

Esercitazione “Altiero Spinelli” in vista della prova teorica

Laboratorio di Informatica specialistica per Scienze dell'Antichità (Dr. Paolo Monella) Anno Accademico 2016/17, 1° semestre. Lezione del 27/11/2016



1. Cos'è il treebanking?
2. Cos'è il PoS tagging?
3. Cos'è la lemmatizzazione?
4. Spiega il significato del seguente codice TEI XML (non c'è bisogno di sciogliere l'abbreviazione “12C---O3---”, ma basta dire a cosa si riferisce):

```
<w ana="12C---O3---"  
lemma="doctor">doctoribus</w>
```

5. Spiega il significato delle componenti del seguente file di testo:

```
Gallia      doctor      12C---O3---  
est         esse        23---3--5-S  
omnis      omnis      23B-3-F--3
```

6. Cos'è un *charset*? Quali *charset* abbiamo studiato?
7. Tramite quale *charset* è possibile codificare il greco politonico in modo da visualizzare correttamente il testo su più computer, a prescindere dal loro sistema operativo (Windows, Microsoft, Linux etc.) e dal software installato in essi?
8. Cosa vuol dire effettuare la *tokenization* su un testo?
9. Questo è un algoritmo? “Prendi tre numeri, i più piccoli che puoi, e combinali creativamente in modo da far venire fuori un numero buffo da pronunciare”. Spiega perché è o non è un algoritmo.
10. Indica e descrivi tre tra i linguaggi di markup che abbiamo studiato.
11. Indica tre *corpora* testuali classicistici (cioè portali o combinazioni corpus/programma di interrogazione) tramite cui sia possibile effettuare ricerche lessicali lemmatizzate.
12. Indica tre *corpora* testuali classicistici (cioè portali o combinazioni corpus/programma di interrogazione) tramite cui sia possibile effettuare il PoS tagging del testo.
13. Cos'è un indice di frequenza?
14. I tre principali operatori delle ricerche booleane sono AND, OR e NOT. Se una ricerca testuale richiede risultati che coprendano “parola1” AND “parola2”, quali risultati saranno visualizzati? Quali, se richiede “parola1” OR “parola2”? Quali, se richiede “parola1” NOT “parola2”?
15. Descrivi alcune delle applicazioni dell'informatica alla filologia testuale.