

Codage

donnée : suite de 0 et de 1 dans l'ordinateur

Codage binaire (en base 2)

1

1

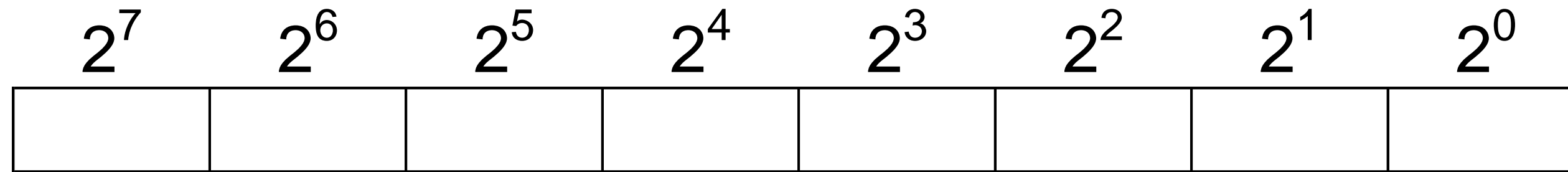
→ le courant passe

0

0

→ le courant ne passe pas

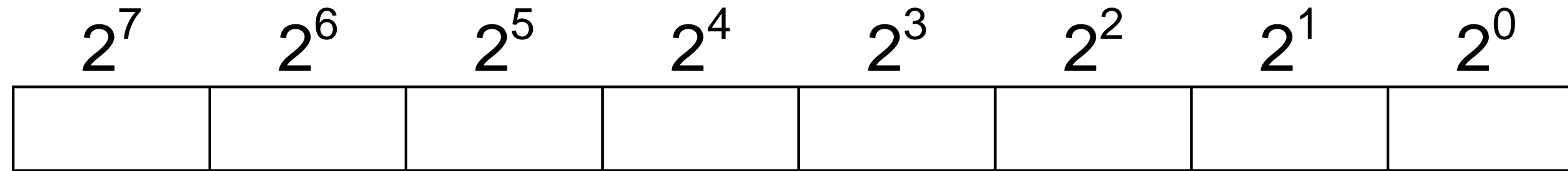
Codage



Codage historique sur 6 bits puis **8 bits** (byte)

Ensuite sur 16 bits, 32 bits , **64 bits** et 128 bits

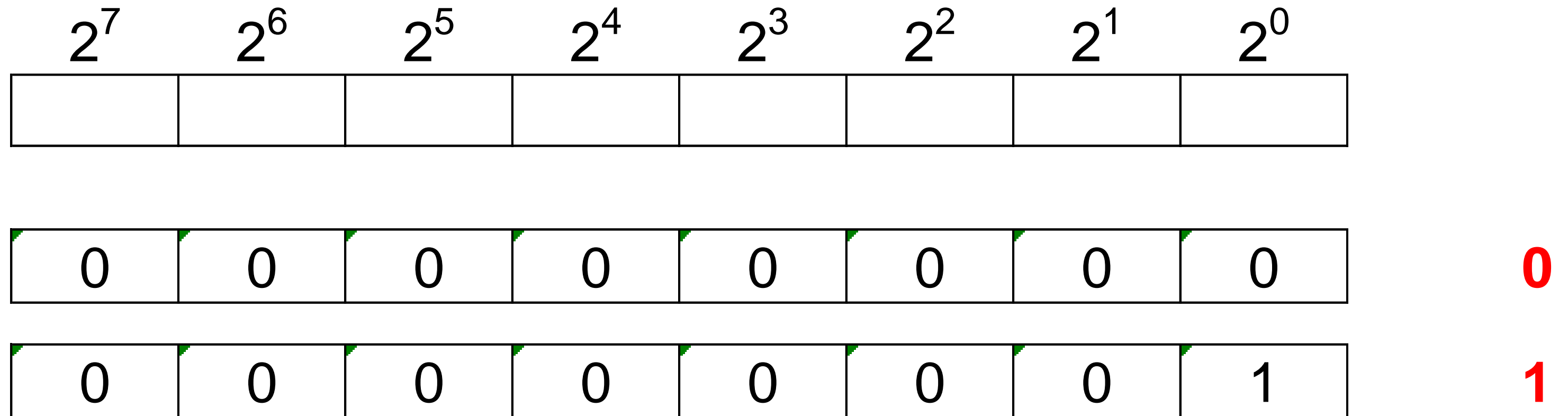
Codage



Principe du codage **ASCII** sur **8 bits** (byte)

ASCII : American Standard Code for Information Interexchange

Codage



1

1

→ le courant passe

0

0

→ le courant ne passe pas

ASCII : American Standard Code for Information Interexchange

Codage

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	1	0	2

ASCII : American Standard Code for Information Interexchange

Codage

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	1	0	2
0	0	0	0	0	0	1	1	3

ASCII : American Standard Code for Information Interexchange

Codage

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	1	0	2
0	0	0	0	0	0	1	1	3

ASCII : American Standard Code for Information Interexchange

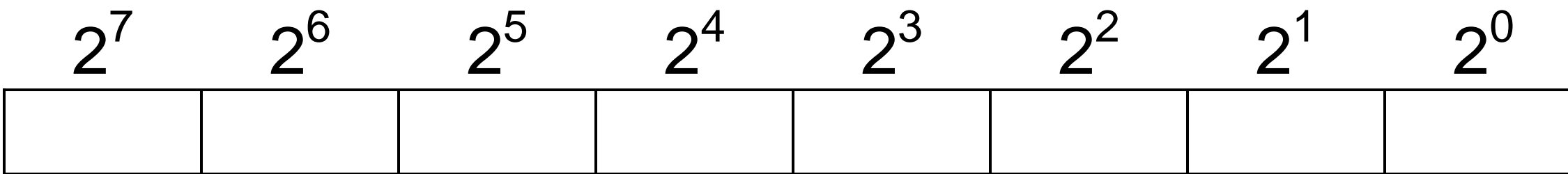
Codage

Code numérique = mot machine = byte

C'est la plus petite structure qui entre dans l'architecture des ordinateurs.

Chaque « mot » contient 8 cases (bits) dans un codage binaire composé de 1 ou 0 dans chaque case (1 : le courant électrique passe ou 0 : il ne passe pas).

Le codage de caractères sur 8 bits autorise 256 caractères

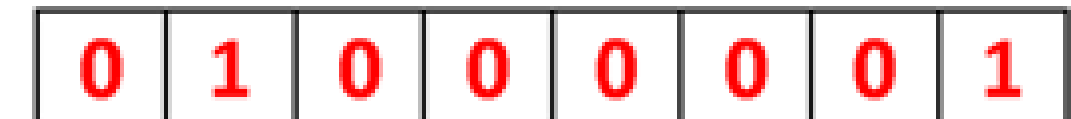


Exemple

Le caractère **A** correspond au code **65** décomposé en **binaire** de la façon suivante :

