

# Informatika v medijih

## 2.3.

### Podatek in informacija. Predstavitev podatkov v računalniku.

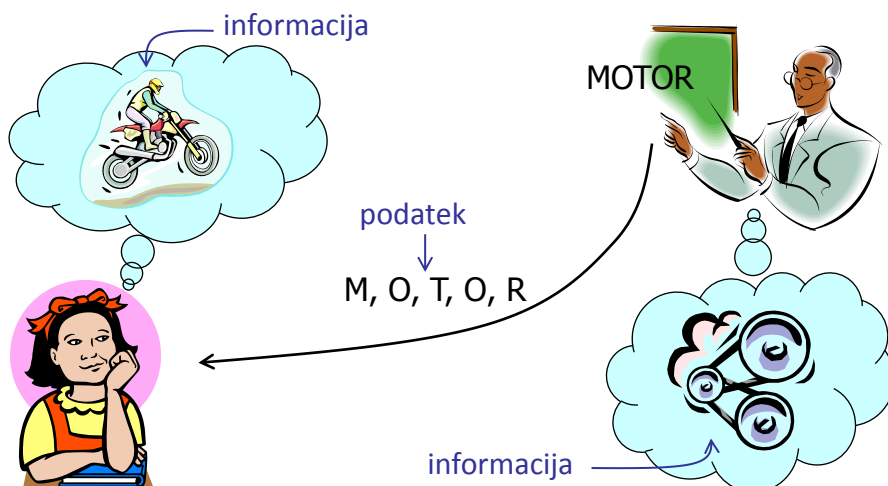
## Podatek ...

- ... je poljubna predstavitev s pomočjo simbolov ali analognih veličin, ki ji je pripisan, ali se ji lahko pripiše nek pomen. (*ANSI in ISO*)
- ... je predstavitev dejstva, koncepta ali instrukcije na formaliziran način, ki je primeren za komunikacijo, interpretacijo ali obdelavo s strani človeka ali stroja. (*ANSI in ISO*)
- ... so dejstva, predstavljena z vrednostmi (številke, znaki, simboli), ki imajo pomen v določenem kontekstu. (*Everest*)

## Informacija ...

- ... je pomen, ki ga človek pripiše podatkom s pomočjo znanih konvencij, ki so uporabljene pri njihovi predstavitvi. (*ANSI in ISO*)
- ... so podatki, postavljeni v kontekst. (*Alter*)
- ... je znanje, pridobljeno iz podatkov. (*Alter*)
- ... je predmet sporočanja in komuniciranja.
- ... je sporočilo, ki je v danem znakovnem sistemu pravilno (sintaktična pravilnost), ima nedvoumen pomen (vsebinska ali semantična pravilnost) in ima za prejemnika uporabno vrednost (pragmatičnost).

## Podatek in informacija



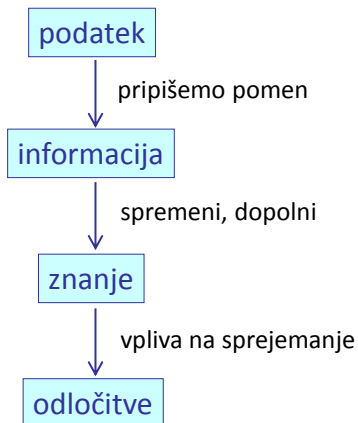
## O podatku in informaciji

- Podatek je poljubna predstavitev s pomočjo simbolov ali analognih veličin, ki ji je pripisan, ali se ji lahko pripiše nek pomen.
- Informacija je znanje, opisano s podatki (dejstva, dogodki, stvari, procesi ali ideje, vključno s koncepti), ki imajo v okviru nekega konteksta določen pomen.
- Karakteristika, ki loči podatek od informacije, je informacijska vsebina; izraža se z uporabno vrednostjo informacije.

## Prejeta informacija

- Informacija je novo spoznanje, ki ga človek doda svojemu poznavanju sveta.
- Prejeta informacija spremeni naše znanje, vpliva na naše odločitve in ravnanje.
- Informacija predstavlja neko novo znanje (zmanjša negotovost, ki je posledica pomanjkanja informacije). Tako mora biti naslovljena na konkretnega uporabnika.

