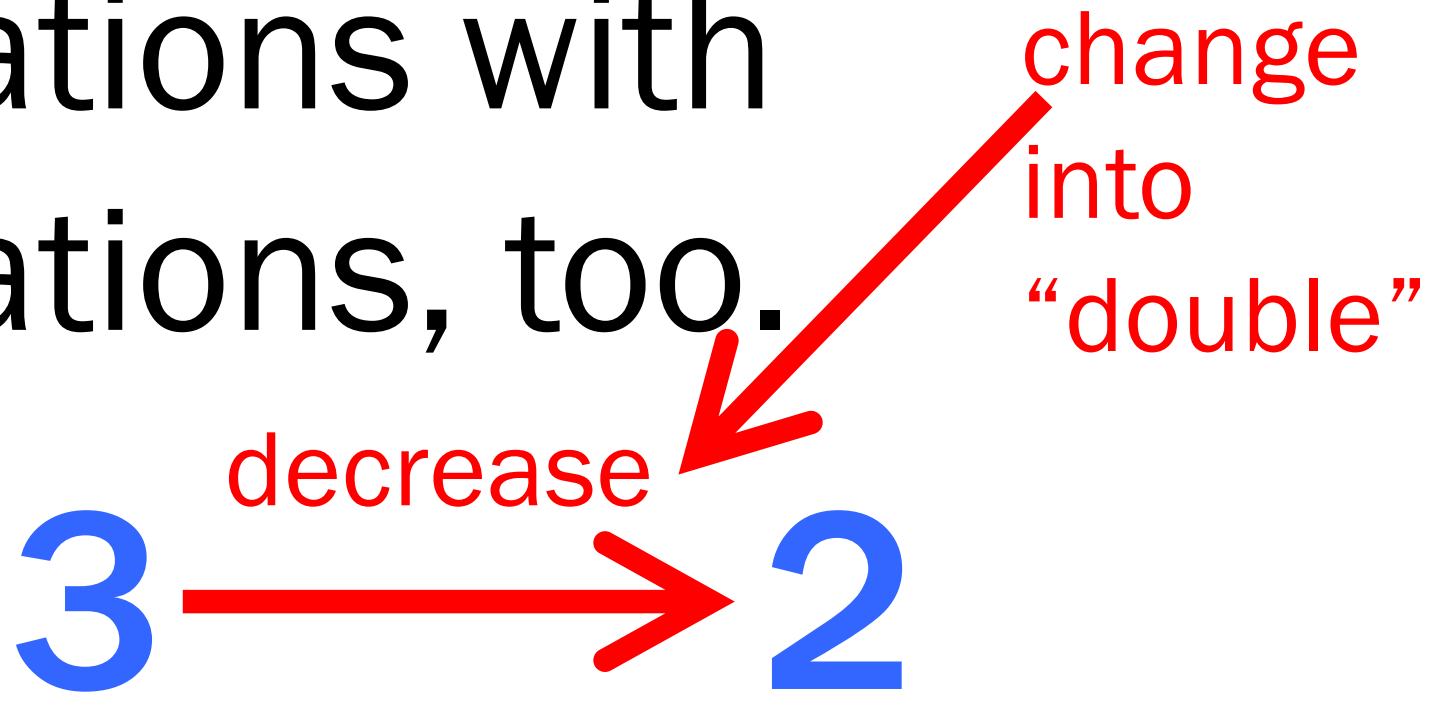
THE STORED PROGRAM

We manipulate operands with operations.



