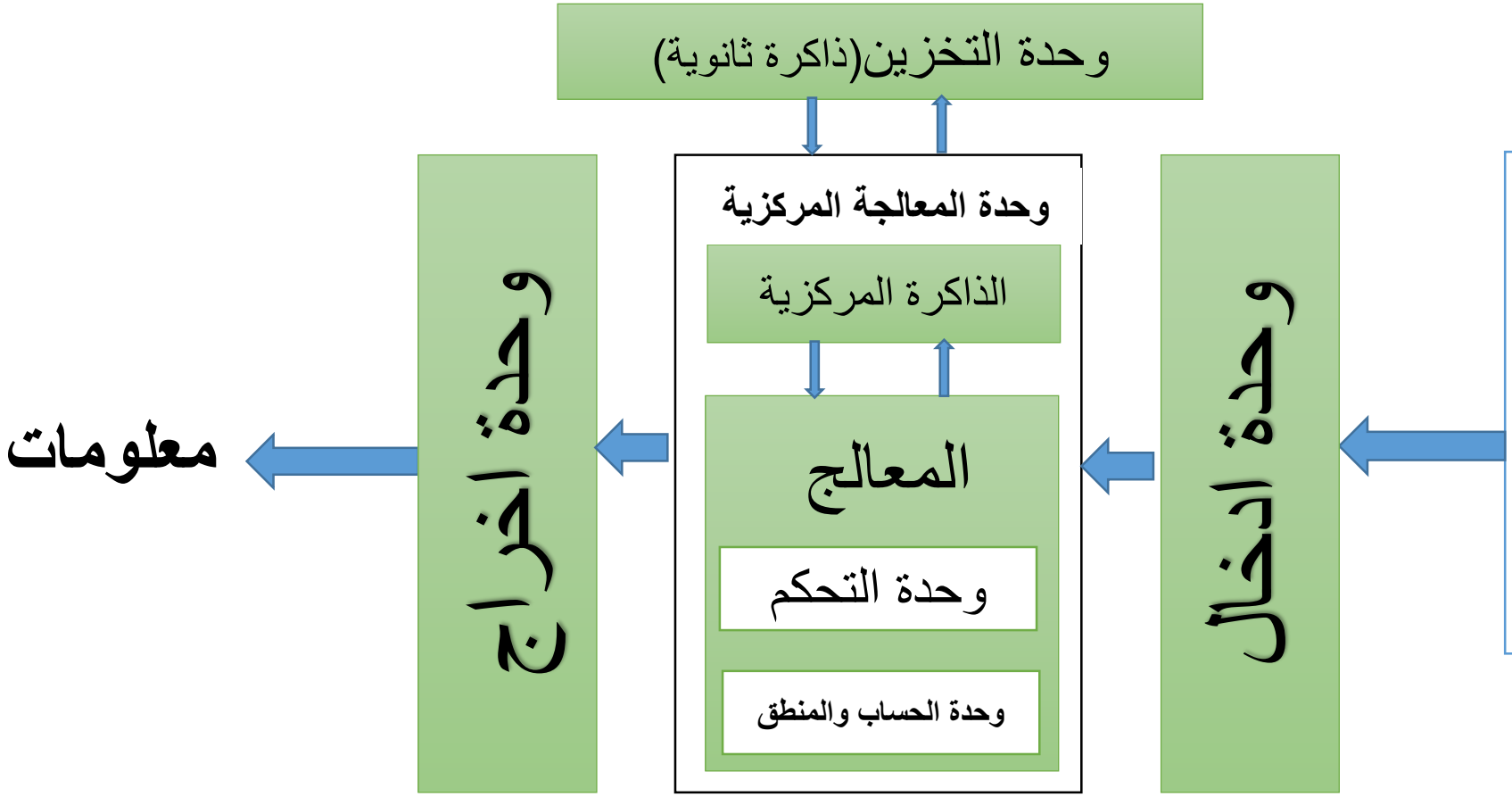
تمثيل البيانات

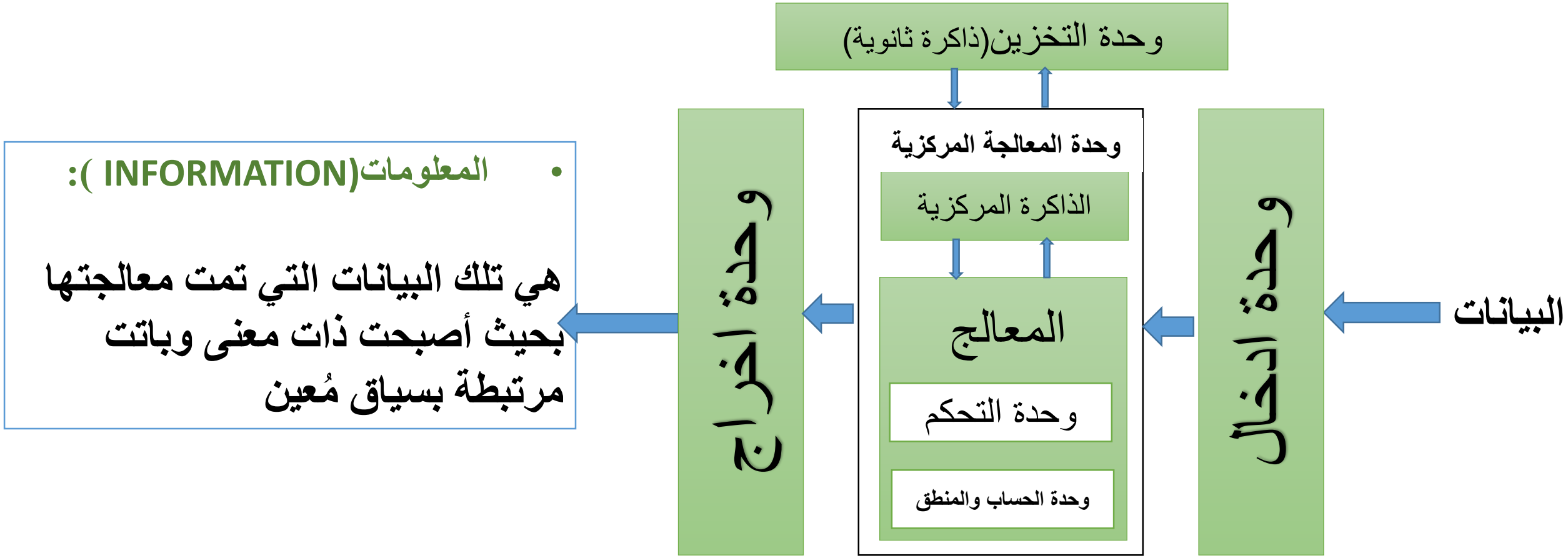
الأستاذة: رحماني سليمة
rahmani.prof.inf@gmail.com

البيانات والمعلومات

- **البيانات(DATA):**