## Podatek, informacija, znanje



## Informacijska enačba

- Börje Langefors – informacijska enačba
  - Informacija je novo spoznanje, ki ga človek doda svojemu poznavanju sveta. Odnos med informacijo, podatki, časom in interpretatorjevim znanjem predstavlja informacijska enačba:

$$I = i(D, S, t)$$

- I – informacija, ki jo posredujejo podatki
- i – informacijska funkcija
- D – podatki
- S – prejemnikovo znanje
- t – čas, ki je na voljo prejemniku za interpretacijo podatkov

## Informacijska enačba

- **Langeforsovi zaključki**
  - Podatki posredujejo informacijo prejemniku, katerega znanje je konsistentno z izbrano predstavitvijo podatkov in modelom sveta, na katerega se nanašajo.
  - Če je količina podatkov tako velika, da se jih v času, ki je na voljo za ukrepanje na njihovi osnovi, ne da interpretirati, se lahko zgodi, da s podatki ni posredovana nobena informacija.

## Kakovost informacij

- **Dostopnost**
  - Brez dostopa do informacij, so le-te brez koristi.
- **Točnost**
  - Le točne informacije pravilno dopolnijo naše znanje.
- **Pravočasnost**
  - Časovno odvisne informacije so koristne le, če jih dobimo pravočasno.
- **Popolnost**
  - Popolnosti informacij ni moč zagotoviti; težimo k čimbolj popolnim.
- **Zgoščenost**
  - Primeren obseg informacij glede na potrebo in razpoložljiv čas.
- **Ustreznost**
  - Prave informacije glede na potrebe; tiste, ki olajšajo sprejem odločitev.
- **Razumljivost**
  - Primerna izbira predstavitve podatkov, konsistentna z našim znanjem.

## Obdelava informacij in podatkov

- Človek in računalnik sprejemata podatke iz okolja.
- Človek podatkom pripiše nek pomen in s tem dopolni svoje znanje.
- Računalnik s pomočjo programov podatke predela:
  - V obliko, kot si jo želi človek.
  - V obliko, ki je primerna za nadaljnjo (strojno) računalniško obdelavo.

## Obdelava informacij in podatkov

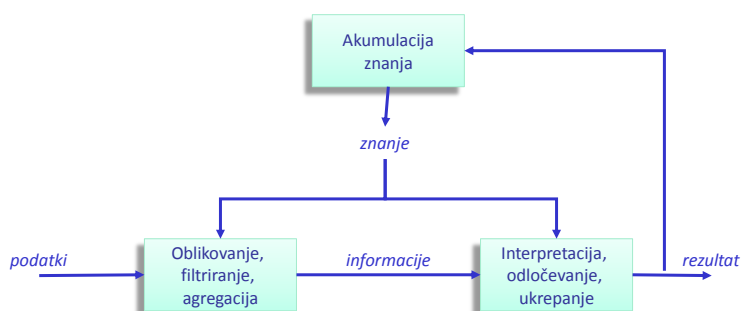


## Obdelava informacij in podatkov

- Če so podatki predstavljeni v obliki, ki ni konsistentna z znanjem prejemnika, jih le-ta ne more postaviti v kontekst in jih zato le stežka obdela (ali jih sploh ne more).
- Podatkom, predstavljenim v primerni obliki, človek avtomatsko pripiše pomen – obdeluje le informacije in ne podatkov!

նօճիւնային և արհեստական ինֆորմատիկա ↔ Informatika v medijih

## Podatek, informacija, znanje



## Informacija in sporočilo

- Z informacijo torej odpravljamo nedoločenost – pove nam nekaj, česar še nismo vedeli.
- Sporočilo:
  - Nastane, ko podatke komu sporočimo, posredujemo.
  - Lahko sporoča nekaj, kar prejemnik že ve (v tem primeru ne zmanjša nedoločenosti, negotovosti).

## Teorija informacij (Claude Shannon, 1948)

- V informacijski teoriji je informacija opredeljena kot:
  - znanje, ki zmanjša negotovost, povezano s pojavom določenega dogodka iz končne množice možnih dogodkov.
- Informacija, ki jo pridobimo s tem, ko zvemo, da se je pripetil določen dogodek, se izračuna po formuli:

$$I = -\log_2 p(x) \text{ [bit]}$$

- Podatek je v zgornji definiciji sporočilo, da se je zgodil dogodek  $x_i$ ; znanje je prepoznavanje verjetnosti nastopa posameznih dogodkov  $p(x_i)$ ,  $i = 1..n$ .



