

Technické podmienky tvorby formátu Comma Separated Values (CSV)

Formát Comma Separated Values (CSV) sa vytvára tak, že

1. sa používa znaková sada podľa § 13 písm. c),
2. príslušný súbor neobsahuje na začiatku signatúru Byte Order Mark (BOM), ktorou je znak U+FEFF kódovaný ako 0xEF, 0xBB, 0xBF,
3. pri prenose prostredníctvom siete Internet sa uvádza hlavička „Content-Type“, a to s hodnotou „text/csv; charset=UTF-8“,
4. ak súbor obsahuje v prvom riadku názvy položiek dátového zoznamu, táto skutočnosť je uvedená v opise údajov a súčasne je pri prenose prostredníctvom siete Internet do hlavičky podľa tretieho bodu doplnené pole „header=present“; názvy položiek v jednom súbore sa odlišujú,
5. názov súboru má obvykle príponu „.csv“,
6. ukončenia riadkov v súbore sa označujú znakom U+000A (LF) alebo dvojicou znakov U+000D a U+000A (CR LF); iné označenia sa nepoužívajú,
7. riadky začínajúce znakom „#“ (U+0023) sú považované za komentáre a pri čítaní súboru sa neinterpretujú ako údaje; ak prvá položka riadku začína znakom U+0023, položka sa uzatvorí do úvodzoviek,
8. položka dátového zoznamu, tvorená dvojicou bezprostredne nasledujúcich znakov čiariok „“ (U+002C) alebo jednou čiarkou na začiatku riadku, sa interpretuje ako „chýbajúci údaj“,
9. položka „prázdny reťazec“ dátového zoznamu sa zaznamenáva dvojicou bezprostredne nasledujúcich úvodzoviek „“ (U+0022),
10. znaky medzera (U+0020) a tabulátor (U+0009) sa interpretujú voľne; ak sa má zachovať v položke určitý reťazec týchto znakov, položka sa uzatvorí do úvodzoviek,
11. ak položka dátového záznamu obsahuje číslo v desatinnom tvare a na oddelenie desatinnej časti sa použije znak čiarka „.“ (U+002C), položka sa uzatvorí do úvodzoviek,
12. na oddelenie položiek sa používa znak čiarka „.“ (U+002C).