  هي قيم أولية (حروف، كلمات، أرقام، رموز أو صور) يتمُّ تمثيلها في هيئة مُحَدَّدَة (Format) قابلة للمعالجة من طرف الحاسوب



البيانات والمعلومات

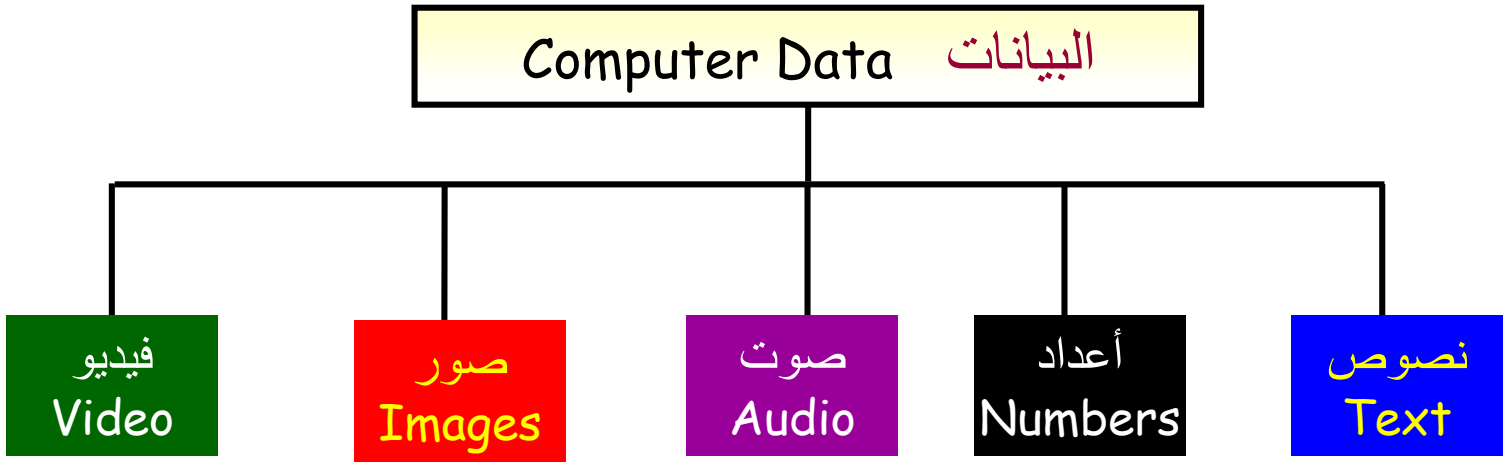


- المعلومات(INFORMATION):

