

정보학개론

1. 오픈엑세스 유형에 대한 설명으로 옳은 것은?

- ① Bronze OA는 출판사 웹사이트에서만 무료로 논문을 이용할 수 있고 공유와 재사용을 허용하지 않는다.
- ② Gold OA는 저자가 자신의 프리프린트를 리포지터리에 셀프 아카이빙할 수 있도록 허용한다.
- ③ Gratis OA는 온라인 접근이 무료이며 재사용도 가능하다.
- ④ Green OA는 OA 학술지에 수록된 모든 논문에 무료 접근을 허용하고, 논문처리비용을 부과하는 경우가 많다.
- ⑤ Libre OA는 접근비용 장벽만을 제거한 상태로 무료 온라인 접근을 허용한다.

2. 대상 영역과 표준 메타데이터 스키마의 연결이 옳지 않은 것은?

- ① 교육 - GEM, KEM, LOM
- ② 도서관 - MARC, MODS, TEI Headers
- ③ 멀티미디어 - MPEG-7, PBCore
- ④ 박물관 - CCO, CDWA, VRA
- ⑤ 지리정보 - CSDGM, GILS

3. 동시출현분석에 대한 설명으로 옳지 않은 것은?

- ① 학문의 지적 구조 파악에 널리 활용된다.
- ② 동시출현단어분석, 공저자분석 등이 포함된다.
- ③ 주제 영역의 식별이나 관련성 분석에 활용된다.
- ④ 문헌 간의 관계를 분석하기 위해 동시인용빈도를 활용한다.
- ⑤ 연구자 간의 협력 관계를 조사·분석하기 위한 도구로 활용된다.

4. 텍스트 마이닝 분석기법으로 옳지 않은 것은?

- ① 감성분석
- ② 군집분석
- ③ 메타분석
- ④ 토픽모델링
- ⑤ 단어빈도분석

5. 전조합 색인과 후조합 색인에 대한 설명으로 옳지 않은 것은?

- ① 대부분의 컴퓨터 검색 시스템은 키워드를 사용한 후조합 색인을 채택하고 있다.
- ② 후조합 색인인 키워드 색인은 한 세트의 용어들로 색인어가 구성되는 색인 방식이다.
- ③ 전조합 색인은 색인어가 탐색 이전에 색인 작성자에 의하여 조합된 것이며, 열거식 분류가 대표적이다.
- ④ 전조합 색인은 복합 주제를 하나의 색인어로 표현하여 한 번의 검색으로 직접 접근할 수 있도록 한다.
- ⑤ 후조합 색인은 문헌의 내용을 여러 개의 단위 개념으로 색인하고, 탐색할 때 각 단위 개념을 조합하여 적합한 문헌을 검색할 수 있도록 하는 색인 방식이다.

6. IFLA가 제안한 메타데이터의 핵심요소 중에서 FRBR의 검색, 식별, 선택, 획득 기능에 전부 사용되는 것만을 <보기>에서 모두 고르면?

_____ < 보 기 > _____

- ㄱ. 날짜
- ㄴ. 저자
- ㄷ. 주제
- ㄹ. 자원명
- ㅁ. 자원식별자

- ① ㄱ, ㅁ
- ② ㄴ, ㄷ
- ③ ㄷ, ㅁ
- ④ ㄱ, ㄷ, ㅁ
- ⑤ ㄴ, ㄷ, ㄹ

7. OAIS 참조모델의 보존기술정보 유형으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. 맥락정보

ㄴ. 참조정보

ㄷ. 출처정보

ㄹ. 행위주체자정보

ㅁ. 관리메타데이터정보

ㅂ. 기술메타데이터정보

- ① ㄱ, ㄴ, ㄷ
- ② ㄱ, ㄷ, ㅁ
- ③ ㄱ, ㄹ, ㅁ
- ④ ㄴ, ㄷ, ㄹ, ㅂ
- ⑤ ㄴ, ㄷ, ㅁ, ㅂ

8. 한국과학기술정보연구원(KISTI)에서 운영하는 연구데이터 공유·관리 플랫폼은?