$$65 = 0*2^7 + 1*2^6 + 0*2^5 + 0*2^4 + 0*2^3 + 0*2^2 + 0*2^1 + 1*2^0$$

⇒



Code ASCII

0	32	64	Q	96	`	128	Ç	160	á	192	L	224	ó		
1	☐	33	!	65	À	97	a	129	ü	161	í	193	Ĺ	225	ß
2	☒	34	"	66	B	98	b	130	é	162	ó	194	T	226	ô
3	♥	35	#	67	C	99	c	131	â	163	ú	195	†	227	ò
4	♦	36	\$	68	D	100	d	132	ä	164	ñ	196	—	228	õ
5	♣	37	%	69	E	101	e	133	à	165	Ñ	197	‡	229	ö
6	♠	38	&	70	F	102	f	134	ã	166	ª	198	ã	230	µ
7	•	39	'	71	G	103	g	135	ç	167	º	199	Ã	231	¶
8	■	40	(72	H	104	h	136	ê	168	¿	200	℄	232	Ⓟ
9	◊	41)	73	I	105	i	137	ë	169	®	201	ℍ	233	ú
10	◼	42	*	74	J	106	j	138	è	170	¬	202	≡	234	û
11	♠	43	+	75	K	107	k	139	ï	171	½	203	≡	235	ü
12	♀	44	,	76	L	108	l	140	î	172	¼	204	≡	236	ý
13	♯	45	_	77	M	109	m	141	ì	173	¡	205	=	237	ÿ
14	♠	46	.	78	N	110	n	142	ÿ	174	«	206	≡	238	·
15	✱	47	/	79	O	111	o	143	ÿ	175	»	207	☒	239	´
16	▶	48	0	80	P	112	p	144	É	176	⋮	208	δ	240	-
17	◀	49	1	81	Q	113	q	145	æ	177	⋮	209	Đ	241	±
18	↕	50	2	82	R	114	r	146	ff	178	⋮	210	Ê	242	=
19	!!	51	3	83	S	115	s	147	ô	179		211	Ë	243	¾
20	¶	52	4	84	T	116	t	148	ö	180	†	212	È	244	¶
21	⊗	53	5	85	U	117	u	149	ò	181	Á	213	É	245	⊗
22	—	54	6	86	V	118	v	150	û	182	Â	214	Í	246	÷
23	±	55	7	87	W	119	w	151	ù	183	À	215	Î	247	·
24	↑	56	8	88	X	120	x	152	ÿ	184	Ⓞ	216	Ï	248	•
25	↓	57	9	89	Y	121	y	153	ö	185	≡	217	J	249	..
26	→	58	:	90	Z	122	z	154	ü	186	≡	218	Γ	250	·
27	←	59	;	91	[123	{	155	ø	187	≡	219	■	251	¡
28	└	60	<	92	\	124		156	£	188	≡	220	■	252	‡
29	✱	61	=	93]	125	}	157	Ø	189	¢	221	¡	253	z
30	▲	62	>	94	^	126	~	158	×	190	¥	222	ì	254	■
31	▼	63	?	95	_	127	△	159	f	191	ł	223	■	255	

Format Texte ASCII (étendu)

(American Standard Code for Information Interexchange)

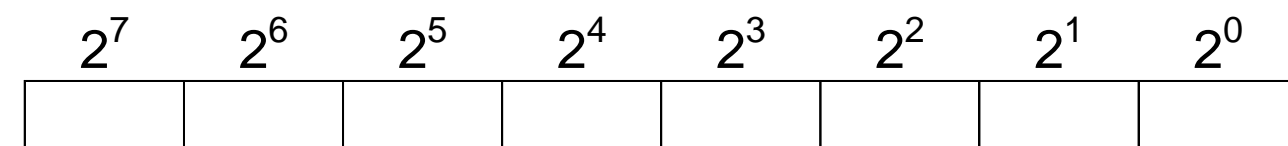
Type de codage standard, adopté dans les années 60, qui fait correspondre un caractère (lettre, chiffre, ponctuation,...) à un code numérique.

Code numérique = mot machine = byte

C'est la plus petite structure qui entre dans l'architecture des ordinateurs.

Chaque « mot » contient 8 cases (bits) dans un codage binaire composé de 1 ou 0 dans chaque case (1 : le courant électrique passe ou 0 : il ne passe pas).

Le codage de caractères sur 8 bits autorise 256 caractères



Exemple

Le caractère **A** correspond au code **65** décomposé en **binaire** de la façon suivante :

$$65 = 0*2^7 + 1*2^6 + 0*2^5 + 0*2^4 + 0*2^3 + 0*2^2 + 0*2^1 + 1*2^0 \Rightarrow \boxed{0\ 1\ 0\ 0\ 0\ 0\ 0\ 1}$$

On a déjà besoin de :

'a' -> 'z' (26 caractères)

'A' -> 'Z' (26 caractères)

0 -> 9 (10 caractères)

Signes de ponctuation (au moins 10 caractères)

Opérateurs arithmétiques (+, -, /, *, ^, ...) (environ 20 caractères)

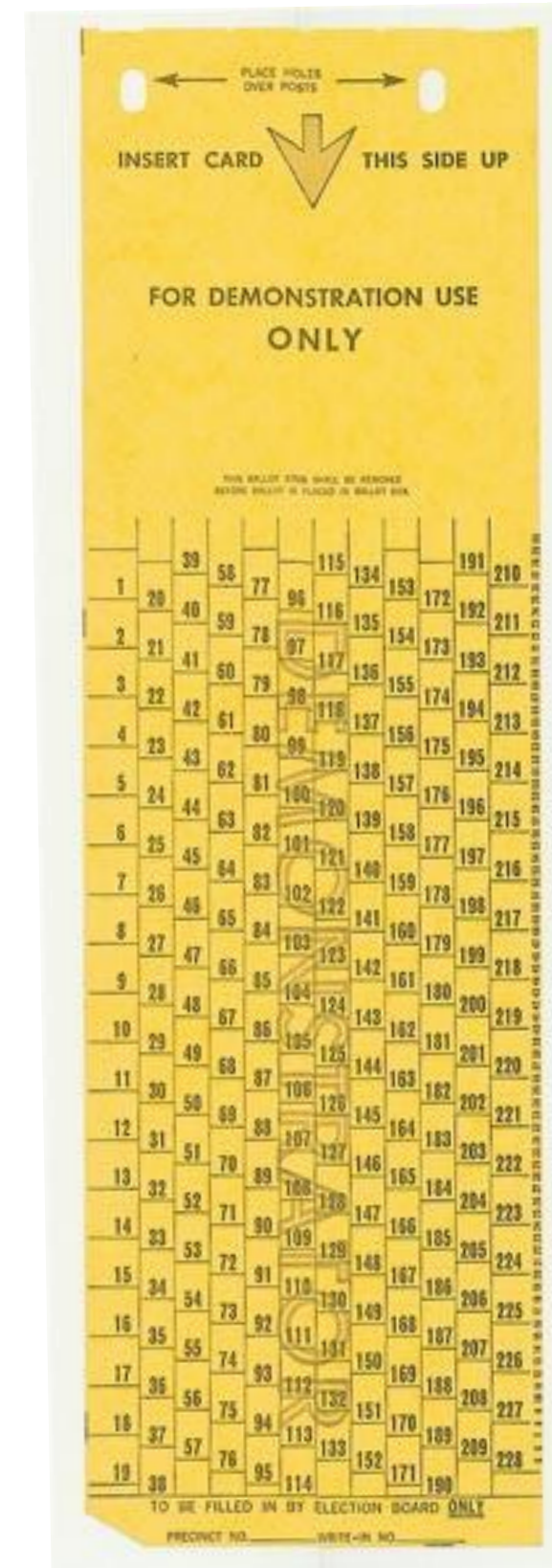
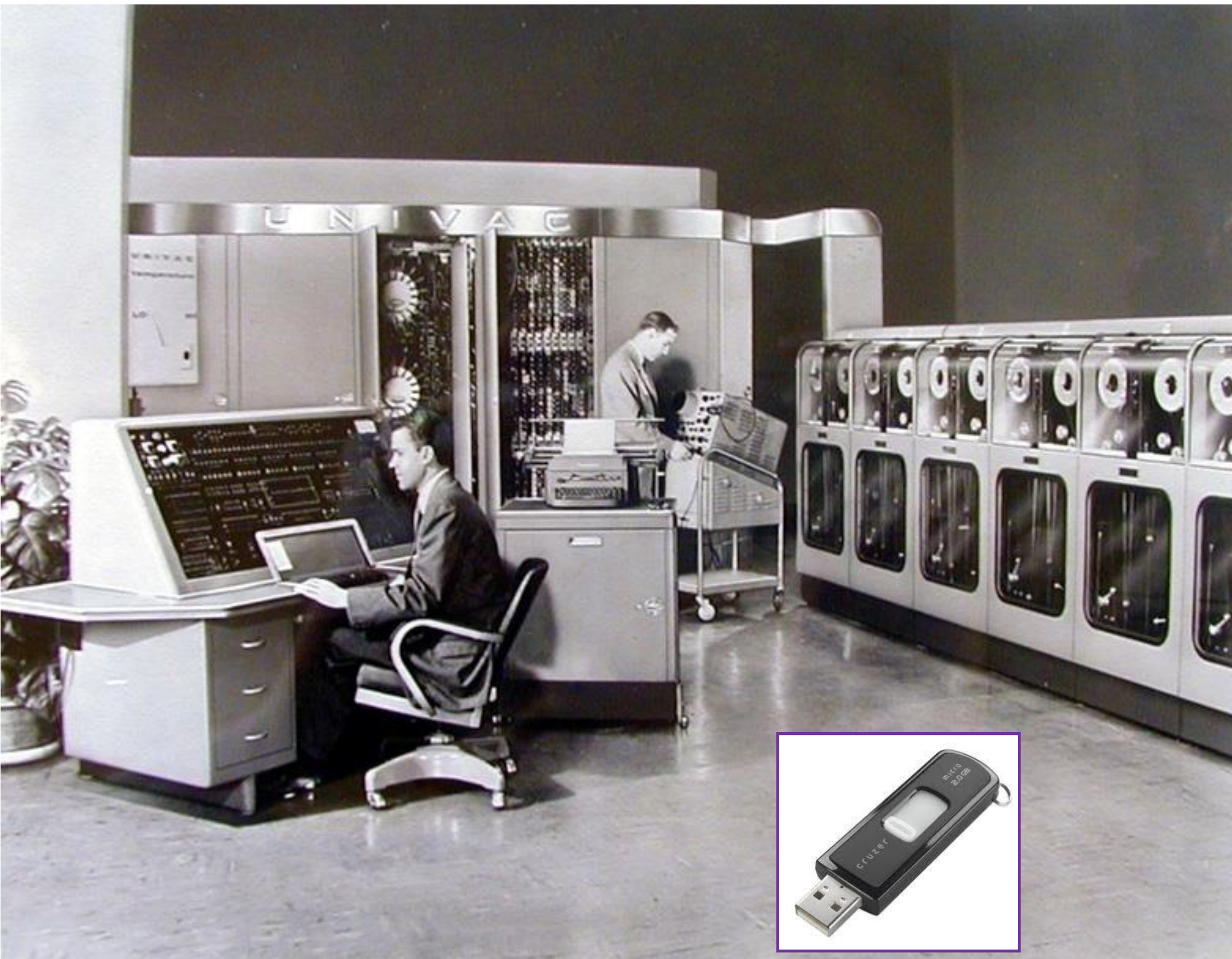
&, %, \$, ... (25 caractères)