هي تلك البيانات التي تمت معالجتها بحيث أصبحت ذات معنى وباتت مرتبطة بسياق مُعين

أنواع البيانات

- ❖ يعرف الحاسب بأنه آلة لمعالجة البيانات.
- ❖ لكن ما هي طبيعة البيانات وكيف يتم التعامل معها داخل الحاسب.
- ❖ تأخذ البيانات أشكالا مختلفة كالأعداد والنصوص والصور والصوت والفيديو.



هل يتعامل جهاز الحاسب مع البيانات بصورتها الأصلية كما هي ؟؟

.. لا طبعا لا يتعامل معها بصورتها الاصلية لان جهاز الحاسب آله الكترونيه تستخدم الطاقة الكهربائية في جميع تعاملاتها

البيانات داخل الحاسب

- كيف يتم التعامل مع كل هذه الأنواع من البيانات.
- لا يمكن أن يكون هناك حاسب يتعامل مع كل نوع من هذه الأنواع لأن هذا ليس عمليا ولا اقتصاديا.
- البنك مثلا يتعامل عادة مع أرقام إلا أنه يحتاج إلى تخزين بيانات العملاء كنصوص.
- الحل الأمثل هو أن نستخدم تمثيلا منسقا ومنتظما للبيانات
- كل بيانات من خارج الحاسب ينبغي أن تنقل إلى داخل الحاسب بهذا التمثيل المنسق عندما تخزن في الحاسب وكذلك عندما يعاد إخراجها من الحاسب يسمى هذا الأمر Bit pattern

البيانات داخل الحاسب(2)

البت Bit

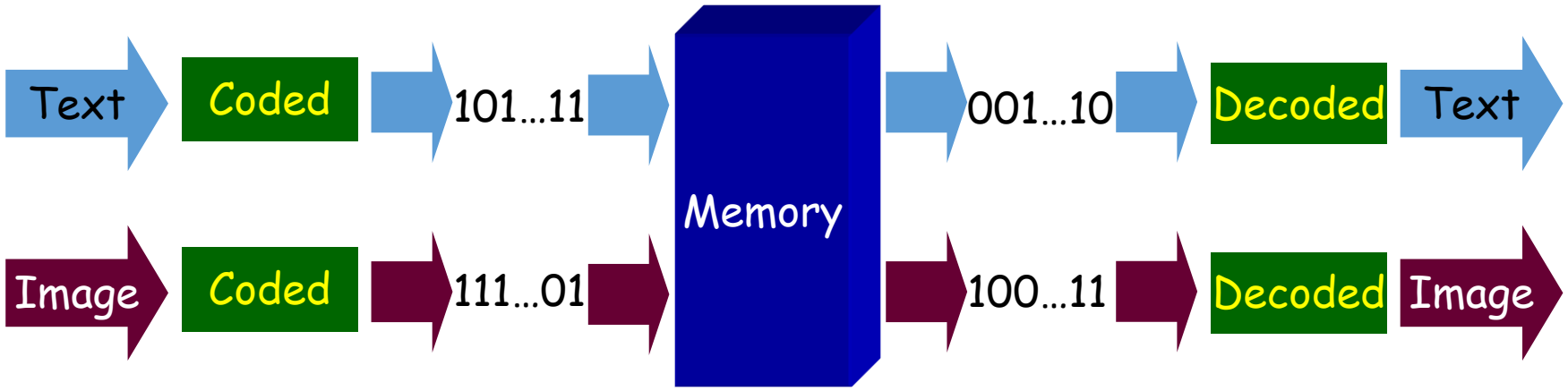
- البت Binary digit (bit) هو أصغر وحدة بيانات يمكن أن تخزن في الحاسب وهو عادة إما 0 أو 1.
- يمثل البت حالة جهاز يأخذ وضعين إما أن يكون شغال on (1) أو متوقف off (0) كمفتاح مغلق أو مفتوح.

