

# Information Communication Technologies

## Lecture 2. Data Representation in a Computer System

Kassymova Aizhan Bakhytzhanovna

PhD, Associate professor

**[a.kassymova@satbayev.university](mailto:a.kassymova@satbayev.university)**



# Lecture 2

---

## 1. Data Representation in a Computer System

1.1 Bits and Bytes

1.2 Number Systems

Text, numbers,  
music, images,  
speech, and  
video are  
distilled down to  
simple pulses of  
electricity and  
stored as 0s and  
1s





---

**Data** refers to the symbols that represent people, events, things, and ideas.

**Data** can be a name, a number, the colors in a photograph, or the notes in a musical composition.

# Data representation

- **Data representation** refers to the form in which data is stored, processed, and transmitted.
- Data can be represented electronically with electrical components being on or off.
- On and off states can be represented using digits 0s and 1s:
  - 0 – off state
  - 1 – on state



# Data representation

- The 0s and 1s are binary digits
- A **bit** is a 0 or 1 used in the digital representation of data



# Representing numbers

Decimal (Base 10)	Binary (Base 2)
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010
11	1011
1000	1111101000

- **Numeric data** consists of numbers that might be used in arithmetic operations (age, price)
- Digital devices can represent numeric data using the binary number system, also called base 2

# Representing text



The diagram illustrates the binary representation of the text "HI!". Each character is shown with its corresponding 8-bit binary code below it. The letter 'H' is represented by 01001000, the letter 'I' by 01001001, and the exclamation mark '!' by 00100001. The characters are colored: 'H' is brown, 'I' is teal, and '!' is dark brown.

01001000 01001001 00100001

- **Character data** is composed of letters, symbols, and numerals that are not used in arithmetic operations (name, color)
- A digital computer uses a series of bits to represent letters, characters, and numerals



# ASCII (American Standard Code for Information Interchange)

00100000	>	00111110	\	01011100	z	01111010	ÿ	10011000		10110110	£	11010100	≥	11110010
00100001	?	00111111	]	01011101	ÿ	01111011	ÿ	10011001	n	10110111	F	11010101	≤	11110011
00100010	@	01000000	^	01011110	!	01111100	ÿ	10011010	ÿ	10111000	π	11010110	ƒ	11110100
00100011	A	01000001	_	01011111	>	01111101	ç	10011011	ÿ	10111001		11010111	J	11110101
00100100	B	01000010	`	01100000	~	01111110	£	10011100		10111010	†	11011000	÷	11110110
00100101	C	01000011	a	01100001	△	01111111	¥	10011101	ÿ	10111011	J	11011001	∞	11110111
00100110	D	01000100	b	01100010	Ç	10000000	℞	10011110	ÿ	10111100	γ	11011010	o	11111000
00100111	E	01000101	c	01100011	ü	10000001	f	10011111	ÿ	10111101	■	11011011	-	11111001
00101000	F	01000110	d	01100100	é	10000010	á	10100000	ÿ	10111110	■	11011100	-	11111010
00101001	G	01000111	e	01100101	â	10000011	í	10100001	ÿ	10111111	■	11011101	√	11111011
00101010	H	01001000	f	01100110	ä	10000100	ó	10100010	ÿ	11000000	■	11011110	∞	11111100
00101011	I	01001001	g	01100111	à	10000101	ú	10100011	ÿ	11000001	■	11011111	z	11111101
00101100	J	01001010	h	01101000	ä	10000110	ñ	10100100	ÿ	11000010	α	11100000	■	11111110
00101101	K	01001011	i	01101001	ç	10000111	ñ	10100101	ÿ	11000011	β	11100001	■	11111111
00101110	L	01001100	j	01101010	ê	10001000	ë	10100110	-	11000100	Γ	11100010	■	11111100
00101111	M	01001101	k	01101011	ë	10001001	ë	10100111	+	11000101	Π	11100011	■	11111101
00110000	N	01001110	l	01101100	è	10001010	è	10101000	†	11000110	Σ	11100100	■	11111100
00110001	O	01001111	m	01101101	ÿ	10001011	ÿ	10101001		11000111	σ	11100101	■	11111101
00110010	P	01010000	n	01101110	ÿ	10001100	ÿ	10101010	ÿ	11001000	μ	11100110	■	11111100
00110011	Q	01010001	o	01101111	ÿ	10001101	½	10101011	ÿ	11001001	τ	11100111	■	11111101
00110100	R	01010010	p	01110000	ñ	10001110	¼	10101100	ÿ	11001010	æ	11101000	■	11111000
00110101	S	01010011	q	01110001	ÿ	10001111	ÿ	10101101	ÿ	11001011	θ	11101001	■	11111001
00110110	T	01010100	r	01110010	é	10010000	«	10101110	ÿ	11001100	Ω	11101010	■	11111010
00110111	U	01010101	s	01110011	æ	10010001	»	10101111	=	11001101	δ	11101011	■	11111011
00111000	V	01010110	t	01110100	Æ	10010010	ÿ	10110000	ÿ	11001110	ω	11101100	■	11111000
00111001	W	01010111	u	01110101	ö	10010011	ÿ	10110001	ÿ	11001111	∞	11101101	■	11111001
00111010	X	01011000	v	01110110	ö	10010100	ÿ	10110010	ÿ	11010000	€	11101110	■	11111000
00111011	Y	01011001	w	01110111	ò	10010101		10110011	ÿ	11010001	π	11101111	■	11111001
00111100	Z	01011010	x	01111000	û	10010110	ÿ	10110100	ÿ	11010010	≡	11110000	■	11110000
00111101	[	01011011	y	01111001	ù	10010111	ÿ	10110101	ÿ	11010011	±	11110001	■	11110001