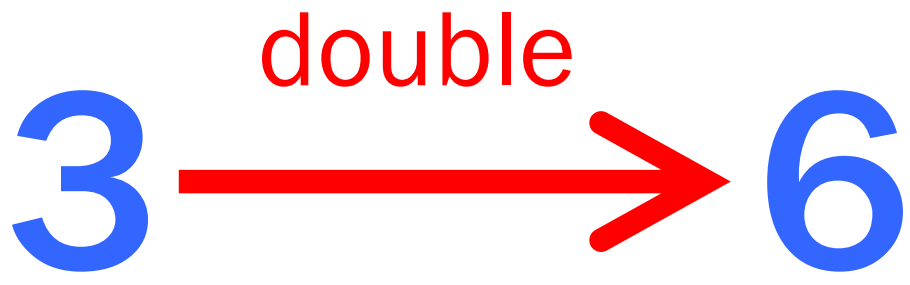
THE STORED PROGRAM

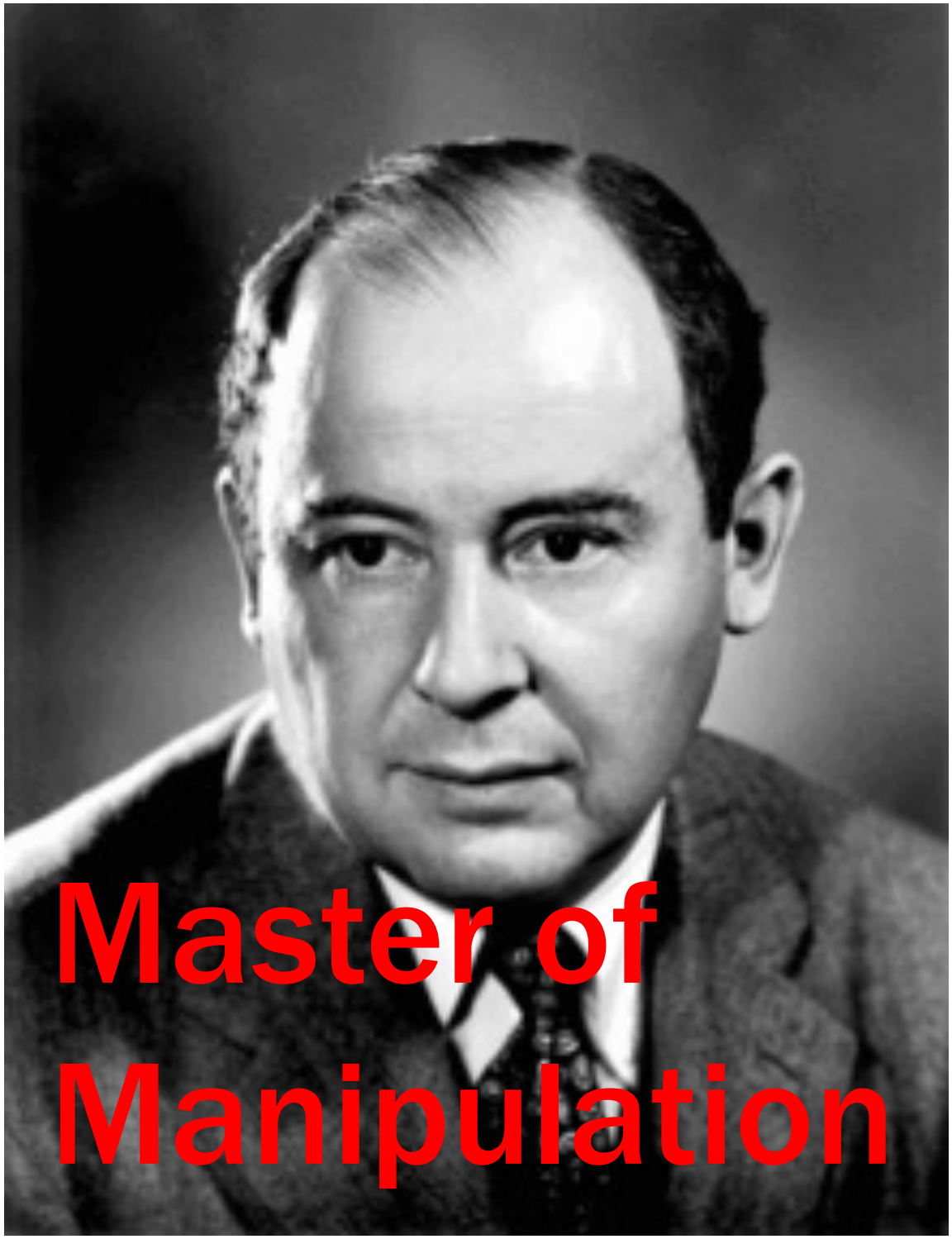
We can manipulate operations with operations, too.



THE STORED PROGRAM

We can manipulate operations with operations, too.




Master of
Manipulation

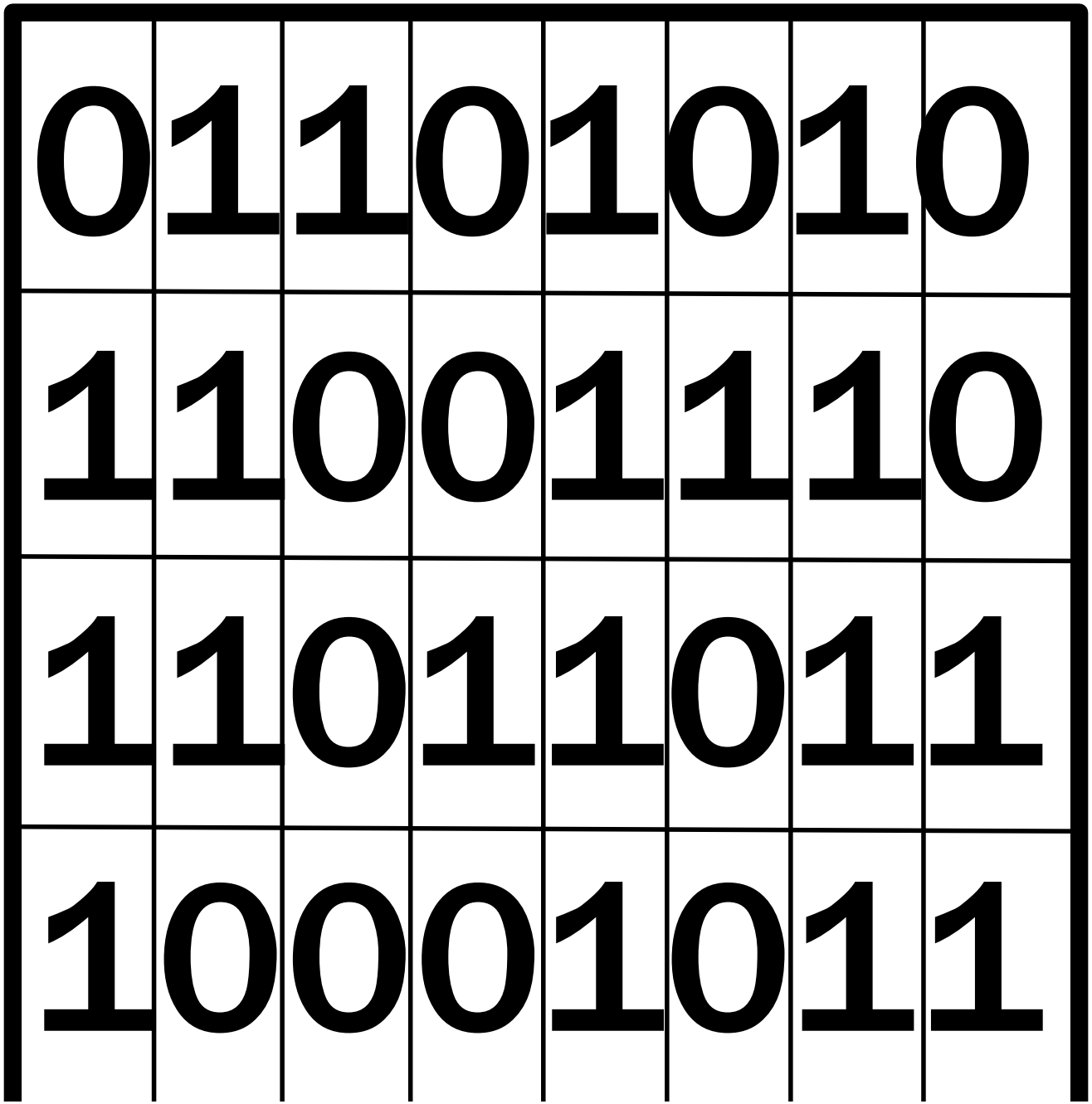
We still have to deal with the issue of knowing where these operations and operands are placed inside the memory.
We know, inside a word, which part is occupied by the operation and which part by the operands, but we still need a way to express where that word is inside the memory.

