

## Prova teorica

Laboratorio di Informatica specialistica per Scienze dell'Antichità  
(Dr. Paolo Monella), Anno Accademico 2018/19, 2° semestre.  
Prova del 30 maggio 2019.

1. Spiega quali informazioni codifica il seguente frammento di codice TEI XML. (non c'è bisogno di sciogliere i sigla #G, #R, #0):  

```
<app>
  <lem wit="#G #R">sed</lem>
  <rdg wit="#0">si</rdg>
</app>
```
2. Spiega quali informazioni visualizza un Wordle come quello della Costituzione Italiana rappresentato qui a lato, e in che modo.
3. In cosa consiste il *distant reading* (rappresentato ad esempio dal portale *Voyant Tools*)?



3. Il *distant reading* è un'analisi\* effettuata con metodi computazionali su vasti corpora senza che ne sia necessaria <sup>la lettura integrale</sup> ~~la lettura integrale~~ (procedura del *close reading*). Il portale *Voyant Tools* usa programmi ~~es~~ <sup>presente</sup> algoritmi *gaze* <sup>di cui</sup> ~~di cui~~ <sup>quali</sup> ~~quali~~ <sup>può</sup> ~~può~~ effettuare il *distant reading*. 2/2

\* unita alla raccolta di dati quantitativi

1. Il tag `<app>` indica <sup>che</sup> ~~che~~ <sup>tratta</sup> ~~tratta~~ di una nota d'apparato. Il tag `<lem>`, <sup>ha</sup> ~~ha~~ <sup>cioè</sup> ~~cioè~~ lemma, <sup>ha</sup> ~~ha~~ <sup>come</sup> ~~come~~ attributo `wit` <sup>cioè</sup> ~~cioè~~ `access witness` il cui valore è costituito da 2 testimoni B ed R. Il tag `<rdg>` <sup>indica</sup> ~~indica~~ una variante rispetto alla lezione presente nel testo edito e <sup>ha</sup> ~~ha~~ <sup>come</sup> ~~come~~ attributo `wit` <sup>il cui</sup> ~~il cui~~ <sup>valore</sup> ~~valore~~ <sup>è</sup> ~~è~~ <sup>il</sup> ~~il~~ <sup>testimone</sup> ~~testimone~~ <sup>0</sup>. 2/2



## Prova teorica

Laboratorio di Informatica specialistica per Scienze dell'Antichità  
(Dr. Paolo Monella), Anno Accademico 2018/19, 2° semestre.  
Prova del 30 maggio 2019.

1. Spiega quali informazioni codifica il seguente frammento di codice TEI XML (non c'è bisogno di sciogliere i sigla #G, #R, #O):  

```
<app>
  <lem wit="#G #R">sed</lem>
  <rdg wit="#O">si</rdg>
</app>
```
2. Spiega quali informazioni visualizza un Wordle come quello della Costituzione Italiana rappresentato qui a lato, e in che modo.
3. In cosa consiste il *distant reading* (rappresentato ad esempio dal portale *Voyant Tools*)?



1) lem: è la variante accettata SED presente nei manoscritti G e R

rdg: è la variante non accettata SI, presente nel manoscritto O

wit = manoscritto

2/2

2) Un wordle visualizza all'interno di un testo digitalizzato le parole più frequenti; più le parole sono frequenti maggiore sarà la grandezza del carattere con cui sono riportate

2/2

3) Il *distant reading* è un metodo d'analisi statistico computazionale su grandi testi digitali. La lettura a distanza permette di analizzare grandi parti di testo senza "leggerle" (close reading)

1,8/2

## Prova teorica

Laboratorio di Informatica specialistica per Scienze dell'Antichità  
(Dr. Paolo Monella), Anno Accademico 2018/19, 2° semestre.  
Prova del 30 maggio 2019.