النسق الثنائي Bit Pattern

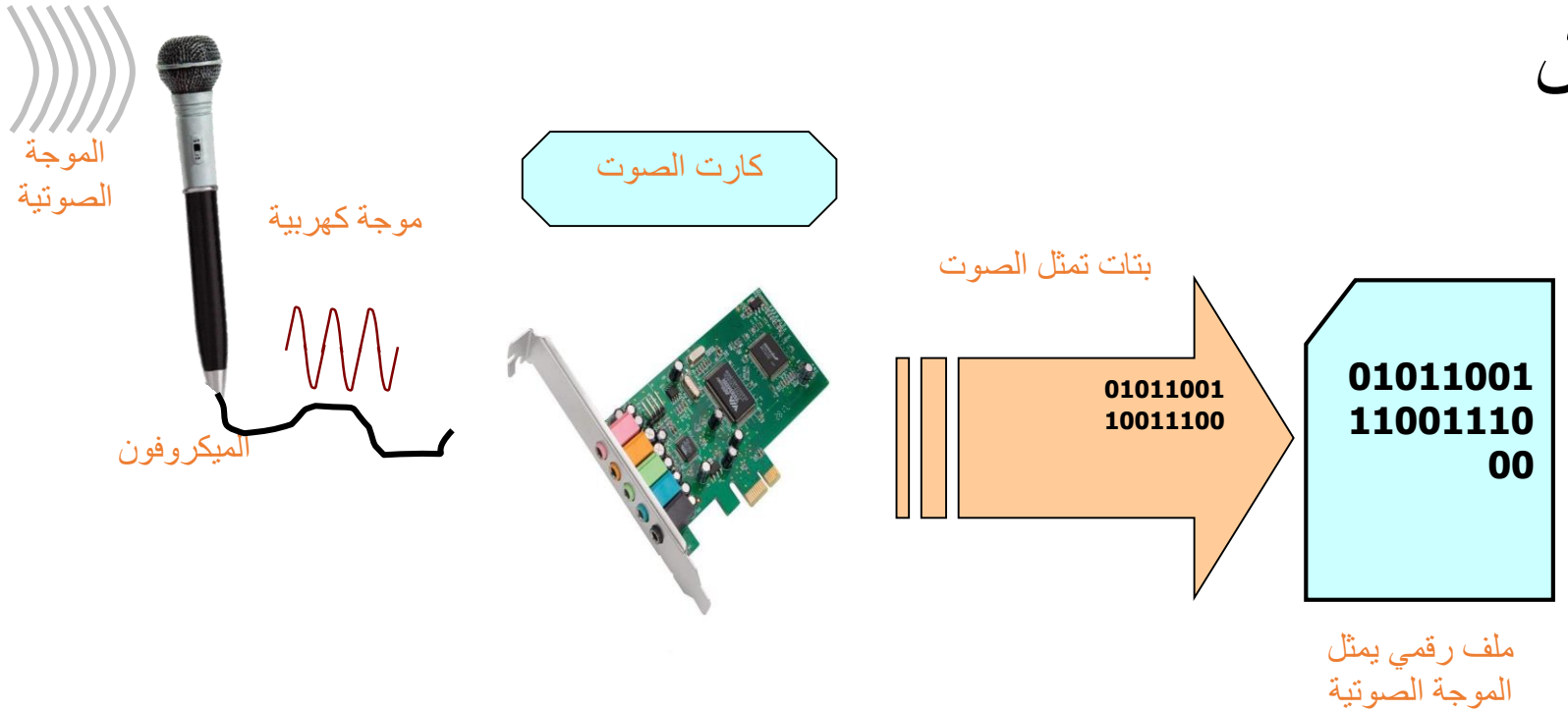
- البت الواحد لا يحل مشكلة تمثيل البيانات.
- لتمثيل الأنواع المختلفة من البيانات تستعمل سلسلة أو تتابع من البتات Sequence or Stream of bits يمكن أن تكون 1 أو 0.
- إذا كانت مجموعة البتات الممثلة للبيانات هي n فانه يمكن تمثيل عدد 2^n من البيانات المختلفة.
- لكن السؤال الآن هو كيف تستطيع ذاكرة الحاسب أن تعرف أي نوع من البيانات يتم تخزينه بها.

النسق الثنائي Bit Pattern

- الجواب أنه في الحقيقة هي لا تدري.. فهي فقط تقوم بتخزين البيانات كسلسلة بتات.
- ترمز البيانات وتشفر Coded عندما تدخل إلى الحاسب ويفك الترميز decoded عندما تقدم للمستخدم.