# Unicode

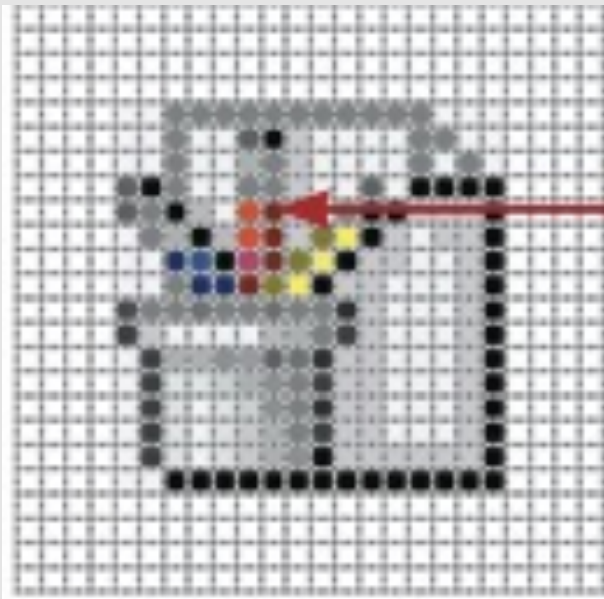
02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	
12	13	14	15	16	17	18	19	1A	1B	1C	1D	1E	1F	
22	23	24	25	26	27	28	29	2A	2B	2C	2D	2E	2F	
32	33	34	35	36	37	38	39	3A	3B	3C	3D	3E	3F	
42	43	44	45	46	47	48	49	4A	4B	4C	4D	4E	4F	
52	53	54	55	56	57	58	59	5A	5B	5C	5D	5E	5F	
62	63	64	65	66	67	68	69	6A	6B	6C	6D	6E	6F	
72	73	74	75	76	77	78	79	7A	7B	7C	7D	7E	7F	
82	83	84	85	86	87	88	89	8A	8B	8C	8D	8E	8F	
92	93	94	95	96	97	98	99	9A	9B	9C	9D	9E	9F	
A2	A3	A4	A5	A6	A7	A8	A9	AA	AB	AC	AD	AE	AF	
B2	B3	B4	B5	B6	B7	B8	B9	BA	BB	BC	BD	BE	BF	
C2	C3	C4	C5	C6	C7	C8	C9	CA	CB	CC	CD	CE	CF	
D2	D3	D4	D5	D6	D7	D8	D9	DA	DB	DC	DD	DE	DF	
E2	E3	E4	E5	E6	E7	E8	E9	EA	EB	EC	ED	EE	EF	
F2	F3	F4	F5	F6	F7	F8	F9	FA	FB	FC	FD	FE	FF	

- Latin scrip
- Linguistic
- Other Euro
- African sci
- Middle East Southwest
- Central Asi
- South Asi
- Southeast
- East Asiar
- Unified Cjt
- American
- Symbols
- Diacritics
- UTF-16 su private us
- Miscellane
- Unallocat