1. Spiega quali informazioni codifica il seguente frammento di codice TEI XML (non c'è bisogno di sciogliere i sigla #G, #R, #O):  

```
<app>
  <lem wit="#G #R">sed</lem>
  <rdg wit="#O">si</rdg>
</app>
```
2. Spiega quali informazioni visualizza un Wordle come quello della Costituzione Italiana rappresentato qui a lato, e in che modo.
3. In cosa consiste il *distant reading* (rappresentato ad esempio dal portale *Voyant Tools*)?



- 1) Nel documento riportato abbiamo i marker di apertura e chiusura **seguiti dall'indicazione della lezione accettata "sed"** di cui G ed R sono testimonianze; **al rigo successivo è invece riportata l'indicazione della lezione "si", una non accettata, di cui O è il testimone.** 1,7/2
- 2) Nel Wordle ~~viene~~ <sup>viene</sup> rappresentato graficamente il numero di occorrenze delle parole di un corpus. **più le parole sono grandi maggiori sono le occorrenze all'interno del testo.** Per esempio, nel caso riportato la parola più ricorrente è "legge". 2/2
- 3) Voyant Tools è un programma di *distant reading*, dunque permette di **venire e conoscere statisticamente delle occorrenze, senza ~~leggere~~ <sup>leggere</sup> il testo** (close reading). 1,5/2



## Prova teorica

Laboratorio di Informatica specialistica per Scienze dell'Antichità  
(Dr. Paolo Monella), Anno Accademico 2018/19, 2° semestre.  
Prova del 30 maggio 2019.

1. Spiega quali informazioni codifica il seguente frammento di codice TEI XML. (non c'è bisogno di sciogliere i sigla #G, #R, #O):  

```
<app>
  <lem wit="#G #R">sed</lem>
  <rdg wit="#O">si</rdg>
</app>
```
2. Spiega quali informazioni visualizza un Wordle come quello della Costituzione Italiana rappresentato qui a lato, e in che modo.
3. In cosa consiste il *distant reading* (rappresentato ad esempio dal portale Voyant Tools)?



1. Il frammento di codice TEI XML dice che il filologo ha scelto la variante "sed" e inserita nel testo è "sed" tramandata dai codici G ed R, mentre il manoscritto O presenta la variante "si", che non è stata accettata dal filologo che ha ricostruito il testo. 2/2

2. Il Wordle riprodotto in figura ci informa che le parole più presenti nella Costituzione italiana sono: "legge", "costituzionale", "Repubblica", "Stato", "Presidente", etc... 2/2

L'algoritmo che lo ha prodotto è basato sull'indice delle frequenze: le parole più frequenti sono stampate in carattere più grande, mentre, gradualmente, diminuisce parallelamente alla diminuzione della frequenza.

Ovviamente sono presenti, in questo caso, sia parole vuote che parole piene. Per ricavare informazioni sul testo possiamo non tener conto di quelle vuote (art., può, nome, etc.)

• Il distant reading è l'analisi letteraria di un testo <sup>letterario</sup> o di interi corpora testuali, svolta in maniera computazionale.

Esso consente agli studiosi di ricevere istantaneamente informazioni sul testo analizzato quali ad esempio l'indice di frequenza, e su questi dati poi si potranno svolgere due tipi di analisi: la prima di tipo stilometrico mediante l'eliminazione di parole ~~piene~~ <sup>piene</sup>, la seconda, a partire dall'eliminazione delle parole vuote, potrà fornire ulteriori informazioni come le relazioni che intercorrono tra le parole più frequenti, che posso essere personaggi, luoghi etc.

2/2

## Prova teorica

Laboratorio di Informatica specialistica per Scienze dell'Antichità  
(Dr. Paolo Monella), Anno Accademico 2018/19, 2° semestre.  
Prova del 30 maggio 2019.

1. Spiega quali informazioni codifica il seguente frammento di codice TEI XML (non c'è bisogno di sciogliere i sigla #G, #R, #0):  

```
<app>
  <lem wit="#G #R">sed</lem>
  <rdg wit="#0">si</rdg>
</app>
```
2. Spiega quali informazioni visualizza un Wordle come quello della Costituzione Italiana rappresentato qui a lato, e in che modo.
3. In cosa consiste il *distant reading* (rappresentato ad esempio dal portale *Voyant Tools*)?