مثال


الموجة
الصوتية
موجة كهربية
الميكروفون
كارت الصوت
بتات تمثل الصوت
01011001
10011100
01011001
11001110
00
ملف رقمي يمثل
الموجة الصوتية

البايت Byte

- جرى العرف على تسمية كل 8-bit بالبايت.
- البايت هو وحدة قياس حجم ذاكرة الحاسب أو وحدة التخزين الثانوية.
- البايت يكفى لتمثيل حرف واحد Character أو رقم.
- مثلا عندما أكتب أنا أحب الحاسب.
- هذا يعد 12 حرفا + فراغين = 14 character.
- يحتاج ذلك سعة تخزين قدرها 14 Byte=.

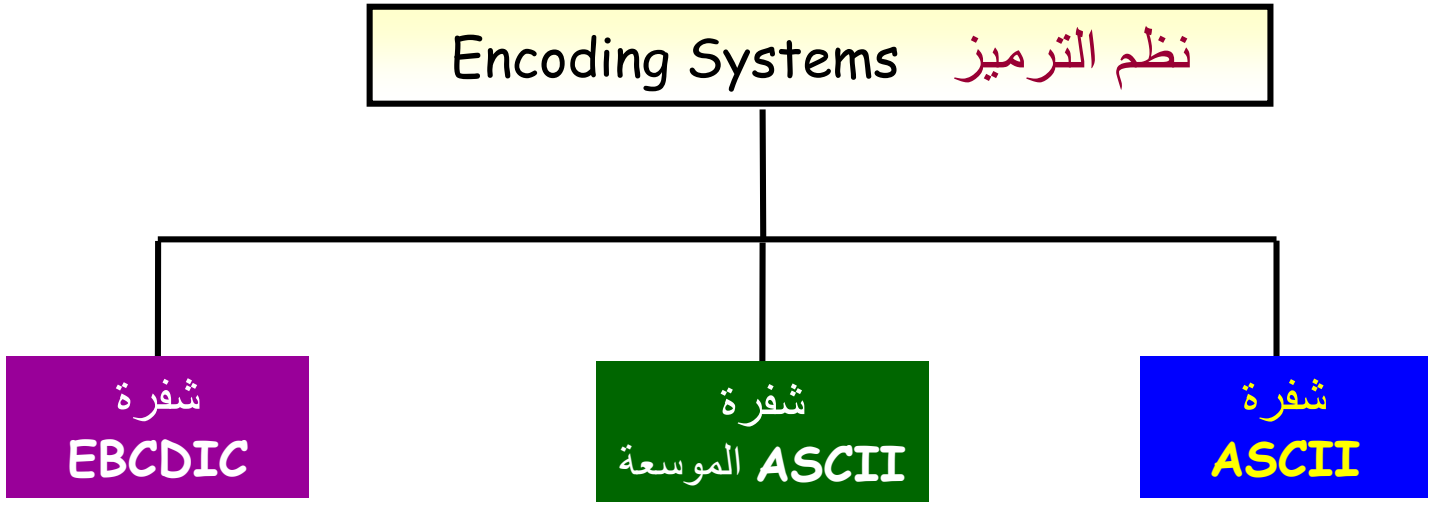
البت Bit	1,0
البايت Byte	8 بت
الكيلو بايت Kilo Byte	1024 بايت
الميجابايت Mega Byte	مليون بايت
الجيجابايت Giga Byte	بليون بايت

النص Text

- أي جزء من النص في أي لغة هو سلسلة من الرموز Sequence of symbols التي تعبر عن فكرة في هذه اللغة.
- الإنجليزية مثلا تحتوي علي 26 symbols كحروف كبيرة ومثلها حروف صغيرة ومجموعة أرقام.
- يمكن أن يمثل كل رمز منها بسلسلة بتات bit pattern.

نظم الترميز والتشفير Encoding Systems

- لكن كيف يميز الحاسب بين الرموز المختلفة وكلها سلاسل ثنائية.
- تم تصميم مجموعة نظم قياسية عالمية للترميز لتمثيل رموز البيانات.