## Teorija informacij

- Enota informacije: **1 bit** (binary digit)
- 1 bit informacije dobimo z odgovorom na vprašanje, pri katerem sta možna natanko dva enako verjetna odgovora.

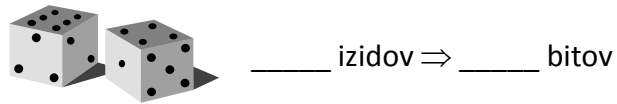
## Merjenje informacij



$$I = \log_2 n$$

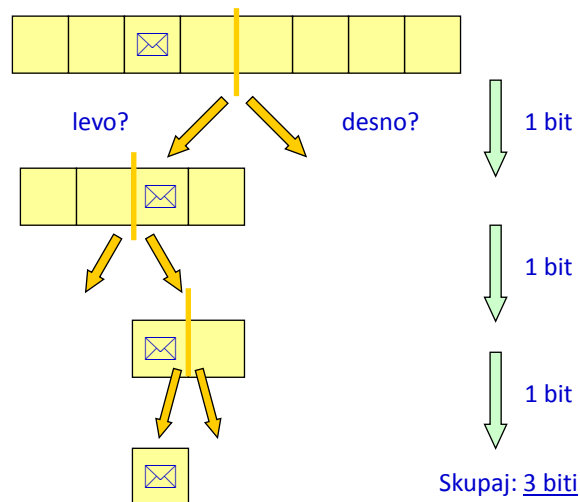
$I$ : količina informacije [bit]  
 $n$ : število *enakovrednih* izidov

## Merjenje informacij: primer



- Enakovrednih izidov je  $n = 6 \times 6 = 36$ .
- Količina informacije:  
 $I = \log_2 n = \log_2 36 = \log_{10} 36 / \log_{10} 2 = 1,556 / 0,301 = \underline{5,17 \text{ bitov}}$

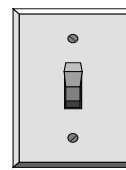
## Zakaj $\log_2$ ?



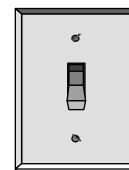
## Bit – enota informacije

### ■ Bit (binarna številka)

- Najmanjša enota informacije.
- Lahko zavzame vrednosti 0 in 1.
- Z biti lahko predstavimo številke, kode in ukaze.



ON



OFF

## Bit in večje enote

1 b bit (zelo majhna enota)

1 B byte (zlog, bajt, oktet) = 8 bitov

1	0	0	1	1	0	1	1
---	---	---	---	---	---	---	---

1 KB kilobyte =  $2^{10}$  B = 1.024 B =  $1.024 \times 8$  b = 8192 b

1 MB megabyte =  $2^{20}$  B = 1.024 KB = 1.048.576 B

1 GB gigabyte =  $2^{30}$  B = 1.024 MB = 1.073.741.824 B

1 TB terabyte =  $2^{40}$  B = 1.024 GB

## Bit in večje enote

Izjemoma:

1 kB kilobyte =  $10^3$  B = 1.000 B = 1.000 × 8 b = 8000 b

Pri prenosu podatkov:

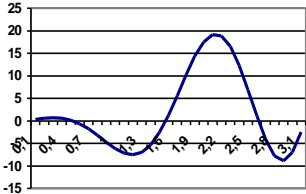
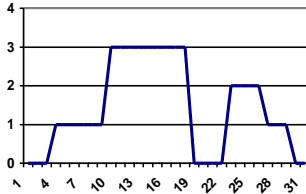
1 Kb kilobit =  $2^{10}$  b = 1.024 bitov

1 kb kilobit =  $10^3$  b = 1.000 bitov

Uporaba: hitrost prenosa podatkov [kbit/s]

## Predstavitev podatkov v računalniku

## Vrste informacij

ANALOGNE	DIGITALNE
<b>zvezne</b>	<b>diskretne</b>
lahko jih izmerimo	lahko jih preštejemo
	