1. lem = lemma - lezione accettata: sed, riportata dai MS G ed R  
 rdg = reading - lezione non accettata: si, riportata dal testimone O  
 wit = witness - testimone

2/2

3. Il distant reading è una tecnica d'analisi di un tipo statistico e computazionale di un testo in grandi corpora digitalizzati. È un metodo per ~~l'analisi del testo~~ analizzare i testi senza leggerli

2/2

2. Un wordle riporta le occorrenze presenti all'interno di un testo o di un'opera. La frequenza delle occorrenze è rappresentata graficamente attraverso la grandezza del carattere della parola. Più grande è il carattere, più frequente è la parola. Nel wordle riportato la parola più frequente è "legge".

2/2



## Prova teorica

Laboratorio di informatica specialistica per Scienze dell'Antichità  
(Dr. Paolo Monella), Anno Accademico 2018/19, 2° semestre.  
Prova del 30 maggio 2019.

1. Spiega quali informazioni codifica il seguente frammento di codice TEI XML. (non c'è bisogno di sciogliere i sigla #G, #R, #O):  

```
<app>
  <lem wit="#G #R">sed</lem>
  <rdg wit="#O">si</rdg>
</app>
```
2. Spiega quali informazioni visualizza un Wordle come quello della Costituzione Italiana rappresentato qui a lato, e in che modo.
3. In cosa consiste il *distant reading* (rappresentato ad esempio dal portale *Voyant Tools*)?



- 1) Lem = variante accettata nelle una edizione  
 rdg = reading / edizione / variante / **accettato / e**  
 wit = witness / Testimone (HS)

I HS #G e #R hanno 'sed'; il HS #O ha 'si'  
 - si tratta di un'applicazione dell'informatica alle filologie.

- 3) Il distant reading è un metodo di analisi di un testo basato su un tipo di analisi statistica e computazionale di grandi corpora testuali digitalizzati. Le letture e distanze causate nel **visualizzare**, viene **figgibile** ("close reading"), **grandi parole** di testi. **1,8/2**

- 2) Il ~~vol~~ Wordle <sup>mostre</sup> ~~mostra~~ le frequenze **delle parole** più utilizzate all'interno delle costituzioni italiane, mettendole in un salto. Per esempio, la parola "legge" è quella più frequente.

## Prova teorica

Laboratorio di Informatica specialistica per Scienze dell'Antichità (Dr. Paolo Monella).  
Anno Accademico 2018/19, 2° semestre. Prova del 30 maggio 2019.

1. Spiega quali informazioni codifica il seguente frammento di codice TEI XML (non c'è bisogno di sciogliere i sigla #G, #R, #0):  

```
<app>
  <lem wit="#G #R">sed</lem>
  <rdg wit="#0">si</rdg>
</app>
```
2. Spiega quali informazioni visualizza un Wordle come quello della Costituzione Italiana rappresentato qui a lato, e in che modo.
3. In cosa consiste il *distant reading* (rappresentato ad esempio dal portale Voyant Tools)?



- 1) Il codice TEI XML rappresenta un'applicazione dell'informatica alla filologia.  
 Il tag `<lem>` mostra che la variante "sed" è quella accettata nella mia edizione.  
 Il tag `<rdg>`, invece, mostra che "si" è un'altra variante (non accettata nella mia edizione).  
 wit = testimoni  
 1,0/2      1,0/2
- 2) Il Wordle rappresentato inolea le parole più utilizzate all'interno della Costituzione Italiana, visualizzando con un formato più grande le parole più frequenti.  
 2/2
- 3) Il *distant reading* è un'analisi effettuata con metodi statistici e computazionali (Es. stella "Amleto"), applicati a grandi corpora. Permette di analizzare tutta la corpora senza leggerla.  
 2/2