

ORGANIZACE DAT V POČÍTAČI

Naučíme se, jak jsou v počítači všechna data uložena. Víme, že v počítači máme programy – např. Word, Excel, že v každém počítači (i telefonu) musí být nějaký operační systém, dále máme v počítači fotografie, hudbu, filmy, hry a mnoho dalších dat. A tohle všechno má počítač uspořádáno, aby měl přehled, kde, co je.

Má tedy něco jako „adresář“ (tabulky FAT) a tam si zapisuje kde je daný soubor uložený. Počítačové viry, nebo hackeři ví, že stačí tuhle tabulku vymazat a počítač už neví kde soubory leží a nemůže je tedy najít a pracovat s nimi. Smazání dat na disku by totiž trvalo třeba i hodiny, zatímco vymazání tabulky (adresáře) trvá zlomek vteřiny.

Tak teď si povíme, co to ten soubor vlastně je.

SOUBOR = anglicky FILE

Soubor je v počítači NEJMENŠÍ SOUVISLÝ BLOK INFORMACÍ.

V počítači je vše uloženo ve formě souborů.

Soubor představuje "knihu" plnou textů, obrázků, zvuků.

Názvy souborů.

Každý soubor má svůj název (jméno a příjmení) a příponu. Podle názvu a přípony se můžeme dovědět, co je soubor zač. Například soubor *Kniha.txt* bude zřejmě nějaký textový soubor, zatímco soubor *Foto_Marie.jpg* bude patrně fotografie Marie.

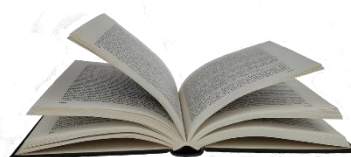
Pravidla pro názvy souborů (většinou si názvy volíme sami).

- Nepoužívejte mezery. Proto všude, kde chcete ve jménu souboru naznačit mezeru, dejte mínus, nebo podtržítka. 2014-dovolena.pdf
- Nepoužívejte češtinu ve jménech souborů a zapisujte jména souborů jen s malými písmeny. Jestli to někde funguje, neznamena to, že to bude fungovat všude. Název obrázku kočička.jpg je lepší nahradit kocicka.jpg
V cizině totiž nemusíte na klávesnici najít češtinu.

- Používejte taková jména, u nichž poznáte, co soubor obsahuje. Soubory se jmény typu sdsfo24x.html nejsou to pravé. Na druhou stranu jsou pochopitelná jména souborů typu kapitola1.html (2, 3 atd.) třeba u stránek seriálů, které se už nebudou měnit.
- V názvu souboru se nesmí používat znaky "/ \ * : ? # < > .

Typy souborů (pro vysvětlení malým dětem představíme typy souborů pomocí obrázků):

Textový soubor



Obrázkový soubor



Fotografie



Zvukový soubor



Filmový soubor

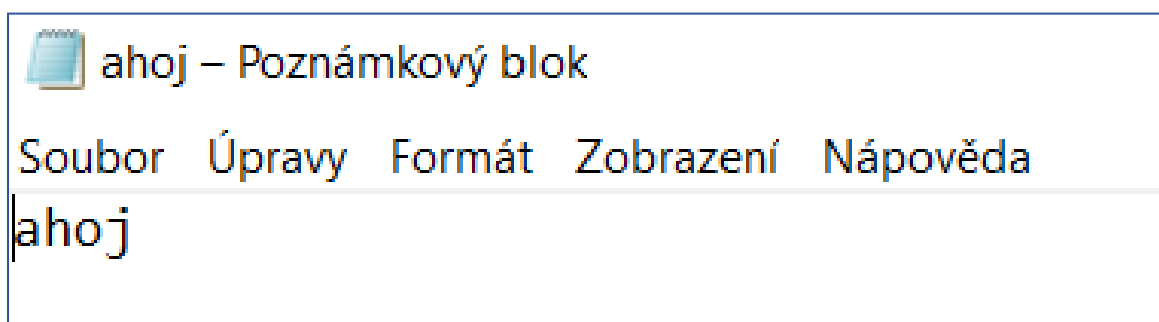


Přípona souboru (napovídá o jaký soubor se jedná a co se dá s takovým souborem dělat):