operations

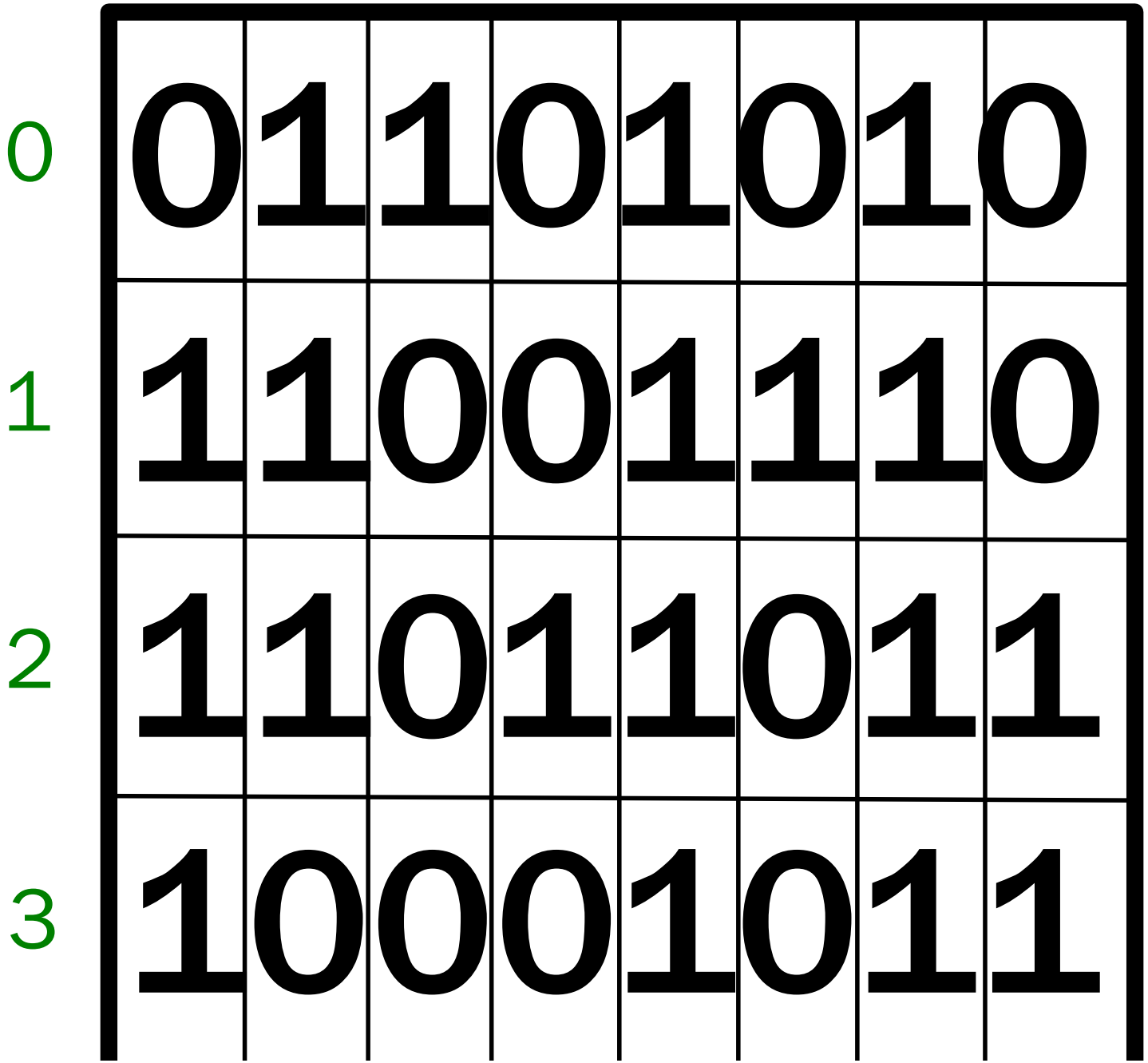
operands

place 352

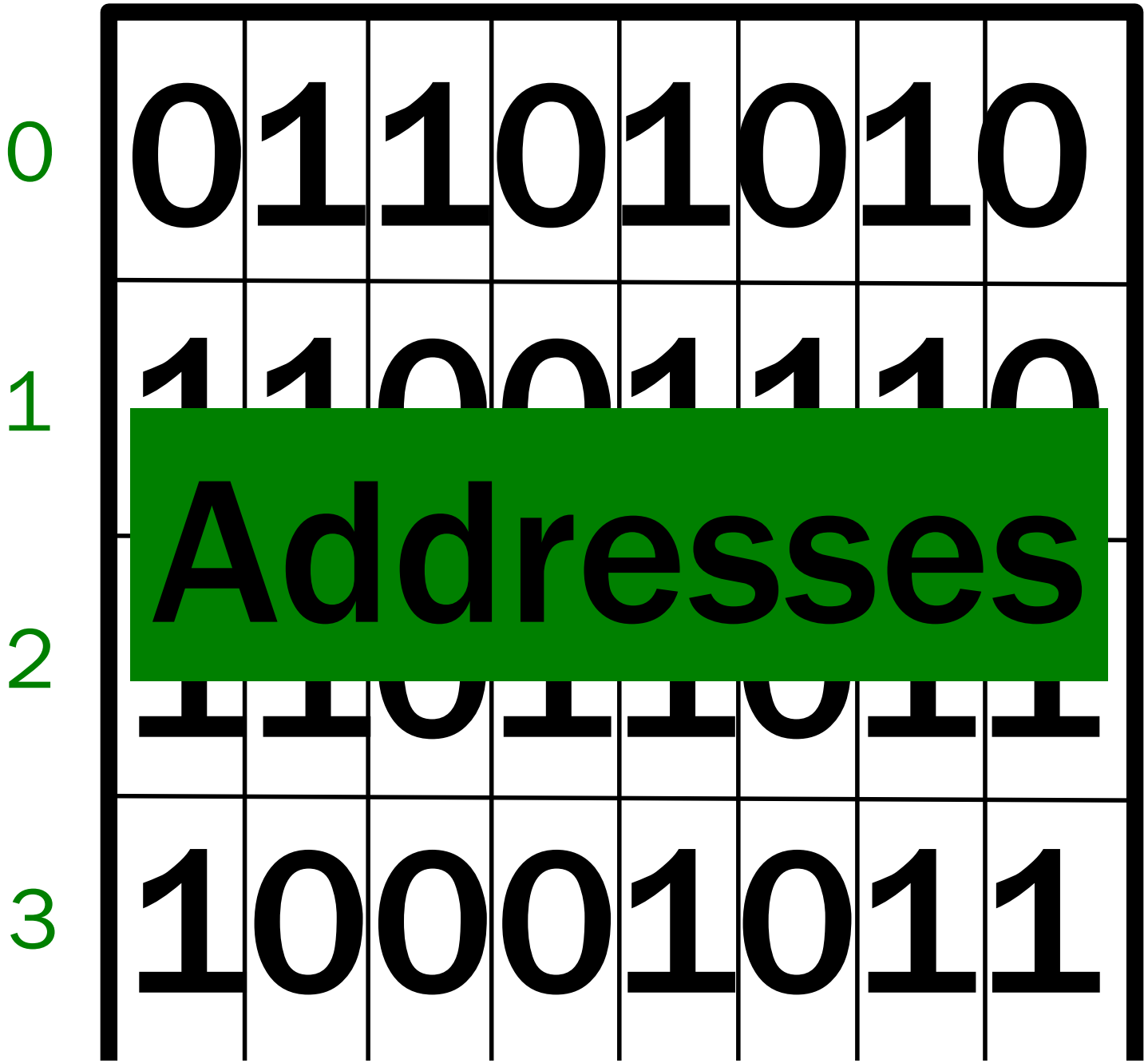
We need a way to refer to each word in the memory.



We rely on the indexes of the matrix. Since rows are deal with as words, we can focus on the row indexes and forget about the column indexes.



Since these indexes work like addresses of each word, we call them so.
Each word has an address.