الشفرة الأمريكية القياسية لتبادل البيانات

- تسمى American Standard Code for Information Interchange (ASCII)
- قام معهد المعايرة الوطني الأمريكي ANSI بوضع شفرة ASCII
- تستخدم ASCII لتمثيل الرمز 7-bits.
- هذا يعني أنها قادرة على ترميز $2^7=128$ رمزا مختلفا.
- أول نسق ثنائي bit pattern هو 0000 000 يمثل null character.
- أخر نسق ثنائي bit pattern هو 1001111 يمثل Delete character.

Dec	Hex	Name	Char	Ctrl-char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char
0	0	Null	NUL	CTRL-@	32	20	Space	64	40	@	96	60	`
1	1	Start of heading	SOH	CTRL-A	33	21	!	65	41	A	97	61	a
2	2	Start of text	STX	CTRL-B	34	22	"	66	42	B	98	62	b
3	3	End of text	ETX	CTRL-C	35	23	#	67	43	C	99	63	c
4	4	End of xmit	EOT	CTRL-D	36	24	$	68	44	D	100	64	d
5	5	Enquiry	ENQ	CTRL-E	37	25	%	69	45	E	101	65	e
6	6	Acknowledge	ACK	CTRL-F	38	26	&	70	46	F	102	66	f
7	7	Bell	BEL	CTRL-G	39	27	'	71	47	G	103	67	g
8	8	Backspace	BS	CTRL-H	40	28	(	72	48	H	104	68	h
9	9	Horizontal tab	HT	CTRL-I	41	29	)	73	49	I	105	69	i
10	0A	Line feed	LF	CTRL-J	42	2A	*	74	4A	J	106	6A	j
11	0B	Vertical tab	VT	CTRL-K	43	2B	+	75	4B	K	107	6B	k
12	0C	Form feed	FF	CTRL-L	44	2C	,	76	4C	L	108	6C	l
13	0D	Carriage feed	CR	CTRL-M	45	2D	-	77	4D	M	109	6D	m
14	0E	Shift out	SO	CTRL-N	46	2E	.	78	4E	N	110	6E	n
15	0F	Shift in	SI	CTRL-O	47	2F	/	79	4F	O	111	6F	o
16	10	Data line escape	DLE	CTRL-P	48	30	0	80	50	P	112	70	p
17	11	Device control 1	DC1	CTRL-Q	49	31	1	81	51	Q	113	71	q
18	12	Device control 2	DC2	CTRL-R	50	32	2	82	52	R	114	72	r
19	13	Device control 3	DC3	CTRL-S	51	33	3	83	53	S	115	73	s
20	14	Device control 4	DC4	CTRL-T	52	34	4	84	54	T	116	74	t
21	15	Neg acknowledge	NAK	CTRL-U	53	35	5	85	55	U	117	75	u
22	16	Synchronous idle	SYN	CTRL-V	54	36	6	86	56	V	118	76	v
23	17	End of xmit block	ETB	CTRL-W	55	37	7	87	57	W	119	77	w
24	18	Cancel	CAN	CTRL-X	56	38	8	88	58	X	120	78	x
25	19	End of medium	EM	CTRL-Y	57	39	9	89	59	Y	121	79	y
26	1A	Substitute	SUB	CTRL-Z	58	3A	:	90	5A	Z	122	7A	z
27	1B	Escape	ESC	CTRL-[	59	3B	;	91	5B	[	123	7B	{
28	1C	File separator	FS	CTRL-\	60	3C	<	92	5C	\	124	7C	|
29	1D	Group separator	GS	CTRL-]	61	3D	=	93	5D	]	125	7D	}
30	1E	Record separator	RS	CTRL-^	62	3E	>	94	5E	^	126	7E	~
31	1F	Unit separator	US	CTRL-_	63	3F	?	95	5F	_	127	7F	DEL

شفرة ASCII الموسعة

- لجعل كل رمز يمثل داخل Byte وسعت شفرة ASCII بإضافة 0 إلى اليسار.
- يصبح بذلك أول تشكيل هو 00 000 000 وآخر تشكيل هو 0111 111 11

النظام العشري الموسع في الشفرة الثنائية

- في بدايات عصر الحاسبات قامت شركة IBM بتصميم شفرة EBCDIC أي EXTENDED BINERY CODED DECIMAL INTERCHARGE CODE
- يستخدم هذا النظام 8 bits لتمثيل الحرف.
- يمكن أن يميز بين $2^8 = 256$ رمزا.
- لا يستعمل إلا في حاسبات IBM فقط .