Bat	spustitelný, textový soubor
Exe	spustitelný soubor – program
Com	spustitelný soubor (OS DOS)
Html	zdrojový kód webové stránky
Doc	textový soubor – Word
XLS	tabulka – Excel
JPG	obrázek, fotografie
MP3	hudební soubor
ZIP	komprimovaný soubor (sbalený, stlačený)
ARJ	komprimovaný soubor
SYS	systemový soubor (většinou součást OS)

Velikost souboru – velikost souborů se měří (udává) v bajtech.

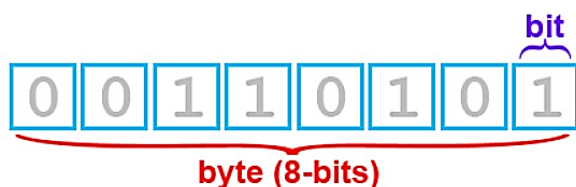
Když v poznámkovém bloku napíšeme slovo Ahoj a uložíme, uvidíme, že vznikne soubor o velikosti 4 bajty.



Bajt, původním, anglickým zápisem byte, zřídka také slabika, je v informatice základní jednotka kapacity počítačové paměti a objemu počítačových dat. Označuje zpravidla osm bitů, které tvoří osmiciferné binární číslo v rozmezí 0–255.

U prakticky všech dnešních počítačů obsahuje 1 bajt právě 8 bitů.

Počítače si mezi sebou posílají sekvence nul a jedniček. Takže když chce jeden počítač poslat druhému počítači písmeno A, tak pošle 01000001.



Aby se to nepletlo, tak existuje tabulka, kde jsou kombinace nul a jedniček pro písmena, číslice i speciální znaky už nastavené:

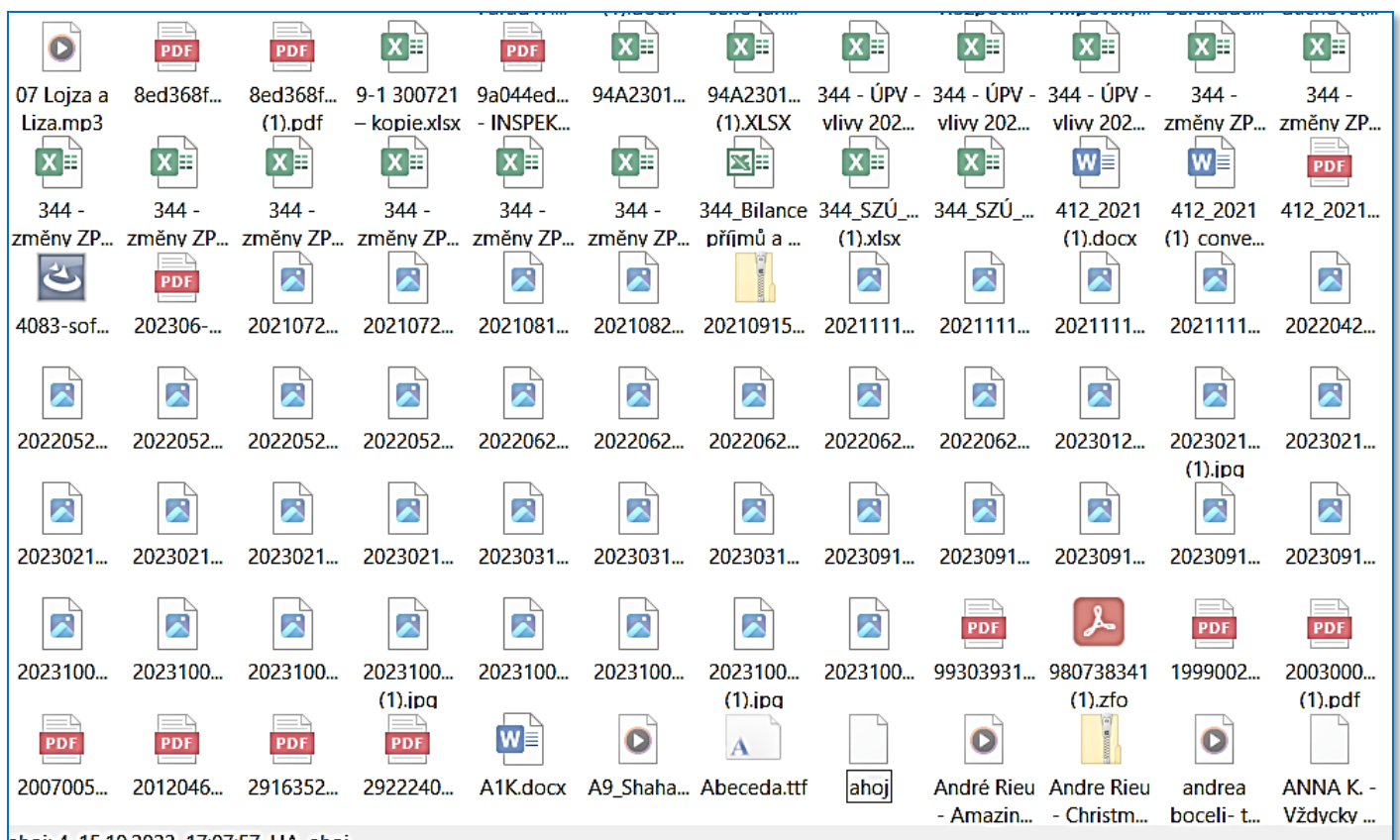
ASCII Alphabet Characters

Symbol	Decimal	Binary	Symbol	Decimal	Binary
A	65	01000001	a	97	01100001
B	66	01000010	b	98	01100010
C	67	01000011	c	99	01100011
D	68	01000100	d	100	01100100
E	69	01000101	e	101	01100101
F	70	01000110	f	102	01100110
G	71	01000111	g	103	01100111
H	72	01001000	h	104	01101000
I	73	01001001	i	105	01101001
J	74	01001010	j	106	01101010
K	75	01001011	k	107	01101011
L	76	01001100	l	108	01101100
M	77	01001101	m	109	01101101
N	78	01001110	n	110	01101110
O	79	01001111	o	111	01101111
P	80	01010000	p	112	01110000
Q	81	01010001	q	113	01110001
R	82	01010010	r	114	01110010
S	83	01010011	s	115	01110011
T	84	01010100	t	116	01110100
U	85	01010101	u	117	01110101
V	86	01010110	v	118	01110110
W	87	01010111	w	119	01110111
X	88	01011000	x	120	01111000
Y	89	01011001	y	121	01111001
Z	90	01011010	z	122	01111010

Složky – složky jsou něco, jako šanony, krabice, pořadače a do nich si ukládáme soubory.



Takhle můžeme ve Windows vidět soubory a složky:

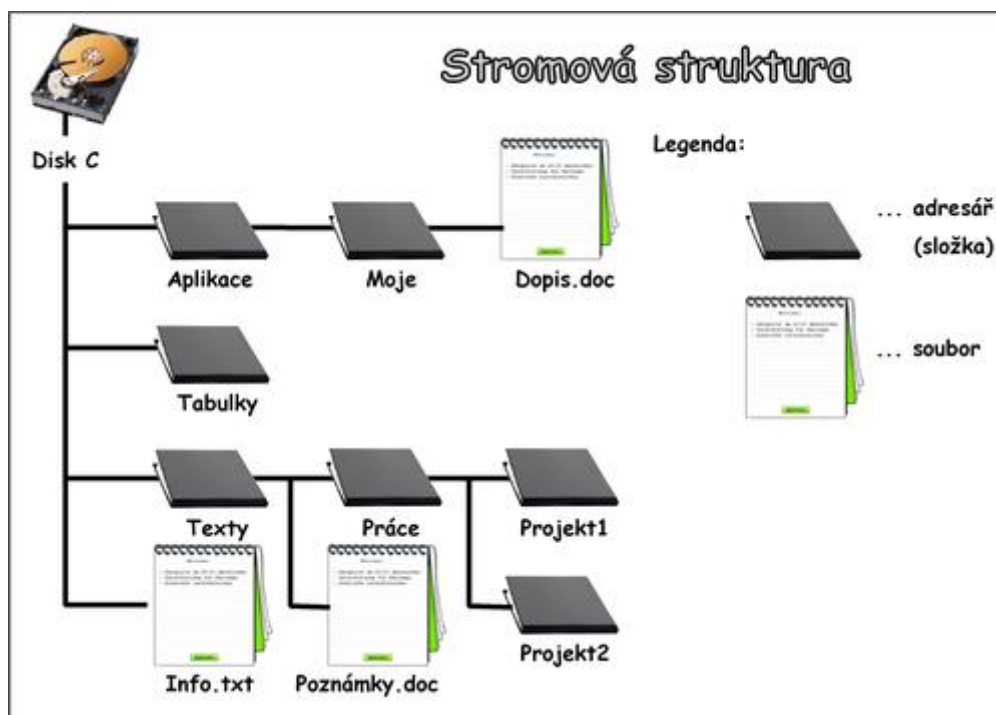


Atributy souboru – většinou atribut nastavují příslušné programy, ale můžeme atribut nastavit i ručně. A co to je? Značka, která určuje, co se může se souborem dělat, jestli je soubor viditelný, nebo skrytý apod.

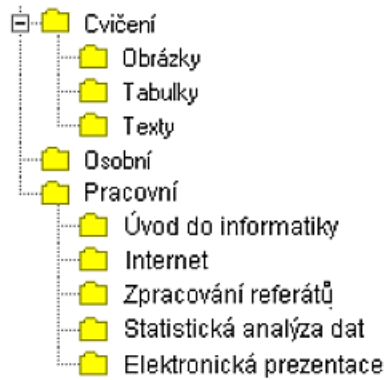
Název / Přípona	Velikost	Typ	Datum	Čas	Atributy
20230911_160912.jpg	2 054 630	JPG Soubor	11.09.2023	16:22:14	HA
20230911_160935.jpg	2 093 521	JPG Soubor	11.09.2023	16:25:35	HA
20230911_160949.jpg	2 057 651	JPG Soubor	11.09.2023	16:23:54	HA
20231001_120557.jpg	2 355 921	JPG Soubor	04.10.2023	10:00:14	HA
20231001_120616.jpg	2 481 561	JPG Soubor	04.10.2023	10:02:31	HA
20231002_122949.jpg	3 524 002	JPG Soubor	02.10.2023	12:34:53	HA
20231002_122958 (1).jpg	4 090 236	JPG Soubor	02.10.2023	12:41:47	HA
20231002_122958.jpg	4 090 236	JPG Soubor	02.10.2023	12:34:58	HA
20231002_123028.jpg	2 778 570	JPG Soubor	02.10.2023	12:38:24	HA
20231002_123117 (1).jpg	3 803 112	JPG Soubor	02.10.2023	12:37:41	HA
20231002_123117.jpg	3 803 112	JPG Soubor	02.10.2023	12:37:37	HA
99303931.pdf	377 657	Microsoft Edge PDF Document	09.03.2023	21:00:07	HA
980738341 (1).zfo	787 226	ZFO Soubor	02.01.2022	12:42:50	HA
1999002087.pdf	168 725	Microsoft Edge PDF Document	06.06.2023	18:31:54	HA
2003000693 (1).pdf	334 843	Microsoft Edge PDF Document	06.06.2023	18:30:54	HA
2007005855.pdf	126 714	Microsoft Edge PDF Document	06.06.2023	18:25:30	HA
2012046277.pdf	357 421	Microsoft Edge PDF Document	06.06.2023	18:24:39	HA
2916352685.pdf	610 940	Microsoft Edge PDF Document	18.01.2022	17:42:38	HA
2922240536.pdf	508 559	Microsoft Edge PDF Document	13.07.2022	21:51:36	HA
A1K.docx	12 086	Microsoft Word Document	12.11.2021	12:14:49	HA
A9_Shaharazad_kankán-final.mp3	7 720 858	MP3 Soubor	20.07.2023	15:06:13	HA
Abeceda.ttf	78 260	TrueType Font file	05.10.2023	7:22:11	HA
ahoj	4	Soubor	15.10.2023	17:07:57	HA

