

Instrukcja przygotowywania danych do zdeponowania w kolekcji instytucjonalnej UPP w Repozytorium Otwartych Danych RepOD

- wszystkie informacje w plikach powinny być prezentowane w języku angielskim;
- w plikach Excela, również w plikach .ods, nie należy stosować wykresów - powinny one zawierać wyłącznie dane liczbowe, a wykresy można ewentualnie dodać w osobnym pliku graficznym;
- pliki należy przygotowywać w dwóch formatach, .xls (.xlsx) i .csv; .doc (.docx) i .odt;
- należy kategorycznie rozwijać lub wyjaśnić zastosowane w plikach skróty;
- pliki .csv powinny być rozdzielone przecinkami (separator pól), z kropką jako separatorem dziesiętnym w przypadku ułamków;
- pliki .csv - kodowanie powinno być w UTF-8 (tak jak i we wszystkich plikach tekstowych);
- jeśli nie ma uzasadnionych powodów do odstępstw, pierwszy wiersz powinien zawierać jedynie nazwy zmiennych, a same dane powinny być zorganizowane w kolumnach;
- nazwy zmiennych nie powinny zajmować kilku wierszy;
- wszystkie niepuste kolumny muszą mieć nazwy;
- tabele powinny być prezentowane w osobnych plikach;
- zmiennych liczbowych nie należy brać w cudzysłów;
- jeśli w plikach znajduje się spora liczba skrótów i oznaczeń symbolicznych, to należy zastosować legendę zawierającą informację, co dany skrót i oznaczenie oznacza. Ewentualnie można te informacje zawrzeć w pliku ReadMe;
- nazwy plików: tylko a-Z, 0-9 i ewentualnie podkreślnik _
- nie należy stosować kropek prócz tych, które rozdzielają nazwę od rozszerzenia; nie należy również stosować polskich znaków, myślników oraz znaków specjalnych. Odnosi się to również do danych wewnątrz archiwów;
- należy unikać prezentowania informacji za pomocą kolorów czy też formatowania tekstu;
- maksymalna wielkość pliku to 5 GB.

Opracowanie:

Wojciech Fenrich - Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego (ICM UW)

Krzysztof Gmerek - Uniwersytet Przyrodniczy w Poznaniu