

Tehnologije i platforme za upravljanje elektronskim sadržajima i dokumentima

1. Objasniti pojam dokumenta, papirnog dokumenta i digitalnog dokumenta.
2. Šta su to metapodaci i navesti nekoliko primera metapodataka?
3. Kakve sve veze između dokumenta i metapodataka mogu postojati?
4. Koje su faze životnog ciklusa dokumenta?
5. Objasniti životnu fazu dokumenta korišćenje.
6. Objasniti životnu fazu dokumenta arhiviranje.
7. Objasniti pojmove upravljanje verzijama dokumenata, sekvencijalno i konkurentno važenje verzija.
8. Koja je osnovna namena sistema za upravljanje dokumentima?
9. Koje su funkcije sistema za upravljanje dokumentima?
10. Opisati Dublin Core format metapodataka.
11. Šta su protokoli za razmenu podataka?
12. Koje su osnovne karakteristike OAI-PMH protokola?
13. Šta su protokoli za udaljeno pretraživanje?
14. Koje su osnovne karakteristike Z39.50 i SRU protokola?
15. Čime se bavi oblast pronalaženja informacija (information retrieval)?
16. Koja je razlika između pronalaženja podataka (data retrieval) i pronalaženja informacija (information retrieval)?
17. Kako se razvijala oblast pronalaženja informacija?
18. Kakve arhitekture mogu imati sistemi za pretragu?
19. Koje vrste sadržaja mogu biti pretraživane putem sistema za pretragu?
20. Koji modeli za pretraživanje se koriste u sistemima za pretragu?
21. Koja je razlika između terma i tokena?
22. Šta je tokenizacija i koji problemi postoje u ovoj fazi pretprocesiranja?
23. Zašto se vrši „normalizacija“ reči?
24. Šta je to stemming?
25. Šta je to lematizacija?
26. Objasniti Bulov model pretraživanja.
27. Šta je to invertovani indeks i kako se kreira?
28. Objasniti procesiranje upita kod Bulovog modela.
29. Šta su pointeri za preskakanje?
30. Šta se može koristiti ako je potrebno podržati upite fraze?
31. Šta je to dvorečni indeks?
32. Šta je to pozicioni indeks?
33. Objasniti Vektorski model pretraživanja.
34. Šta je ocena relevantnosti?
35. Šta je frekvencija terma?
36. Šta je frekvencija dokumenta?
37. Šta je tf-idf?
38. Objasniti kreiranje težinske matrice.
39. Koje su razlike između Bulovog i Vektorskog modela pretraživanja?
40. Da li se relevantnost odgovora meri u odnosu na informacionu potrebu ili upit?
41. Šta je preciznost (eng. precision)?
42. Šta je povrat (eng. recall)?
43. Šta je F mera i zašto je ona relevantnija od korišćenja preciznosti i povrata?
44. Kako se može vršiti evaluacija performansi sistema za pretraživanje?
45. Šta je kapa mera?
46. Opisati A/B testiranje?
47. Šta je to Lucene?
48. Šta predstavljaju klase Document i Field u Lucene biblioteci?
49. Navesti osnovne karakteristike upitnog jezika Lucene-a.
50. Kako se implementira analiza (procesiranje) teksta pomoću Lucene-a?