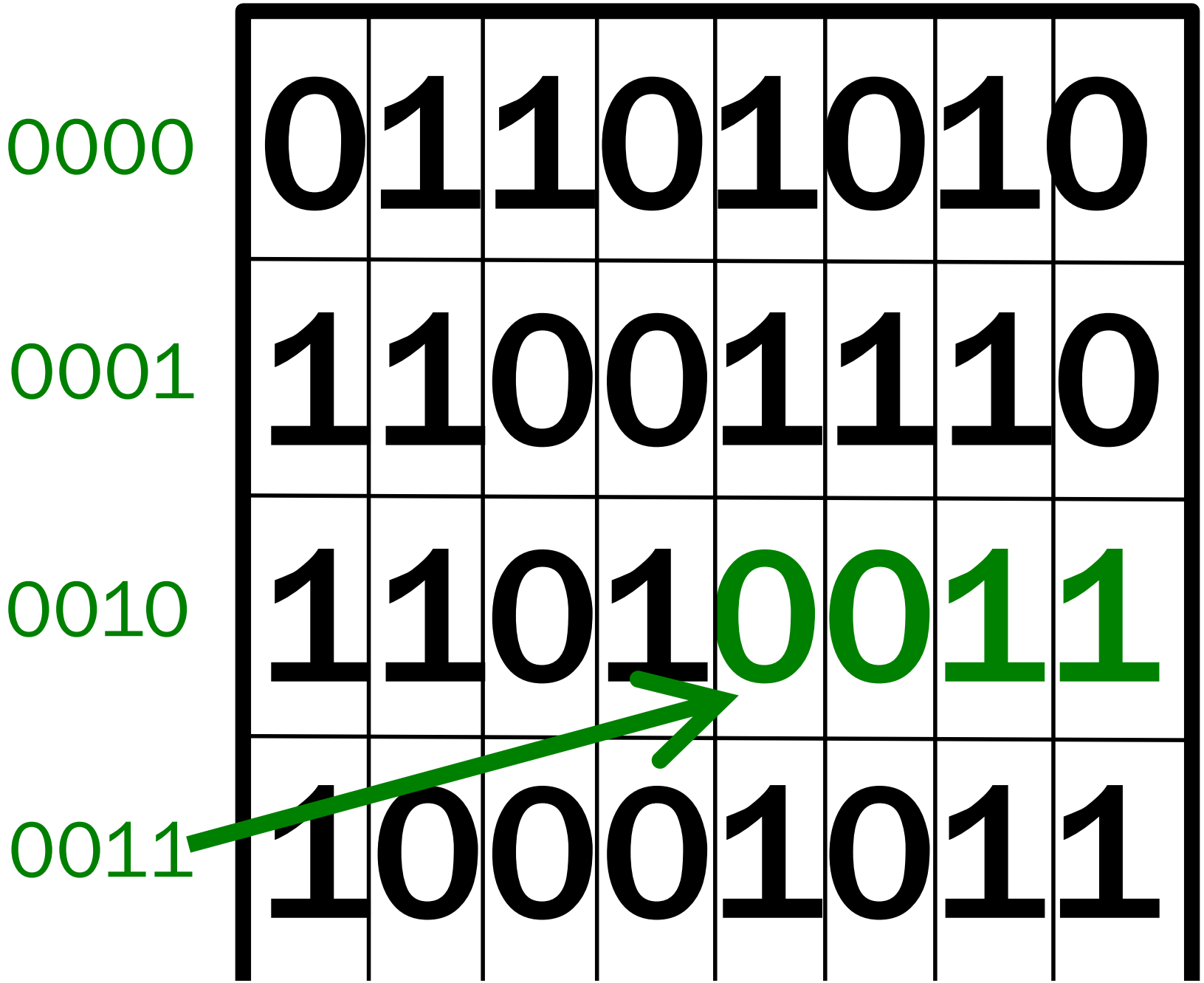


Depending on how many bits we use for each address, the number of words we can assign an address to varies. If we use 4 bit (like in this slide), we have a range of addresses from 0000 to 1111. A total of 16 addresses, which means that the addressable memory can only have 16 words.

0000	01101010
0001	11001110
0010	11011011
0011	10001011

If we used 32 bits to compose addresses, then we could address up to 2^{32} words, that is, more than 4 billion words.

Since addresses are also bits, they can be stored inside the memory as data, and become a parameter of some operations, like “assign 5 to x”, where x is the address of the destination word.



Hence, all sorts of data can be stored inside the memory: operations to perform, operands to perform operations to, and the addresses of the words where these operations and these operands are written.

operations

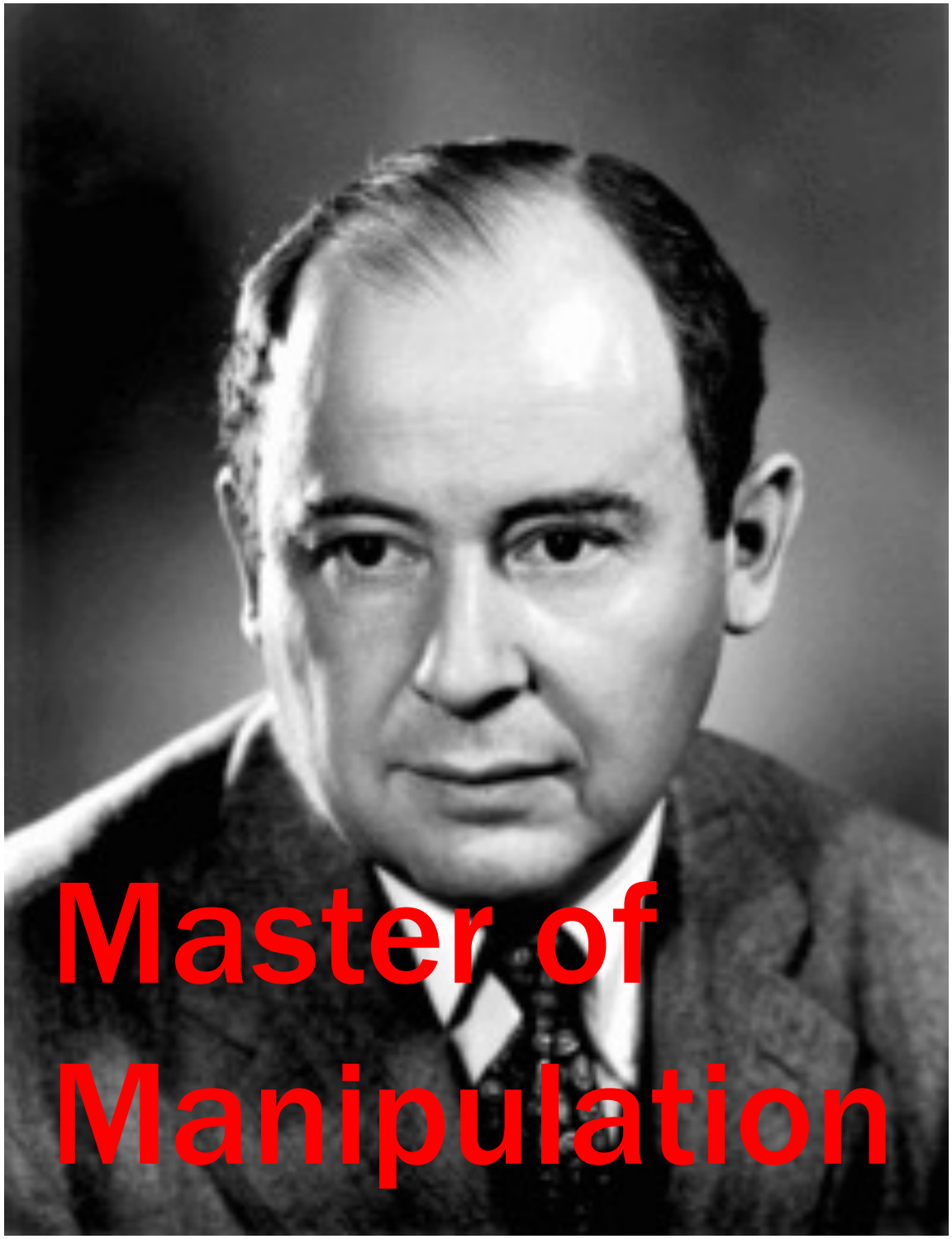
operands

place 4527

THE STORED PROGRAM

- We manipulate operands.
- We manipulate operations.
- We manipulate addresses.