CTRL C, CTRL X, ...

Code ASCII

L'Univac (1951)

Codage historique sur 6 bits puis **8 bits** (byte) ⇒ « mot machine »



Entrée/Sorties

Instructions

Fichiers

Format

Algorithme

Programme

Données

Variables

Sorties : cartes perforées

Organisation des données



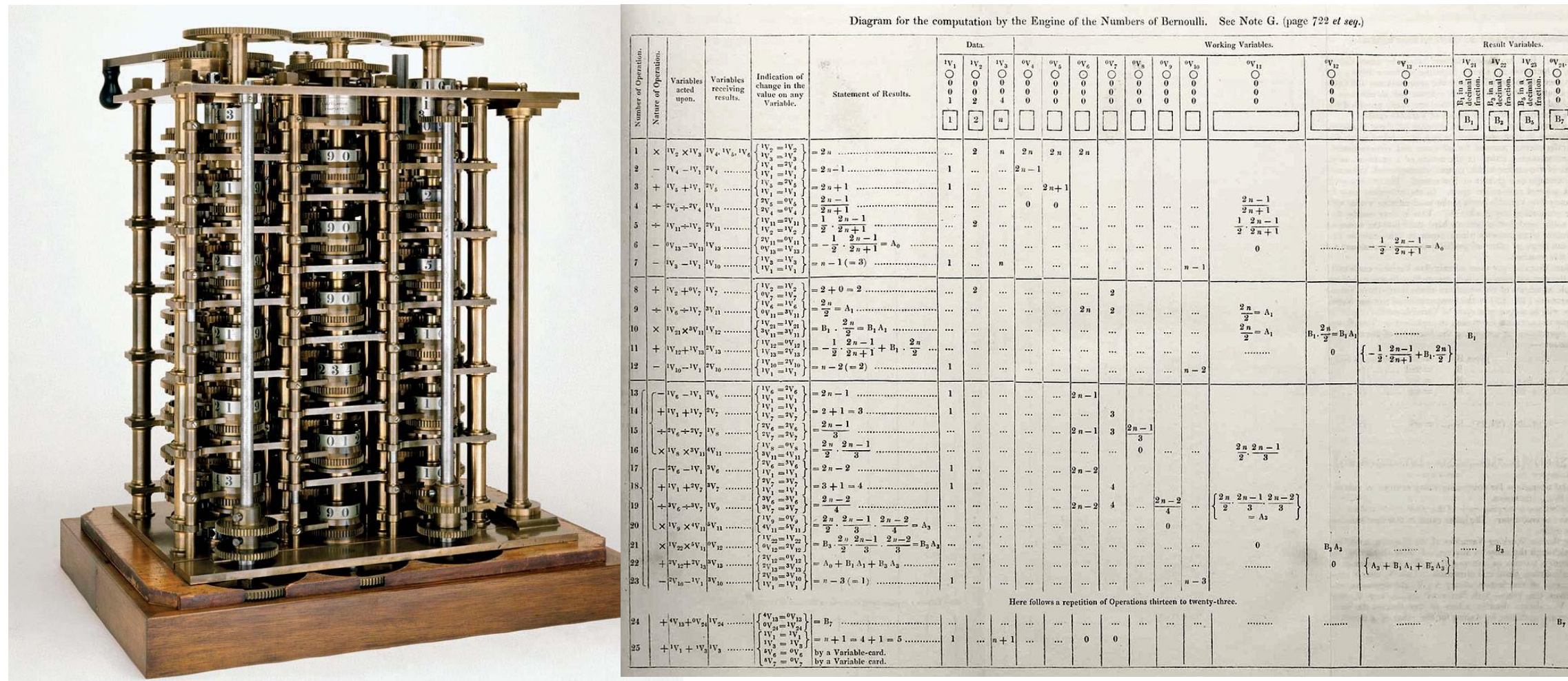
Ada Lovelace (1815-1852)

Pionnière de l'informatique.

En 1842 elle a réalisé le **premier programme** informatique

pour l'ancêtre de l'ordinateur :

la machine analytique de Charles Babbage.



