Course title: Digital Humanities   Prof. J. Savoy

Serie 2. Literature and the French Theatre

_____

You will have two weeks to perform the following tasks.

1. Cite four differences between XML and HTMP standards?
2. Are both standards fully declarative one?

The practical questions: French Theater

With the code available and the folder /data/theatre-classique/, you can extract many plays (stored in XML format). However this folder is limited to storing plays, without additional information.

Focusing on the author names. The name is provided according to the pattern "LastName**,** Firstname" (e.g., "Smith**,** Adam").

How many author names can you find?

Are you sure that all these distinct names refer to distinct authors?

Can you reduce the variability around the author names?

Can you count the number of plays per author?

PS: The name "Rousseau" is the same as "ROUSSEAU"