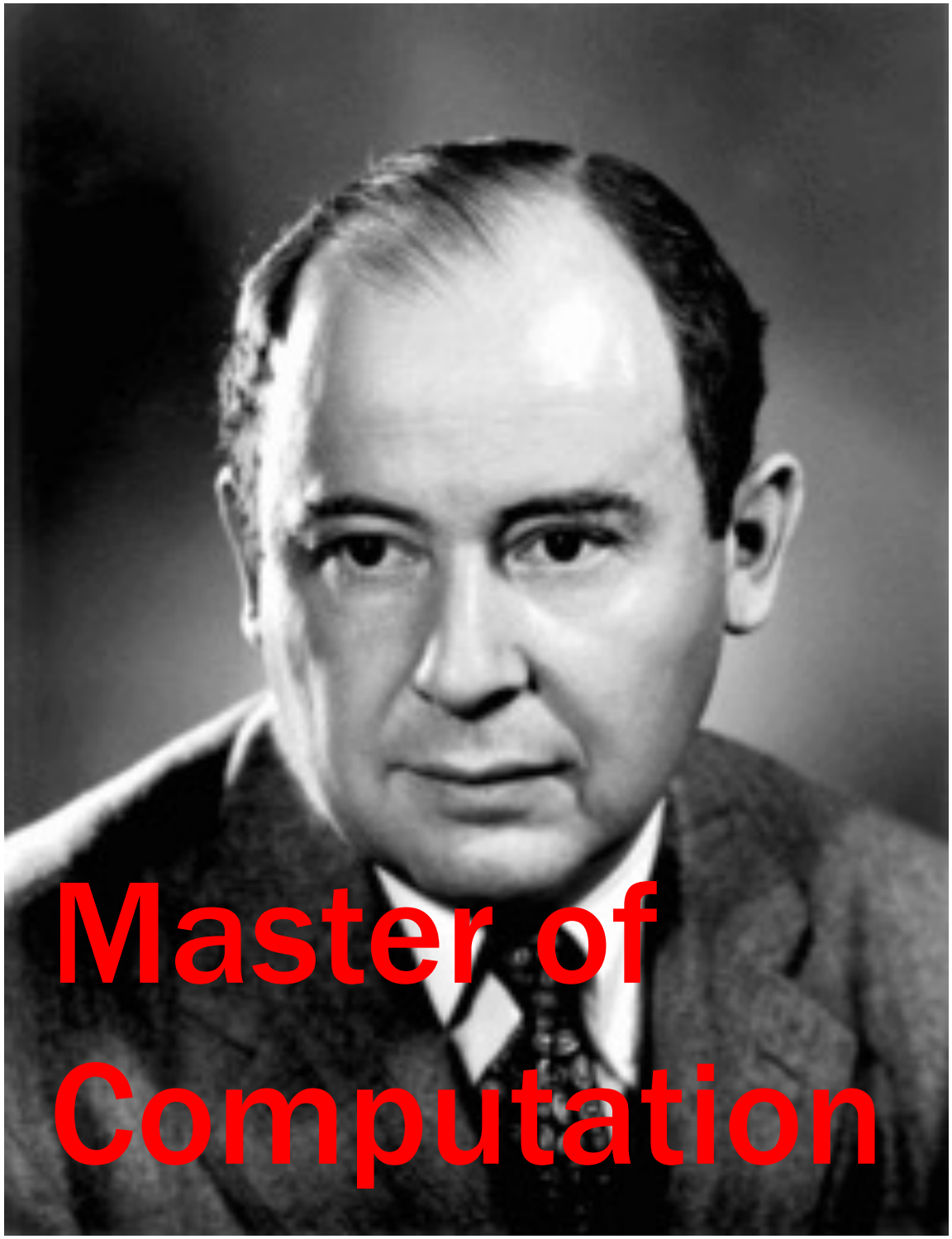
The paradigm of the stored program enables us to manipulate all these items, because they are all bits inside the memory.


Master of
Manipulation

THE STORED PROGRAM

In other, more abstract terms…

- We elaborate data.
- We create and modify programs.
- We transfer data and programs.



(Still
an awful
person.)

Speaking of bits stored in the memory, here are two examples of items stored in the memory of my computer.

They are both items stored in the memory, but they are different from each other, despite being strongly connected.

The Digital in
Digital A...ion.docx

PROGRAM

DATA

They are both files, but not of the same kind.