Kořenový adresář (root directory)

- v unixových systémech se kořenový adresář označuje znakem lomítko (/) a je společný pro všechna připojená média
- v operačních systémech Microsoft Windows (též DOS) má každý svazek (logická disková jednotka) svůj kořenový adresář a jeho označení se skládá z označení jednotky (písmeno latinky a dvojtečka) a zpětného lomítka (\), například C:\



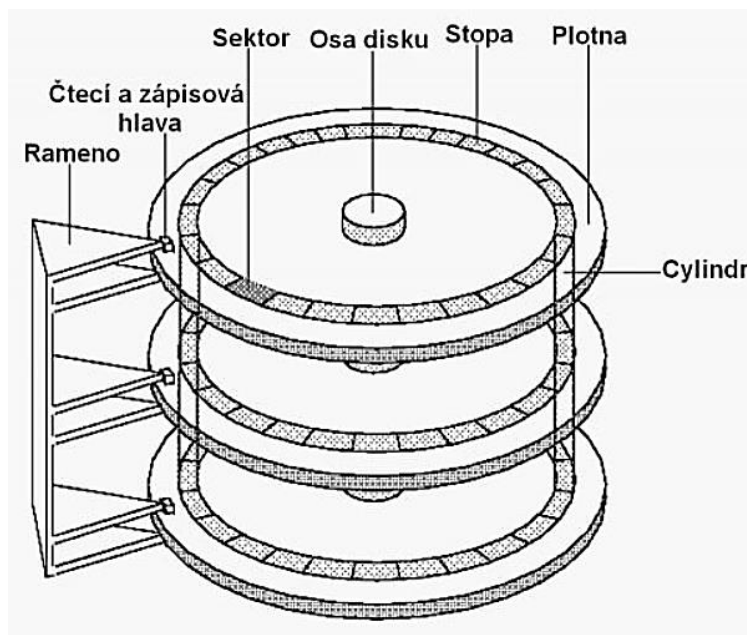
H:



A jak jsou data uložena na disku? Ukážeme si to na klasickém hard disku.

Na každé stopě je několik sektorů (dříve obvykle po 512 bajtech, dnes je běžná velikost sektoru 4096 bajtů), přičemž starší disky mají stejný počet sektorů na obvodu i na vnitřní stopě, z čehož vyplývá, že hustota záznamu je u středu plotny větší.

Protože se soubory ukládají do sektorů s velikostí 4096 bajtů – tak i náš soubor Ahoj o velikosti 4 bajty zabere ve skutečnosti místo 4096 bajtů.



Co je to fragmentace disku?

Disky zapisují soubory větší než jeden cluster nahodile na volná místa disku, tím vzniká fragmentace disku ("rozházení" dat po disku). Po nějaké době jsou volná místa po disku roztroušena a promixována s již plnými, a tak vznikají při zápisu a čtení dat čím dál větší prodlevy (čtecí hlavička hledá správné místo) a klesá i výkon disku.

Tento problém řeší defragmentace disku, která se provádí z prostředí operačního systému a jejímž účelem je poskládat roztroušené fragmenty jednotlivých souborů tak, aby byly co nejbližší sobě. Data na disku se zahustí a tím se zrychlí i operace s daty.