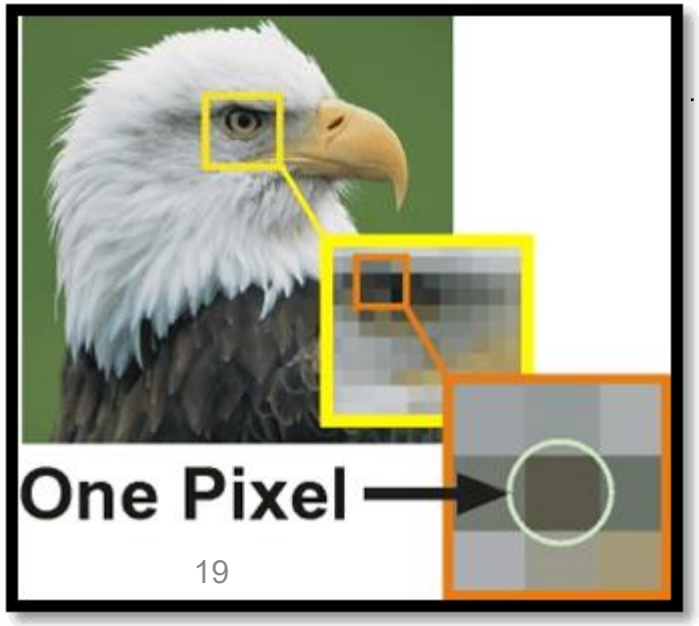
أنظمة التشفير والترميز

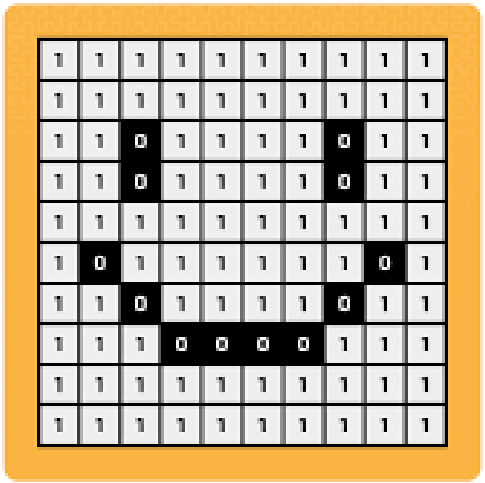
	EBCDIC		ASCII	
Character	Binary	Decimal	Binary	Decimal
A	1100 0001	193	100 0001	65
B	1100 0010	194	100 0010	66
a	1000 0001	129	110 0001	97
b	1000 0010	130	110 0010	98
0	1111 0000	240	011 0000	48
1	1111 0001	241	011 0001	49
<	0100 1100	76	010 1100	60

الصور

خريطة البتات Bitmap

- تقسم الصورة إلى مصفوفة من النقاط الصغيرة Dots تسمى نقاط مضيئة Pixels.
- تعتمد مساحة النقطة على درجة الوضوح والدقة Resolution.
- يخصص لكل نقطة تشكيل ثنائي Bit pattern.
- إذا كانت الصورة غير ملونة (أبيض/ أسود فقط) يكفي 1-bit لتمثيل النقطة فيكون White(1) و Black(0) .
- يتم تسجيل التشكيلات الثنائية للنقطة واحدة تلو الأخرى وتخزن بالحاسب.





خريطة البتات

- إذا كانت الصورة تحتوي على تدرج من اللون الرمادي أيضا فيمكن أن تمثل النقطة بـ 2-bitفيكون مثلا White(11) ويكون light gray(10) ويكون dark gray(01)
- لتمثيل الصور الملونة تستخدم ثلاثة ألوان R,G,B
- تقاس شدة كل لون بدلالة هذه الألوان الثلاثة
- تمثل كل نقطة بثلاثة تشكيلات ثنائية كل تشكيل يمثل لونا إما Green أو Red أو Blue

الصوت Audio

- الصوت بطبيعته هو إشارة تماثلية متصلة زمنيا.
- لابد من تحويله إلى إشارة رقمية حتى يمكن تخزينه في الحاسب.
- يتم ذلك عبر العمليات الآتية:
  - تؤخذ عينات من الصوت على فترات زمنية متقاربة ومتساوية.
  - تحدد قيمة العينات وتخصص لها قيمة Value من مجموعة قيم معدة سلفا.
  - تحول القيمة التي تم تحديدها إلى تشكيل ثنائي bit pattern.
  - يخزن هذا التشكيل الثنائي في الحاسب.