PROGRAM DATA

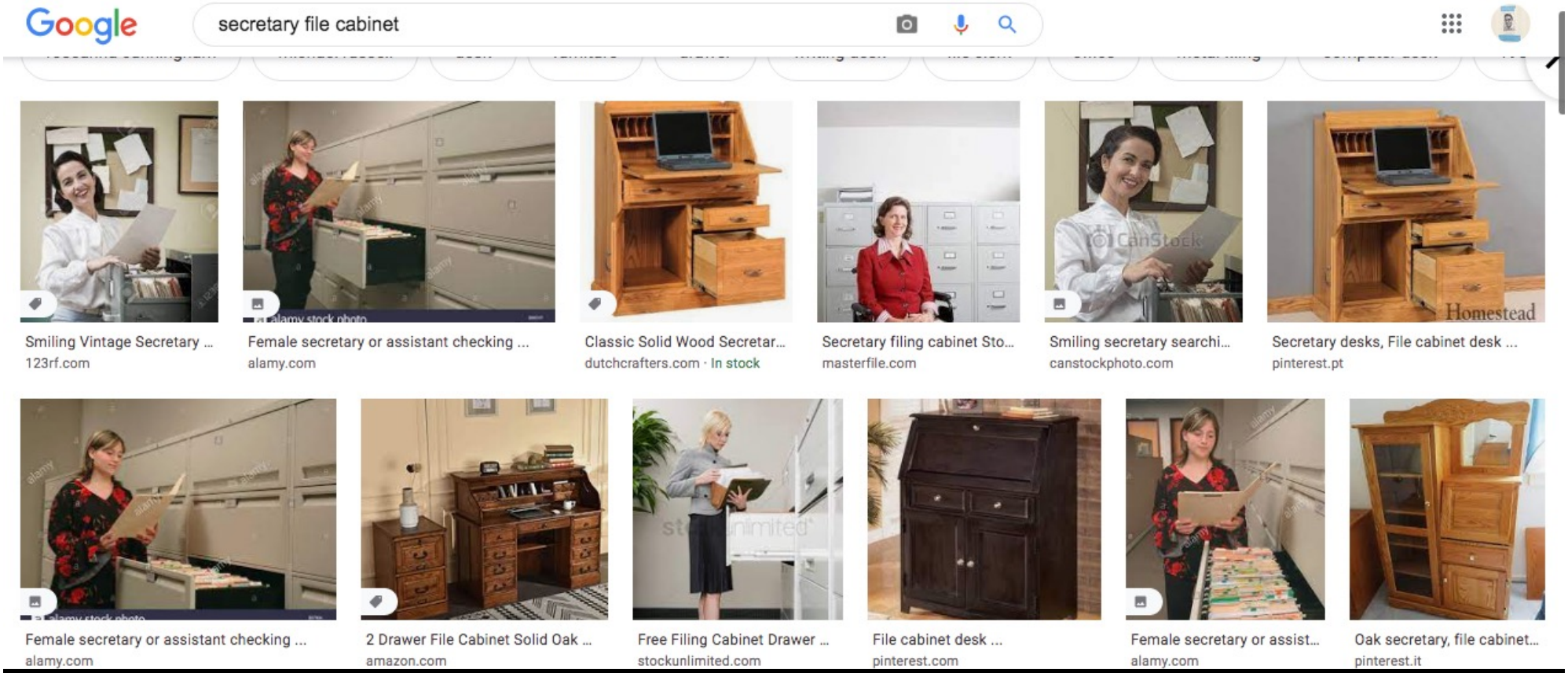
FILES

"File" is a term we borrow from real life, inside an office.

A file cabinet.

Here is a person who manages the file cabinet in an office.

The image is one of many results of an image search with the keywords “secretary file cabinet.”



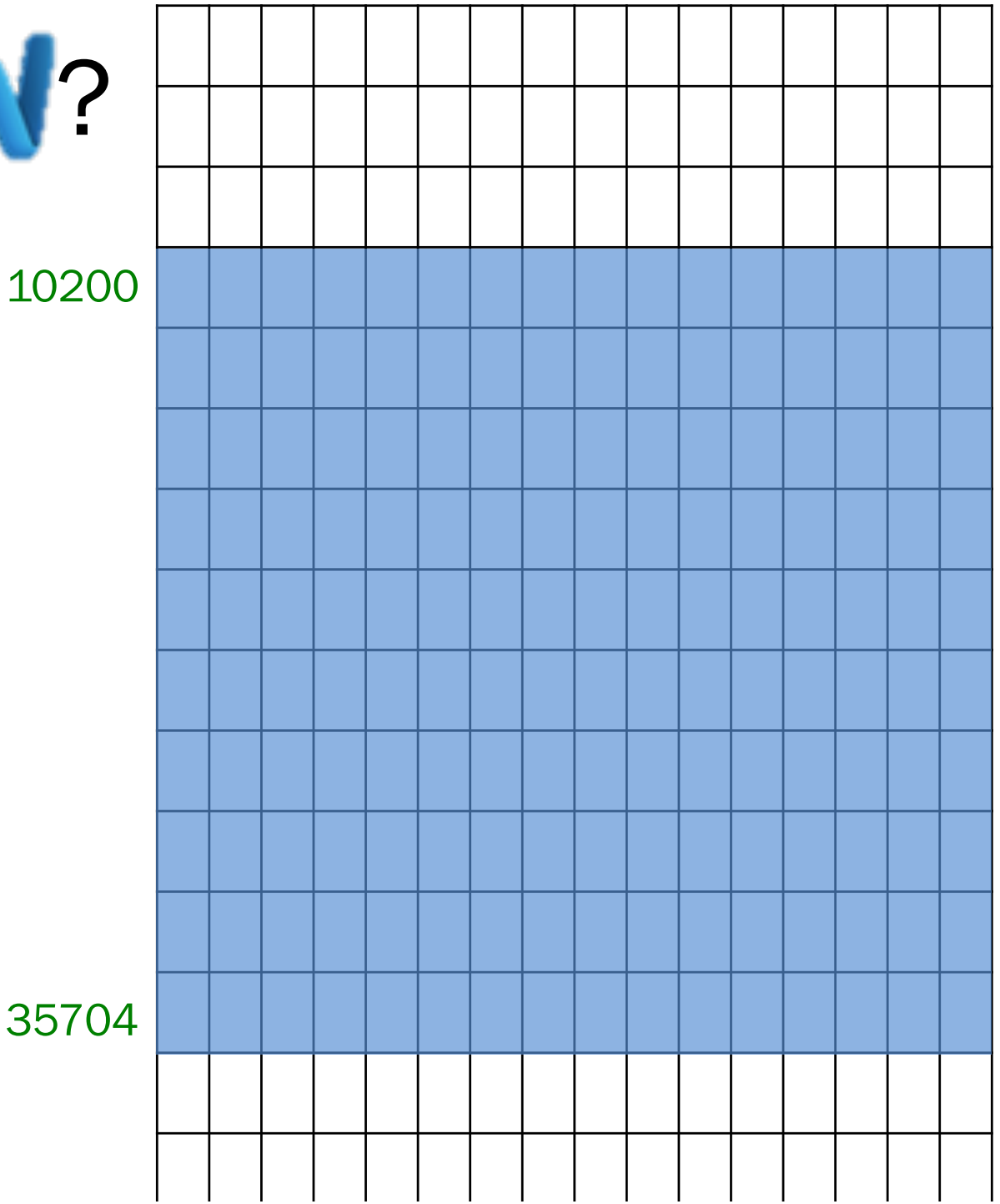
If you notice a stereotype in the images, do not blame the search technology: it just gives you what is available online. If what is online is problematic and search technology makes us keep on working with what is online, then the problem gets amplified and extended by Information Technology (this is a foreshadowing of current and future problems between society and technology).

Where is  ?

How are files managed inside a computer memory? A part of the operating system that manages your computer is called “file system”, and deals with storage and retrieval of files.

Where is W?

From word 10200 to word 35704.



Data files have extensions (the letters after the '.') that tell the operative system what kind of program files need to be launched to work with those data.