## Predstavitev informacij

ANALOGNA	DIGITALNA
ura s kazalci 🕒	digitalna ura [16:01:33]
merilnik hitrosti	števec kilometrov
živosrebrni termometer	digitalni termometer
velikost napetosti	napetost je/ni
svetlobna jakost	svetloba je/ni
jakost magnetnega polja	orientacija mag. polja
globina/odmik zarez	vdolbina je/ni



## Predstavitev podatkov v današnjih računalnikih

- diskretna predstavitev
- z dvema stanjema (binarno, dvojiško): 0,1
- Zlog (byte): 8 bitov

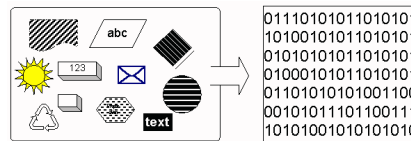
1	0	0	1	1	0	1	1
---	---	---	---	---	---	---	---

## Vrste podatkov

- števila
- znaki (besedila)
- slike
- zvok in glasba
- video
- ...

## Kodiranje podatkov

- Vsak podatek lahko predstavimo v dvojiški (binarni) številski obliki.
- Podatek razgradimo na niz elementarnih vprašanj, na katere odgovorimo z DA ali NE in potem ta niz predstavimo z dvojiškim številom.
- Za vsak bit vemo, kateremu vprašanju ustreza.
  - Kodiranje števil
  - Kodiranje pisnih znakov
  - Kodiranje slike
  - Kodiranje zvoka



## Predstavitev (celih) števil

- Dvojiški (binarni) številski sistem

Število (desetiško)	Dvojiško	Zlog
0	0	0000000
1	1	0000001
2	10	0000010
3	11	0000011
4	100	0000100
5	101	0000101
15	1111	0001111
55	110111	0011011
255	1111111	1111111

## Predstavitev znakov

- Kodiranje:
  - enolična preslikava: znaki → števila

Primer: 'A' ↔ 65

- ASCII (American Standard Code for Information Interchange) – 7-bitni standard, ki vsebuje:
  - male in velike črke angleške abecede
  - številke '0'..'9'
  - posebne znake '+', '-', '.', '\*', '#', '@'...
  - presledek ' ' (koda 32)
  - kontrolni znaki (npr. <CR>, <LF>)

065	A
066	B
067	C
068	D
069	E
070	F
071	G

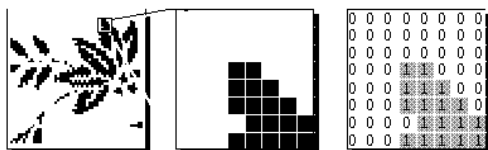
## Kodne tabele

- ISO Latin1
  - 8-bitna koda
  - Razširitev ASCII (ASCII kompatibilna)
  - Vsebuje znake za zahodnoevropske države
- ISO Latin2
  - Enako kot ISO Latin1, vendar za srednjeevropske države
- Cirilica
  - Mnogo različnih kodiranj. Najpogostejši:
    - KOI8 na UNIX sistemih
    - Windows-1251 pod Windowsi
- Unicode (16 bitna)
  - vse znake iz vseh jezikov(!) (UTF8, UTF16, ...)



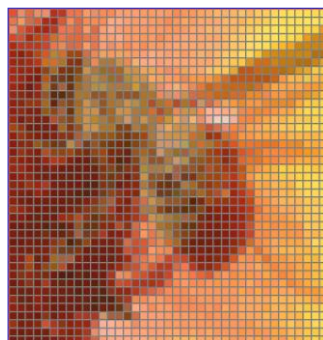
## Predstavitev slik

- Dve (osnovni) vrsti slik
  - Bitne slike, vektorske slike
- Večina slik je bitnih, njihova predstavitev v računalniku je zelo preprosta in neposredna
  - Velikost (ločljivost, *rezolucija*), barvna globina



## Predstavitev slik

- Primer bitne slike:
  - Velikost: 128×128 slikovnih elementov (pikslov, pik)
  - Barvna globina: 256 barv (vsaka pika en zlog)



## Predstavitev pik



Monokromatsko, 2 barvi  
(1 bit)



256 nivojev sivine  
(8 bitov, 1 zlog)



16 barv  
(4 biti, 1/2 zloga)



256 barv  
(8 biti, 1 zlog)



"True color" **RGB** (16,7 mil. barv)  
(24 bitov, 3 zlogi: **8+8+8 b**)

## Vaja dela... (1)

- Koliko pomnilnika potrebujemo za sliko ločljivosti 800x600 pikselov v 4096 barvah?