# EBCDIC (Extended Binary- Coded Decimal Interchange Code)

Most Sig. Bits V	Least Significant Bits															
	0 0000	1 0001	2 0010	3 0011	4 0100	5 0101	6 0110	7 0111	8 1000	9 1001	A 1010	B 1011	C 1100	D 1101	E 1110	F 1111
0 0000	NUL (0) 00	SOH (1) 01	STX (2) 02	ETX (3) 03	PF (4) 04	HT (5) 05	LC (6) 06	DEL (7) 07	GE (8) 08	RLF (9) 09	SMM (10) 0A	VT (11) 0B	FF (12) 0C	CR (13) 0D	SO (14) 0E	SI (15) 0F
1 0001	DLE (16) 10	DC1 (17) 11	DC2 (18) 12	TM (19) 13	RES (20) 14	NL (21) 15	BS (22) 16	IL (23) 17	CAN (24) 18	EM (25) 19	CC (26) 1A	CUI (27) 1B	IFS (28) 1C	IGS (29) 1D	IRS (30) 1E	IUS (31) 1F
2 0010	DS (32) 20	SOS (33) 21	FS (34) 22		BYP (36) 24	LF (37) 25	ETB (38) 26	ESC (39) 27			SM (42) 2A	CU2 (43) 2B		ENQ (45) 2D	ACK (46) 2E	BEL (47) 2F
3 0011			SYN (50) 32		PN (52) 34	RS (53) 35	US (54) 36	EOT (55) 37				CU3 (59) 3B	DC4 (60) 3C	NAK (61) 3D		SUB (63) 3F
4 0100	SP (64) 40										€ (74) 4A	.(75) 4B	< (76) 4C	( (77) 4D	+ (78) 4E	Note1 (79) 4F
5 0101	& (80) 50										! (90) 5A	\$ (91) 5B	* (92) 5C	) (93) 5D	; (94) 5E	¬ (95) 5F
6 0110	- (96) 60	/ (97) 61									 (106) 6A	, (107) 6B	% (108) 6C	_ (109) 6D	> (110) 6E	? (111) 6F
7 0111										~ (121) 79	: (122) 7A	# (123) 7B	@ (124) 7C	' (125) 7D	= (126) 7E	" (127) 7F
8 1000		a (129) 81	b (130) 82	c (131) 83	d (132) 84	e (133) 85	f (134) 86	g (135) 87	h (136) 88	i (137) 89						
9 1001		j (145) 91	k (146) 92	l (147) 93	m (148) 94	n (149) 95	o (150) 96	p (151) 97	q (152) 98	r (153) 99						

# Representing images

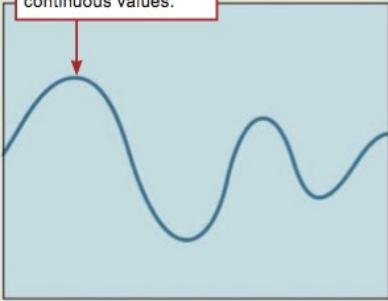


A red dot  
might be  
digitized  
as 1100.

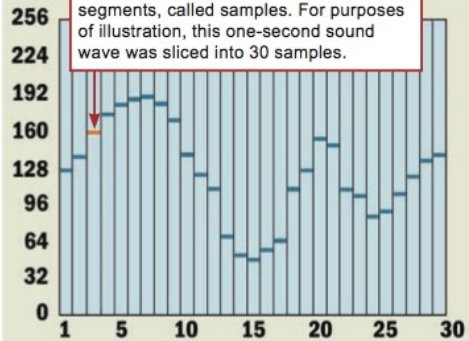
- A digital image is simply a list of color numbers for all the dots it contains

# Representing sound

An analog sound wave is a smooth curve of continuous values.




To digitize a wave, it is sliced into vertical segments, called samples. For purposes of illustration, this one-second sound wave was sliced into 30 samples.



**Sound, such as music and speech, is characterized by the properties of a sound wave.**

Bit	One binary digit	Gigabit	$2^{30}$ bits
Byte	8 bits	Gigabyte	$2^{30}$ bytes
Kilobit	1,024 or $2^{10}$ bits	Terabyte	$2^{40}$ bytes
Kilobyte	1,024 or $2^{10}$ bytes	Petabyte	$2^{50}$ bytes
Megabit	1,048,576 or $2^{20}$ bits	Exabyte	$2^{60}$ bytes
Megabyte	1,048,576 or $2^{20}$ bytes		

## Quantifying Digital Data

A thin, vertical black line is positioned to the left of the text.

End of Lecture 2