Entrée/Sorties

Instructions

Fichiers

Format

Algorithme

Programme

Données

Variables

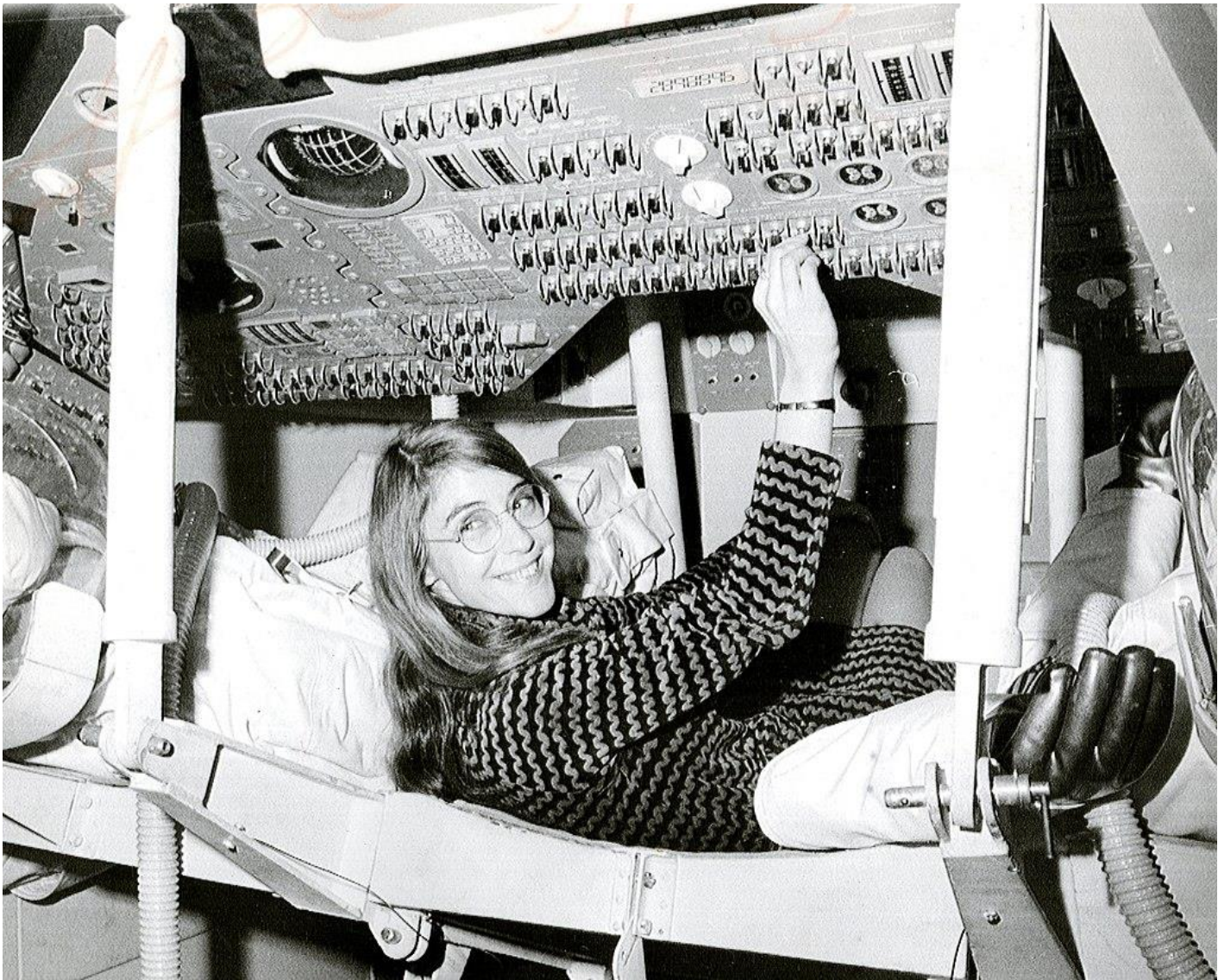
Entrées

Programme

Sorties

De nombreuses autres femmes ont contribué au développement de l'informatique

Organisation des données



Margareth Hamilton (née en 1936)

Programme spatial Apollo



Mary Jackson (1921-2005)

Informaticienne et ingénieure aérospatiale (NASA).

Le film « Les figures de l'ombre » retrace les destins de Katherine Johnson, Mary Jackson et Dorothy Vaughan qui ont permis de grandes avancées pour la NASA

Organisation des données

codage en groupes 8 bits

Données - Data

Quantitatives / Qualitatives



▷ unité d'information

Bit (0/1)

(codage binaire)

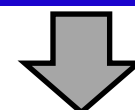
Organisation des données

Enregistrements (*individus*)



Données - Data

Quantitatives / Qualitatives



Bit (0/1)

(codage binaire)

codage en groupes 8 bits

▷ unité d'information

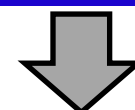
Organisation des données

Enregistrements (*individus*)



Champs (*données, data, attributs*)

Quantitatives / Qualitatives



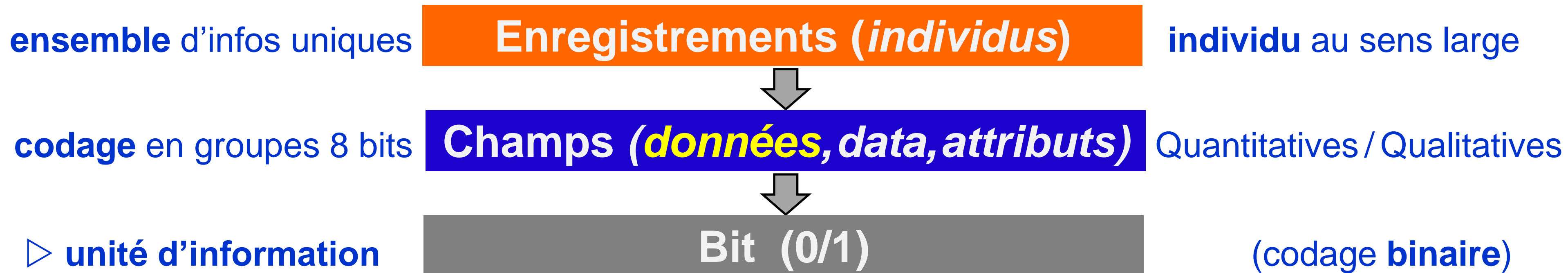
Bit (0/1)

(codage **binaire**)

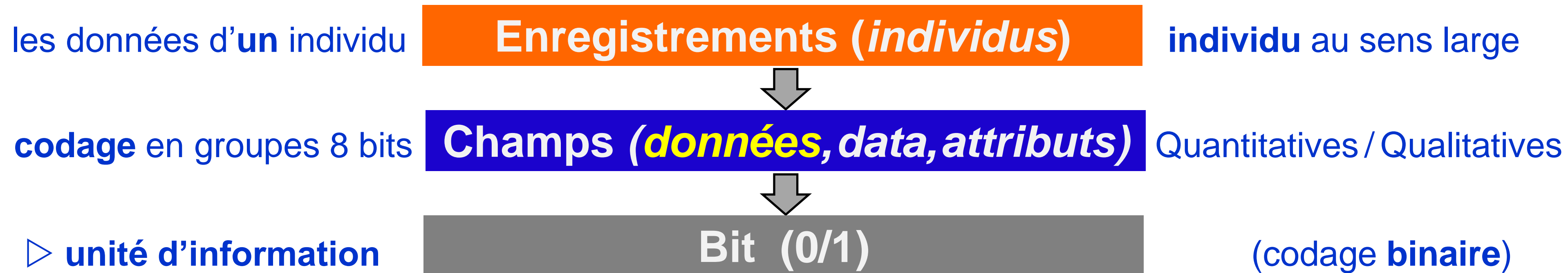
codage en groupes 8 bits

▷ **unité d'information**

Organisation des données



Organisation des données



000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

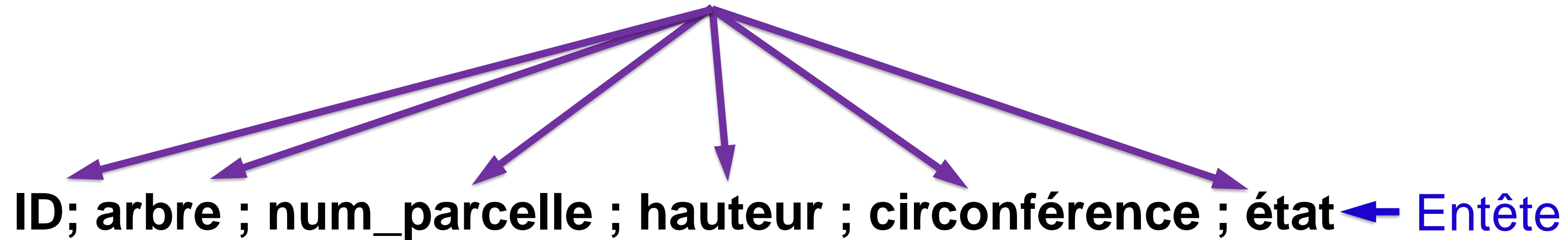
000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

Identificateurs des champs

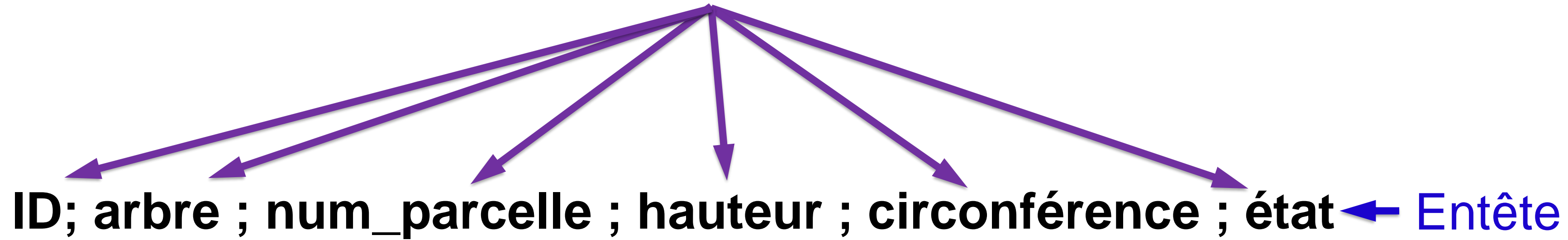


000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

Identificateurs des champs



000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

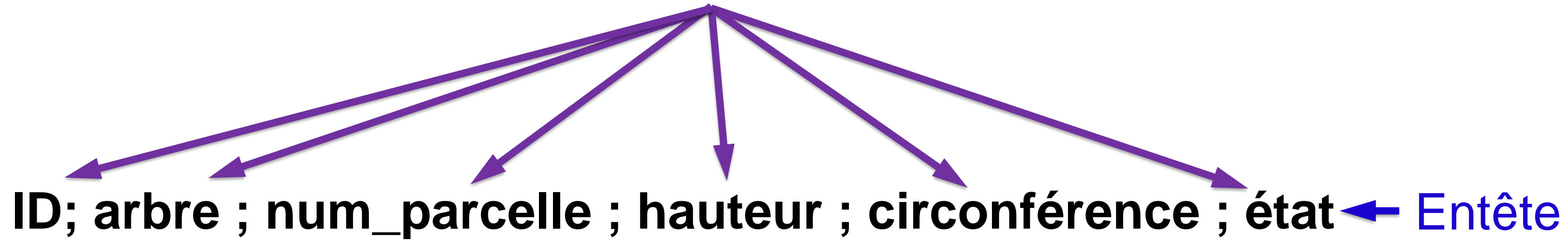
000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

Champs (*données, data*)

Five green arrows point upwards from the text 'Champs (données, data)' to the fields in the third record: ID, arbre, num_parcelle, hauteur, and circonférence.