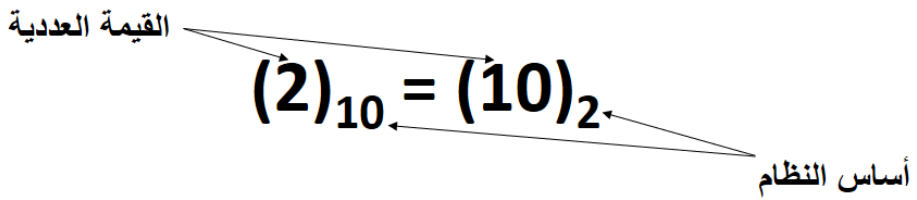
الفيديو Video

- الفيديو هو تمثيل للصور (تسمى الصورة frame إطار) في الزمن.
- الفيلم movie هو تسلسل من الإطارات الواحد تلو الآخر.
- كل صورة أو إطار تحول إلى تشكيل ثنائي وتخزن في الحاسب.
- مجموعة الصور المتتابعة تمثل الفيديو.

الأعداد Numbers

- تمثيل الأعداد يختلف عن تمثيل البيانات غير العددية Non-number data.
- يتم تحويل الأعداد (العشرية)الى نظام العد الثنائي أو النظام الثنائي و هو نظام عد ذو رقم أساس 2، يستخدم لتمثيل القيم العددية باستخدام رمزين، عادة ما يكونان 0 و1

نظام العد العشري	نظام العد الثنائي
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000



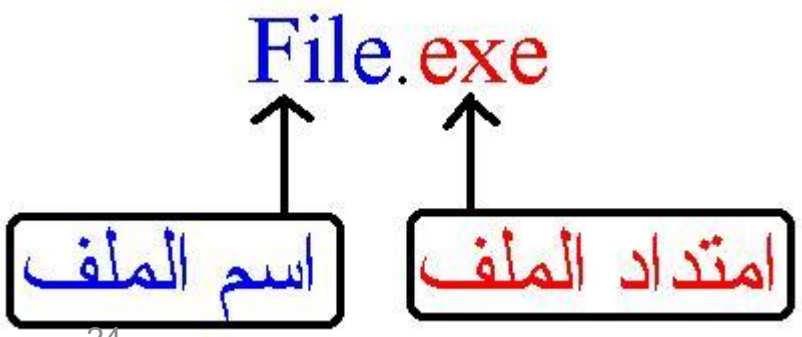
تخزين البيانات في الذاكرة الثانوية:

تكمن أهمية الذاكرة الثانوية في تخزينها لمجموعات من البيانات تحتاج لأن تحفظ بعيدا عن ذاكرة الحاسوب الرئيسية.

وهذه المجموعات تعرف بالملفات **Files** وتمتاز بحجمها الكبير وديمومتها , فهي توجد مستقلة عن البرامج التي تصنعها وتعدل عليها وتستخدمها.

إن استخدام الملفات ضروري في تخزين البيانات التي لا يتناسب حجمها مع حجم الذاكرة الرئيسية والتي يجب أن تبقى سليمة و كاملة حتى بعد انتهاء عمل الحاسوب.

كل ملف له اسم و امتداد , والاسم يميز كل ملف عن الآخر و للمستخدم الحرية في اختيار الاسم الذي يرغبه لملفه أما الامتداد فهو خاص بجهاز الحاسوب حسب نوع الملف , فالملفات التي تحمل الامتداد **doc** هي وثائق , **wave** ملفات صوت , **mov** ملفات فيديو , **gif** صور وهكذا.



بغض النظر عن نوع الملفات فإنها جميعا تخضع لنفس العمليات وهي:

1. صناعة الملف وتسميته وحفظه Create, Name, Save
2. نسخ الملف و تحريكه وحذفه Copy, Move, Delete
3. عرض الملف على الشاشة و طباعته Display, Print
4. تنفيذ الملف Execute
5. تحميل الملف من القرص للذاكرة الرئيسية لإمكانية نسخه من قبل الآخرين على الشبكة Upload و العكس يسمى Download

7. تصدير الملف من البرنامج الذي تعمل عليه إلى برنامج آخر **Export** و العكس يسمى **Import**
8. ضغط الملف بحيث يخزن دون فراغات و بالتالي تصغير حجمه **Compress**
9. حماية الملف من عبث الآخرين أو الوصول غير المخول أو الفيروسات من خلال كلمات السر أو تغيير خصائصه كإخفائه أو جعله للقراءة فقط ومن خلال برامج الوقاية من الفيروسات **Protect**

الى اللقاء في المحاضرة القادمة