## Vaja dela... (2)

- Na voljo imamo 10MB pomnilnika. Koliko slik ločljivosti 1024x768 pikselov v TrueColor načinu lahko shranimo?

## Vaja dela... (3)

- Največ koliko barv ima lahko slika, da v 5MB pomnilnika shranimo 10 slik ločljivosti 800x600 pikselov?

## Formati shranjevanja slik

- Piksle, ki tvorijo bitno sliko, je potrebno shraniti (v pomnilnik, npr. na trdi disk) na način, ki omogoča naknadno nedvoumno rekonstrukcijo slike.
- V ta namen uporabljamo različne (standardne) formate zapisa bitnih slik
  - Shranjevanje brez izgube informacij
  - Shranjevanje z izgubo informacij (stiskanje)

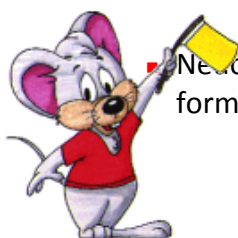
## Formati shranjevanja slik

- BMP
  - Osnovni format za shranjevanje bitnih slik v operacijskih sistemih Windows.
- TIFF (Tagged Image File Format)
  - Standarden format za shranjevanje bitnih slik (v različnih sistemih in okoljih)
  - Primeren za shranjevanje (predvsem skeniranih) slik – podpira poljubne velikosti, ločljivosti in barvne globine

## Formati shranjevanja slik

### ■ GIF (Graphics Interchange Format)

- Omogoča stiskanje podatkov o sliki brez izgub (po algoritmu LZW)
- Omogoča prozorno (transparentno) ozadje, animacije
- Pogosto uporabljen format v spletu



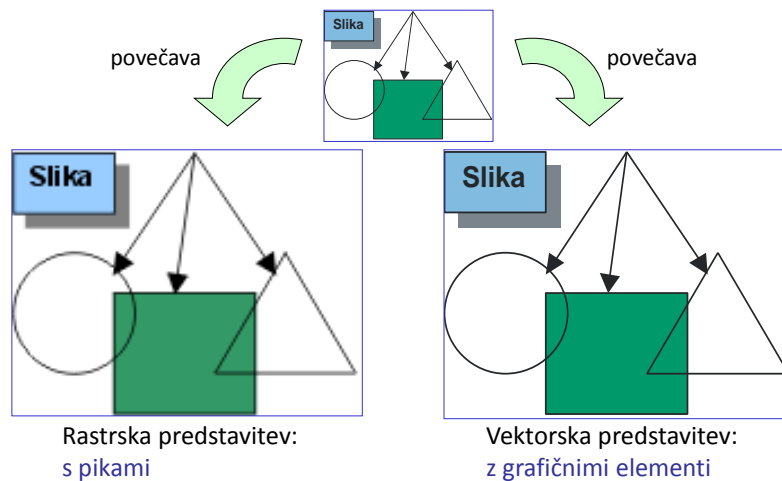
- Navedeni problemi s patentom oz. licenciranjem formata, danes ga nadomešča format PNG

## Formati shranjevanja slik

### ■ JPEG (Joint Photographic Experts Group)

- Shranjuje podatke o slikah z izgubami (shranjena slika ni enaka originalu!)
- Namenjen shranjevanju "naravnih" slik (fotografije), barvnih ali sivinskih
- Ni primeren za shranjevanje "nerealističnih" slik (diagrame, risani elementi, ...)
- Najpogosteje uporabljen v spletu za fotografije

## Vektorska predstavitev slik



## Shranjevanje vektorskih slik

- Shranjuje se (objektni) opis slike
  - Posamezne grafične elemente shranjujemo s pomočjo koordinat, dolžin, usmeritev, formul za izrisovanje, ...
  - Uporabljajo se različni formati zapisa (odvisno od programa)
  - PostScript, PDF (Portable Document Format)
    - Formata, ki ju je razvil Adobe Systems. PostScript je najpogosteje uporabljen v okoljih Unix, PDF pa je postal de-facto standard za dokumente na internetu.