Identificateurs des champs



Séparateur de champs

000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

Orange arrows point from a central point above the semicolons to each semicolon in the first line of data.

000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

Champs (données, data)



Identificateurs des champs

ID; arbre ; num_parcelle ; hauteur ; circonférence ; état ← Entête

Séparateur de champs

000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

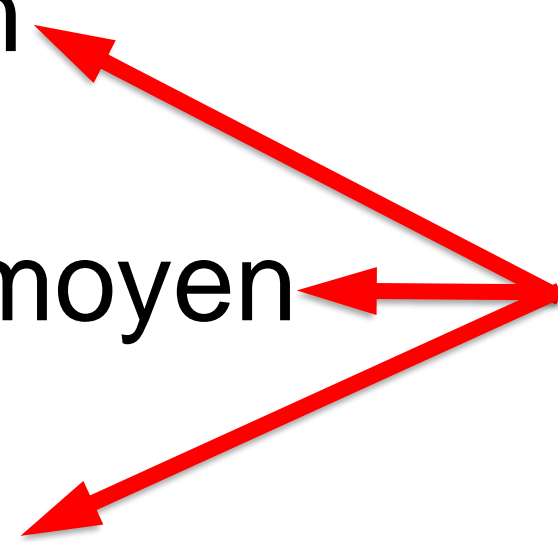
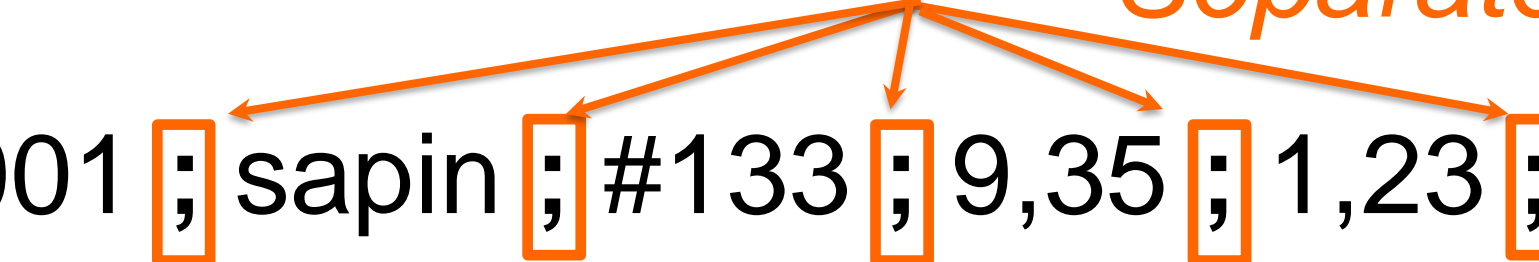
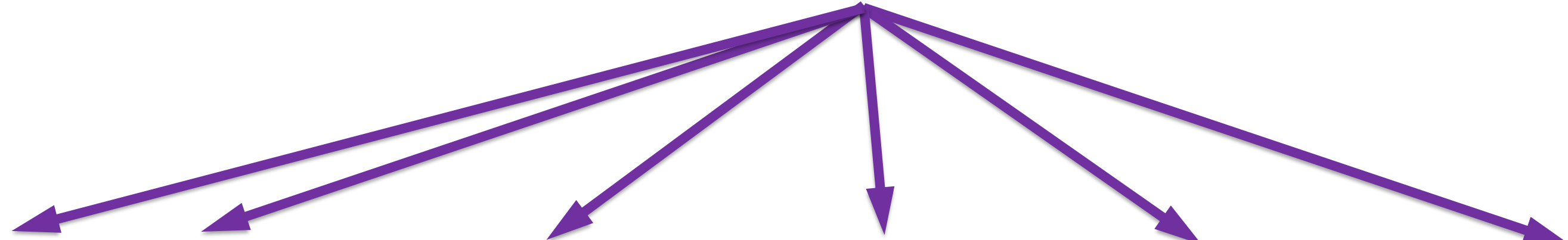
000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

**Enregistrements
(individus)**

Champs (données, data)

Individus



Identificateurs des champs

ID; arbre ; num_parcelle ; hauteur ; circonférence ; état ← Entête

Séparateur de champs

000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

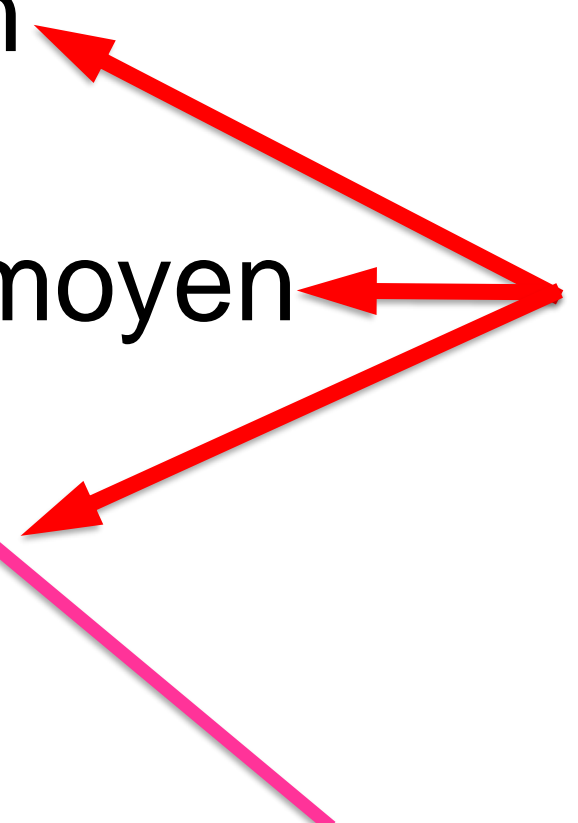
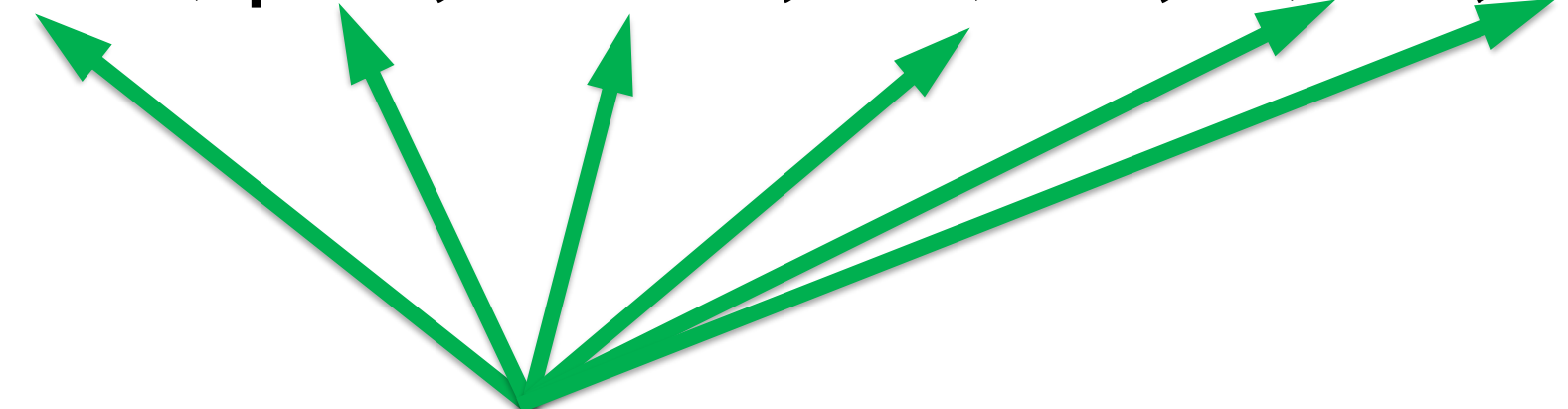
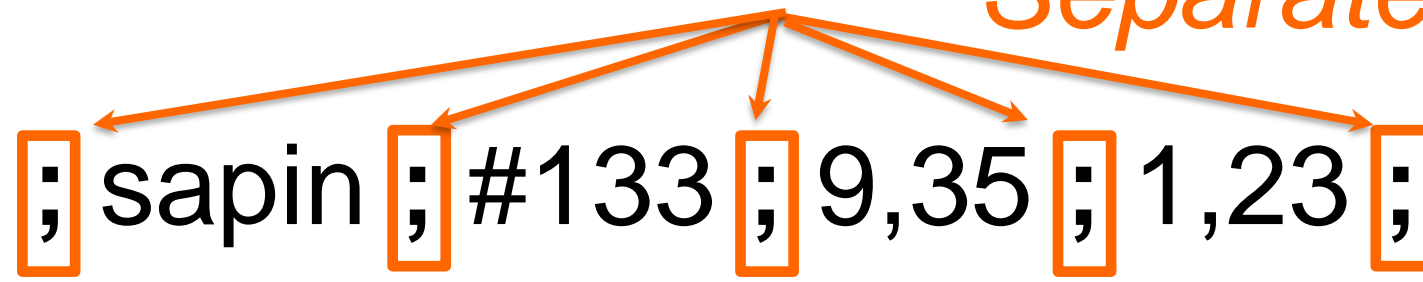
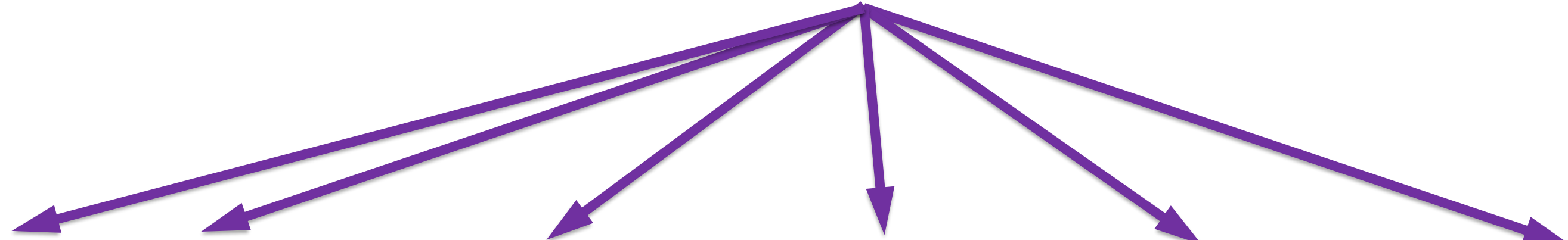
000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