11 Save Your Tears.m4a

enso.JPG

The Digital in Digital A...ion.docx

PROGRAM DATA

FILES

FILES

- A file is a group of bits that are logically treated as a unit.
- A file may be comprised of data, program instructions, or addresses.

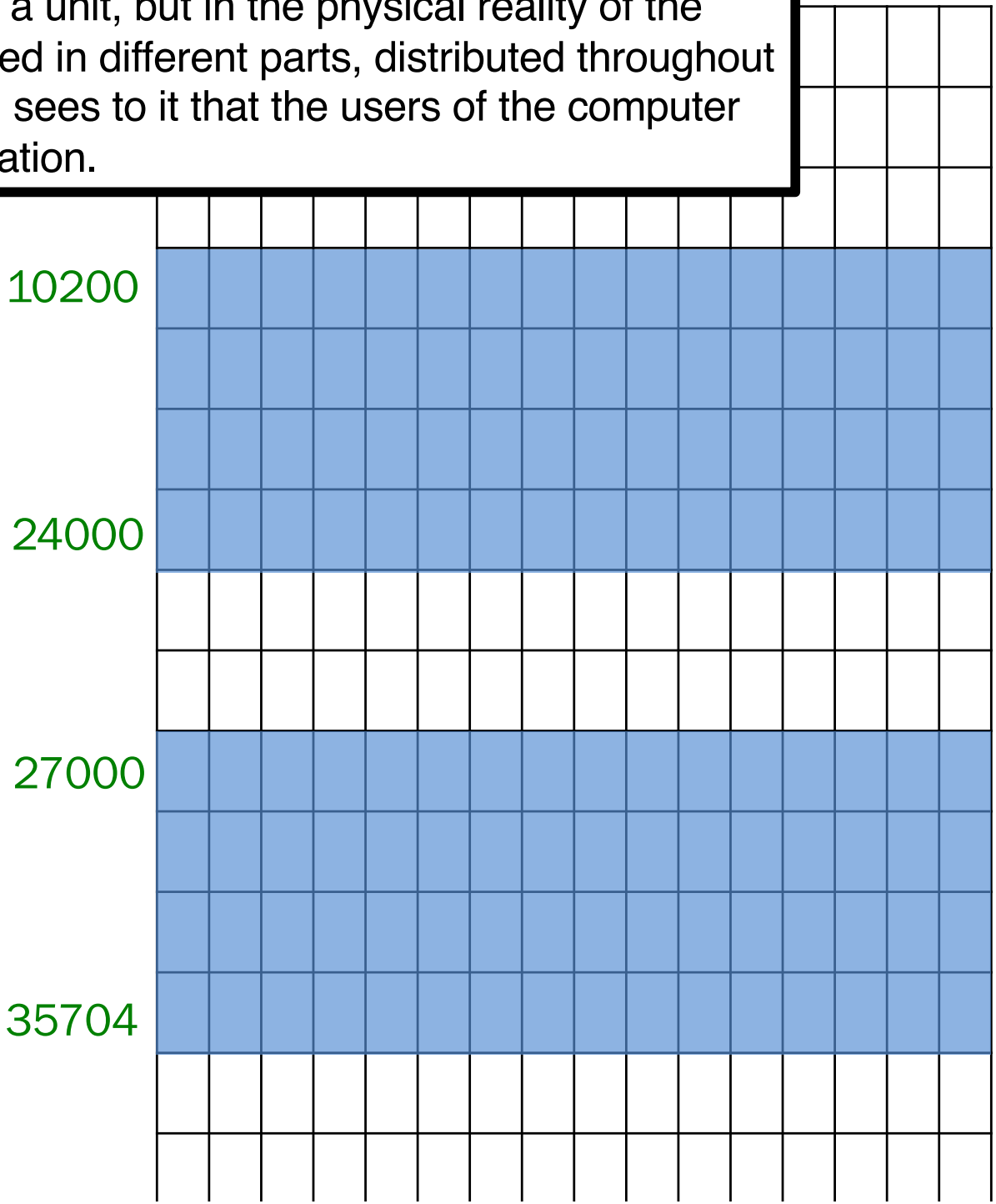
FILES

- A file is a group of bits that are logically treated as a unit.
- A file may be comprised of data, program instructions, or addresses.

"Logical" in the context of files does not mean that it is connected to solid and elemental reasoning; it means "abstract" as opposed to "physical".

A file is “logically” treated as a unit, but in the physical reality of the memory, it may be fragmented in different parts, distributed throughout the memory. The file system sees to it that the users of the computer do not notice this fragementation.

From word 10200 to word 24000 and from word 27000 to 35704.



FILES

- A file is a group of bits that are logically treated as a unit.
- A file may be comprised of data, program instructions, or addresses.

"My_thesis.docx" is a file of data.
Microsoft Word is a file of instructions.
What is a file of addresses? It is actually a very familiar thing.

The folders that the File System of your computer shows you, are actually files containing addresses of other files (the ones that are metaphorically “contained” in the folder).

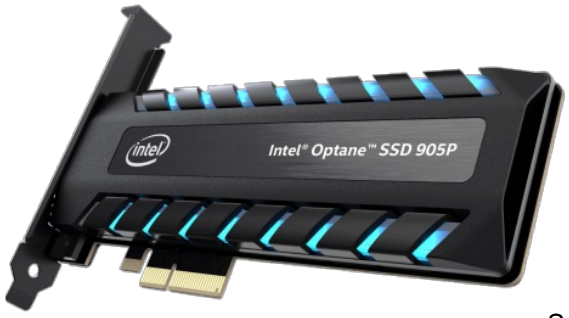
Digital Humanities
@UniBG

FOLDER

This way of organizing bits applies to all memory devices inside your computer, or that you can insert into or connect to your computer.

Digital Memory Devices

USB key



Solid State Disk

RAM

CD/DVD

SD card

Magnetic
Hard Disk