## Besedilo: znakovna in grafična predstavitev

Znakovna predstavitev:  
npr. koda ASCII

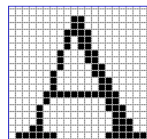
Grafična predstavitev:  
rastrsko, kot slika

$$A = 65_{10} = 01000001_2$$

- Bistvena razlika v:
  - Količini podatkov
  - Zmožnosti obdelave



OCR: Optical Character Recognition  
optično razpoznavanje znakov

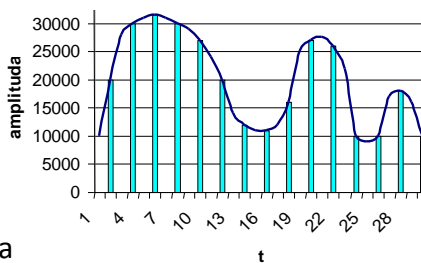


```

000000000000000000000000
000000001100000000000000
000000001100000000000000
000000001110000000000000
000000001110000000000000
000000010110000000000000
000000010011000000000000
000000100011000000000000
000000100000110000000000
0000011111111000000000
0000010000001100000000
0000010000001100000000
0000010000001100000000
.....
0111110000001111110
0000000000000000000000
    
```

## Predstavitev zvoka

- Vzorčenje
  - A/D pretvornik
- Frekvence vzorčenja
  - 11 kHz, 22 kHz, kHz, 48 kHz, 96 kHz
- Predstavitev amplitude zvoka
  - 8 ali 16 bitov
- Predstavitev ene sekunde zvoka (16 bitov, 44.1 kHz):
  - $1\text{ s} \times 44100\text{ meritev/s} \times 2\text{ B} \approx 88\text{ KB}$  (stereo: 176 KB)



## Vaja dela... (4)

- Kakšno je razmerje med količino potrebnega pomnilnika za predstavitev zvoka v CD audio in zvoka, posnetega v 128kbit/s MP3 formatu?

## Predstavitev glasbe

- S podatki predstavimo note, odigrane na instrumentu.
- MIDI (Musical Instrument Digital Interface)

### Primer:

10010001	zaigraj na kanalu 1
00111100	ton C1
01000000	z glasnostjo 64



## Ostale predstavitve

- Specifične predstavitve podatkov za specifične potrebe
- Kombinacije naštetih načinov predstavitev
  - Video: slika (bitna) in zvok
  - Dokumenti: besedilo, slike (bitne in vektorske)
  - Multimedijske predstavitve: besedilo, slike, zvok, ...

## Vaja dela... (5)

- Koliko pomnilnika je potrebno za (nestisnjen) video dolžine 1,5 h, če je ločljivost slike 640x360 pikselov, št. slik na sekundo 25, slike so v TrueColor načinu?

## Vaja dela... (5)

- Koliko pomnilnika je potrebno za (nestisnjen) video dolžine 1,5 h, če je ločljivost slike 640x360 pikselov, št. slik na sekundo 25, slike so v TrueColor načinu?
- Kako hitro omrežno povezavo potrebujemo, da bi lahko video v taki obliki spremljali preko omrežja?

## Različni podatki v pomnilniku

$01010010_2 = 52_{16} = 82_{10}$



Število  
(8-bitno)

Znaki H,A,L,O: **HALO**



Besedilo  
(ASCII)

- Računalniški program "ve", kakšne vrste podatki so shranjeni v posameznih delih pomnilnika !



Slika  
16x16 pik  
(1-bitna,  
črno-bela)

### POMNILNIK

0	1	1	0	0	1	1	0
0	1	0	1	1	1	1	0
1	0	0	0	0	1	0	1
1	1	0	1	1	0	1	1
0	0	0	0	0	0	0	0
0	1	0	1	0	0	1	0
1	0	0	0	1	0	1	1
0	1	1	0	0	1	1	0
0	1	0	0	1	0	0	0
0	1	0	0	0	0	0	1
0	1	0	0	1	1	0	0
0	1	0	0	1	1	1	1
1	1	0	1	1	0	1	1
0	0	0	0	0	0	0	0
0	1	0	1	0	0	1	0
1	0	0	0	1	0	1	1
0	1	1	0	0	1	1	0
0	1	0	1	1	1	1	0
1	0	0	0	0	1	0	1
0	0	0	0	0	0	0	0
0	1	0	1	0	0	1	0
1	0	0	0	1	0	1	1

H  
A  
L  
O