Enregistrements
(individus)

Séparateur de décimales

Champs (données, data)

Individus



Identificateurs des champs \Rightarrow variables

ID; arbre ; num_parcelle ; hauteur ; circonférence ; état ← Entête

Séparateur de champs

000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

Individus

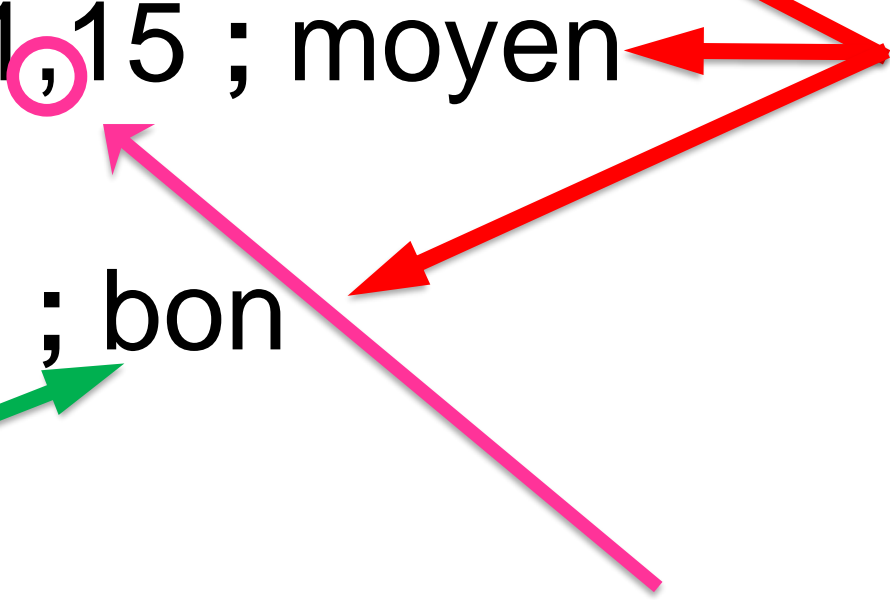
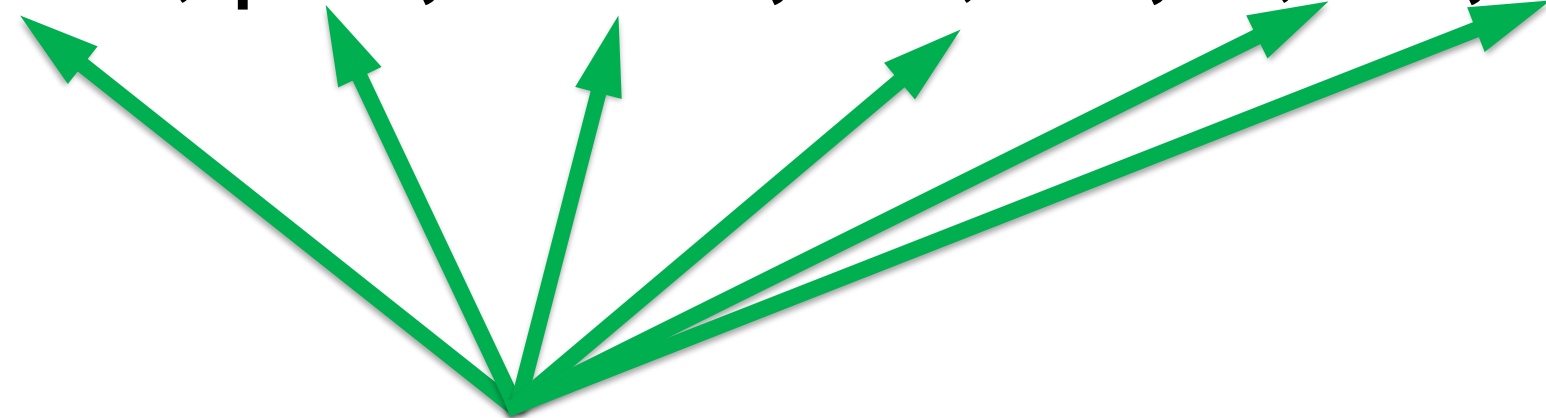
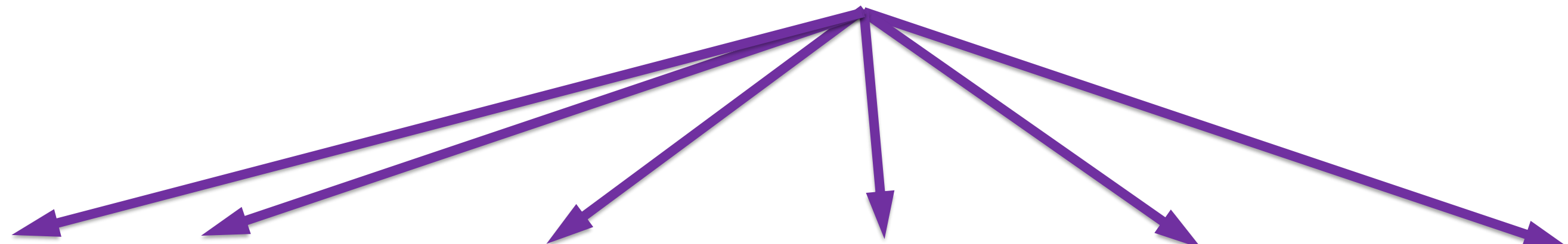
000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen

Enregistrements
(individus)

000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

Séparateur de décimales

Champs (données, data)



Enregistrement

Enregistrement : compilation d'informations relatives à l'individu considéré

- un enregistrement par individu ;
- découpé en **champs** de différente nature \Rightarrow **Variables**

*Pas de limite quant au nombre de **variables** engagées, ni à la diversité de leurs types*

Fichier

Lignes : individus

Colonnes : variables

Enregistrement



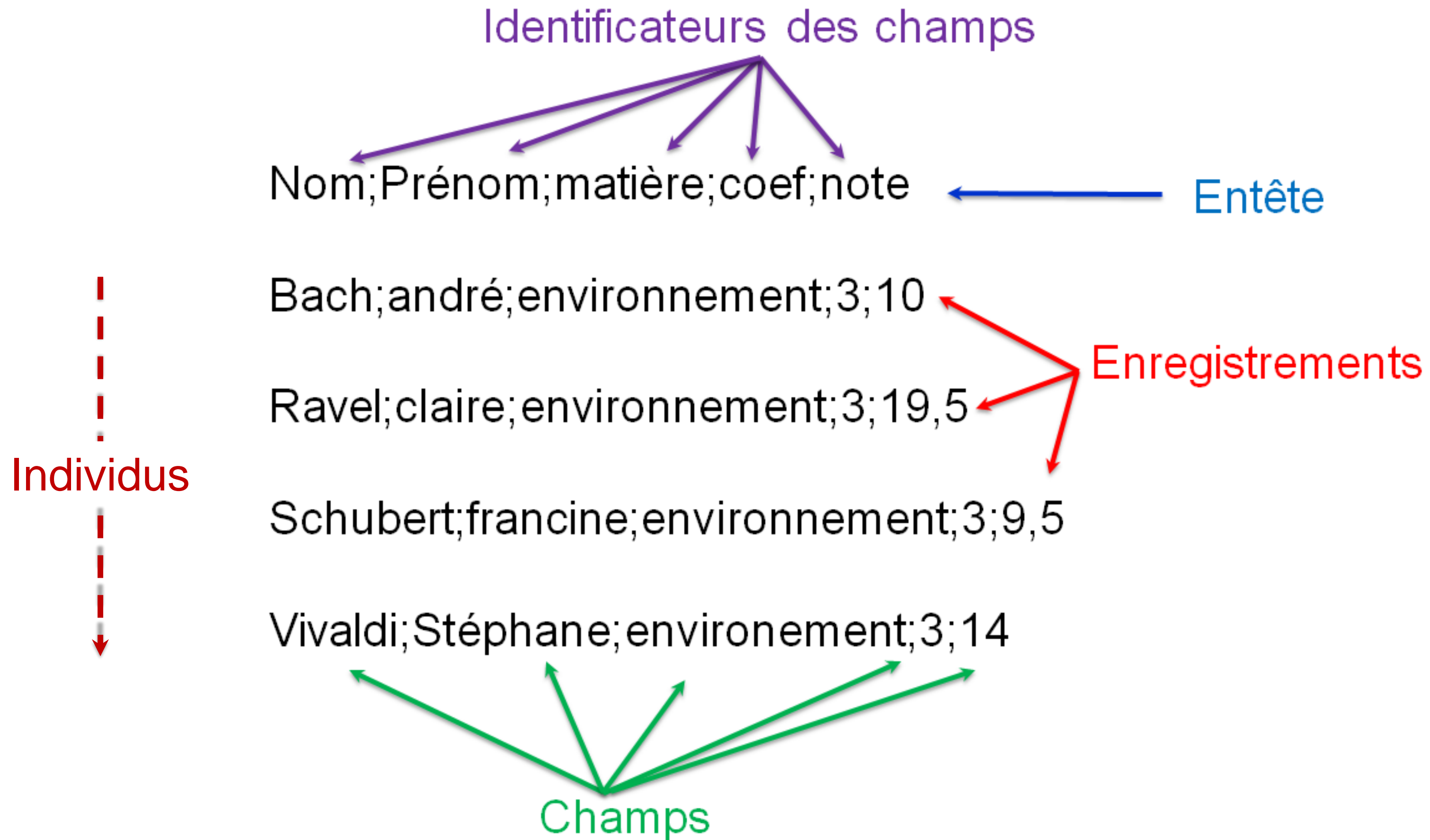
*Pas de limite quant au nombre de **variables** engagées, ni à la diversité de leurs types*

Bases de données

Extraction d'une partie de l'information (par des *requêtes*) pour établir (à l'aide d'une *analyse statistique* ou par *Machine Learning* en utilisant des *logiciels extérieurs*) de nouvelles *connaissances* et d'éventuelles *liaisons entre variables*

Relations de cause à effet

Exemple d'un fichier de notes



Enregistrement

Enregistrement : compilation d'informations relatives à l'individu considéré

- un enregistrement par individu ;
- découpé en **champs** de différente nature ;
- un champ donné contient **une** donnée, **une** valeur pour la variable décrivant ce champ ;
- un **champ** correspond à **une variable** d'un type donné ;

Un enregistrement regroupe donc des informations de différents types, propres à un individu.

Pas de limite quant au nombre de variables engagées, ni à la diversité de leurs types (quanti/quali)

Science

Observations

Technique

Mesures

Description

Caractéristiques

Contexte

Paramètres

Enquête

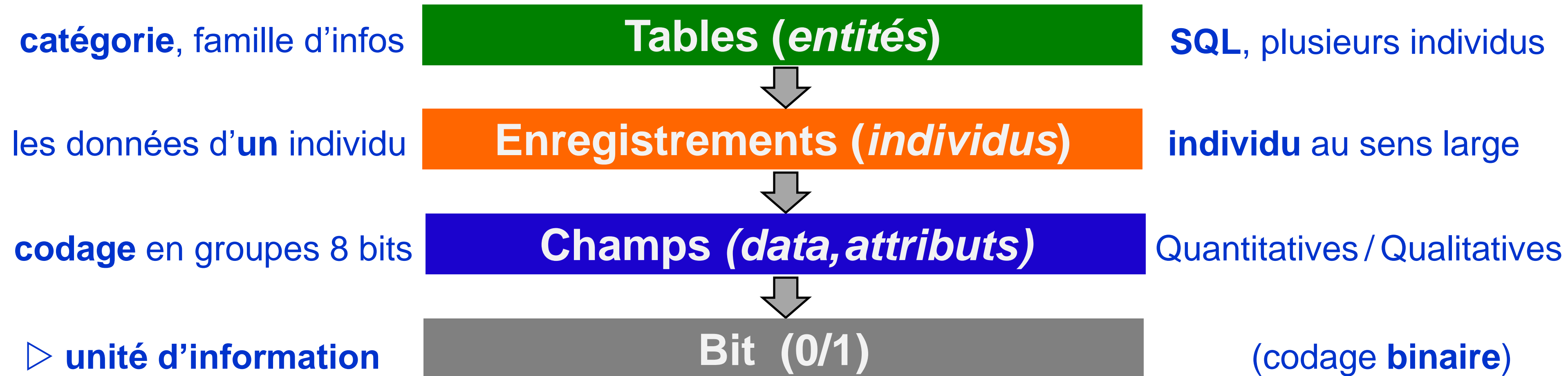
Archives

individu : pris au sens large

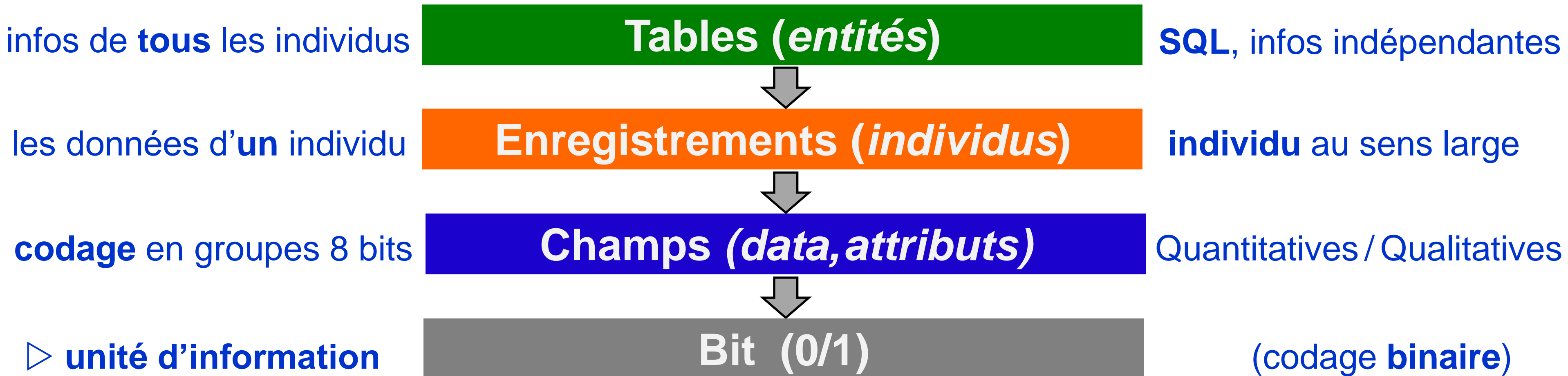
= unité statistique

Attention : bien contrôler la source des informations, leur crédibilité et leur précision !

Organisation des données



Organisation des données



ID; arbre ; num_parcelle ; hauteur ; circonférence ; état

000001 ; sapin ; #133 ; 9,35 ; 1,23 ; bon

000002 ; bouleau ; #095 ; 8,21 ; 1,15 ; moyen


000003 ; pin ; #022 ; 12,56 ; 1,50 ; bon

000004 ; bouleau ; #095 ; 9,51 ; 1,03 ; bon

000005 ; peuplier ; #022 ; 7,22 ; 0,95 ; mauvais

000006 ; chêne ; #133 ; 10,08 ; 1,25 ; bon

Nom du champ	Type de données	Description (facultative)
ID	Numérique	identificateur de l'individu
arbre	Texte court	type / espèce
num_parcelle	Texte court	code de la parcelle sur laquelle se trouve l'arbre
hauteur	Numérique	hauteur (en mètres)
circonférence	Numérique	circonférence (en mètres)
état	Texte court	santé de l'arbre

Arbres	
 ID	
arbre	
num_parcelle	
hauteur	
circonférence	
état	

Propriétés du champ

Général Liste de choix

Taille du champ	255
Format	
Masque de saisie	
Légende	
Valeur par défaut	
Valide si	
Message si erreur	
Null interdit	Non
Chaîne vide autorisée	Oui
Indexé	Non
Compression unicode	Non
Mode IME	Aucun contrôle
Mode de formulation IME	Aucun
Aligner le texte	Général

ID	arbre	num_parcel	hauteur	circonférence	état
000001	sapin	#133	9,35	1,23	bon
000002	bouleau	#095	8,21	1,15	moyen
000003	pin	#022	12,56	1,5	bon
000004	bouleau	#095	9,51	1,03	bon
000005	peuplier	#022	7,22	0,95	mauvais
000006	chêne	#133	10,08	1,25	bon

Mode Création. F6 = Autre volet. F1 = Aide.

Organisation des données

Bases de données



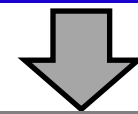
Tables (*entités*)



Enregistrements (*individus*)



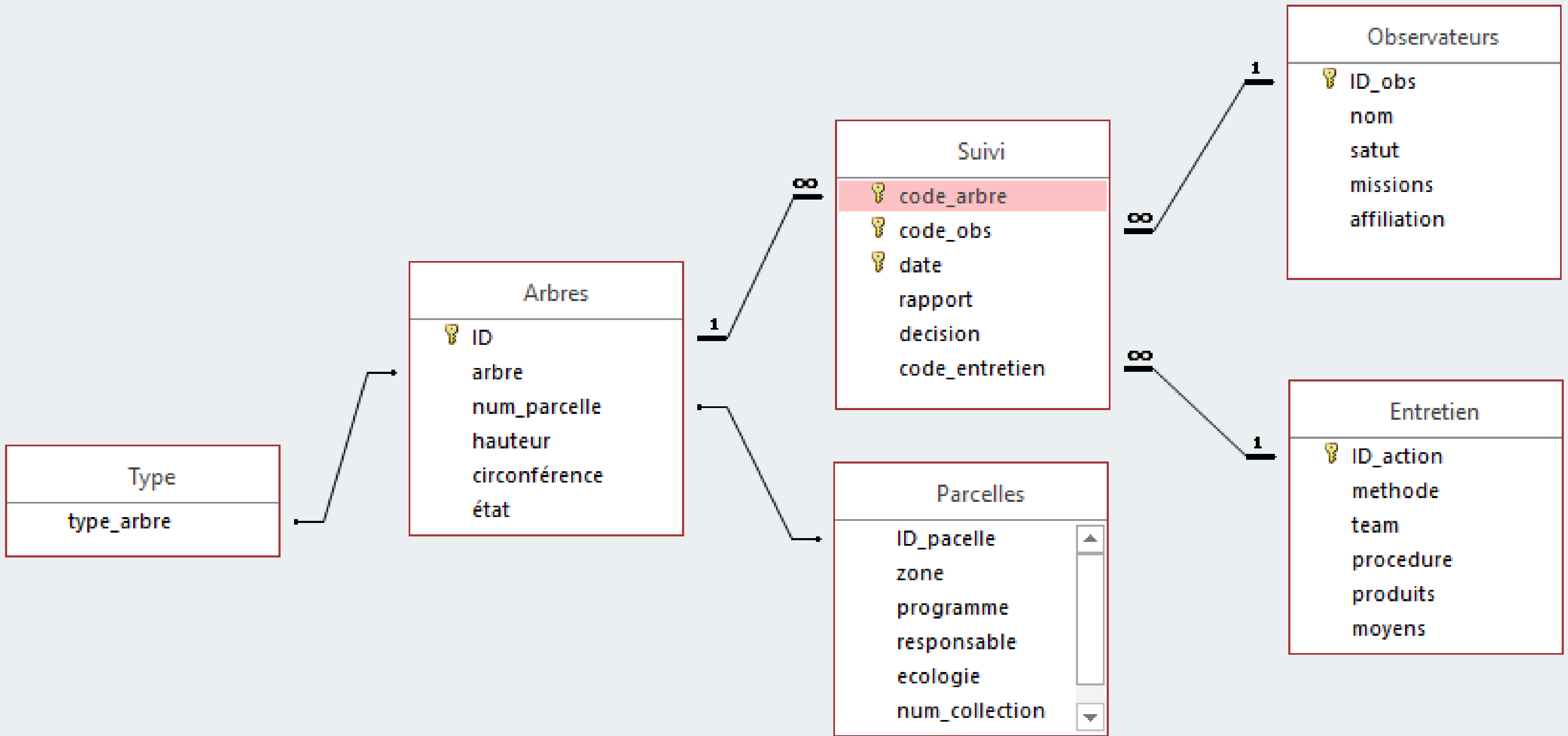
Champs (*attributs, data*)



Bit (0/1)

(produits numériques)

(codage binaire)



Organisation des données

Banques de données (*contenu*)

Datawarehouse (*entrepôts*)

Bases de données (*structure*)

Tables (*entités*)

Enregistrements (*individus*)

Champs (*attributs, data*)

Bit (0/1)

(produits numériques)

(codage binaire)

Organisation des données

Serveurs (*et réseaux*)

Banques de données (*contenu*)

Datawarehouse (*entrepôts*)

Bases de données (*structure*)

Tables (*entités*)

Enregistrements (*individus*)

Champs (*attributs, data*)

Bit (0/1)

(produits numériques)

(codage binaire)

Organisation des données

Big Data (*les 5 V*)

volume, vitesse, variété,
véracité, valeur.

Serveurs (*et réseaux*)

Banques de données (*contenu*)

Datawarehouse (*entrepôts*)

Bases de données (*structure*)

Tables (*entités*)

Enregistrements (*individus*)

Champs (*attributs, data*)

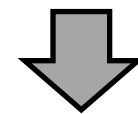
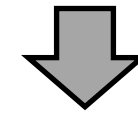
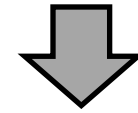
Bit (0/1)

(produits numériques)

(codage binaire)

Organisation des données

Big Data



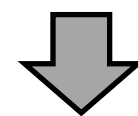
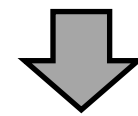
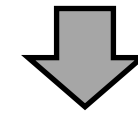
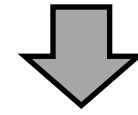
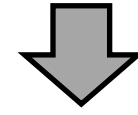
Bit (0/1)

(produits numériques)

(codage binaire)

Organisation des données

Big Data



Bit (0/1)

*Leibnitz, père de la notation binaire,
ne s'en serait sûrement pas étonné,
lui qui cherchait à tout expliquer
par les mathématiques....*

(produits numériques)

(codage binaire)