- ① DataON
- ② dCollection
- ③ OAK
- ④ RISS
- ⑤ ScienceON

9. 섀넌(Shannon)의 정보이론에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. 정보량과 엔트로피는 반비례한다.

ㄴ. 정보의 단위를 비트(bit)로 표기한다.

ㄷ. 전달되는 정보를 선택한 이후 불확실성은 강화된다.

ㄹ. 정보량이 많다는 것은 무작위성과 불확실성이 크다는 것을 의미한다.

ㅁ. 정보의 의미론적 측면을 배제하고 효율적인 정보 전송에 관한 문제에 집중한다.

- ① ㄱ, ㄴ, ㄷ
- ② ㄱ, ㄹ, ㅁ
- ③ ㄴ, ㄷ, ㄹ
- ④ ㄴ, ㄹ, ㅁ
- ⑤ ㄷ, ㄹ, ㅁ

10. 단어빈도와 역문헌빈도에 대한 설명으로 옳지 않은 것은?

- ① 색인어 가중치는 단어빈도와 역문헌빈도의 값을 조합하여 계산한다.
- ② 역문헌빈도는 문헌빈도가 낮은 단어일수록 높은 가중치가 부여된다.
- ③ 단어빈도는 문헌집단의 크기, 텍스트 길이 등을 고려하여 정규화한 빈도를 사용하기도 한다.
- ④ 단어빈도는 발생빈도가 높은 특정 단어가 해당 문헌의 주제를 잘 나타내는 것으로 판단하는 근거가 된다.
- ⑤ A라는 용어가 전체 문헌 5개 중에서 4개의 문헌에 출현할 때 계산식의 로그를 적용하지 않을 경우 A의 역문헌빈도는 0.8이다.

11. 자연어 색인과 통제어 색인에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. 통제어 색인은 용어 사용의 모호성을 해결할 수 있다.

ㄴ. 통제어 색인은 상위어, 하위어 등 관련 개념에 대한 검색이 용이하다.

ㄷ. 자연어 색인은 관련 개념에 대한 검색이 누락되는 것을 최소화하기 쉽다.

ㄹ. 자연어 색인은 소셜 태깅(social tagging)과 같이 새로운 개념에 대한 표현이 용이하다.

- ① ㄱ, ㄴ
- ② ㄴ, ㄹ
- ③ ㄷ, ㄹ
- ④ ㄱ, ㄴ, ㄹ
- ⑤ ㄱ, ㄴ, ㄷ, ㄹ

12. 다음의 계량서지법칙에 대한 설명으로 옳지 않은 것은?

n편의 논문을 생산한 저자의 수는 한 편의 논문을 생산한 저자의 수의 약 $1/n^2$ 이다. 그리고 한 편의 논문을 생산한 저자의 수는 논문을 생산한 모든 저자의 수의 약 60%이다.

- ① 로트카(Lotka)의 역제곱 법칙으로도 불린다.
- ② 1926년 로트카가 화학자와 물리학자의 논문 건수를 조사하여 분석하였다.
- ③ 다양한 학문 분야에서 저자의 생산성을 분석하고 예측하는 데 활용되고 있다.
- ④ 학문 분야나 연구 방식에 따라 위 법칙의 적용 가능성에 대해 다양한 의견이 존재한다.
- ⑤ 위 법칙에 따르면 전체 논문 생산 저자의 수가 2,000명인 경우 2편의 논문을 생산한 저자의 수는 약 500명이다.

13. OAI-PMH(Open Archives Initiative Protocol for Metadata Harvesting)에 대한 설명으로 옳지 않은 것은?

- ① SP(Service Provider)와 DP(Data Provider) 사이에서 작동한다.
- ② MARC가 아닌 더블린 코어를 채택하고 있어 상대적으로 저비용이다.
- ③ 메타데이터를 수확하지 않는 분산검색 방식에 비해 최신의 정보를 검색할 수 있다.
- ④ HTTP 프로토콜을 이용한 요청과 XML 레코드의 응답이라는 단순한 형식으로 커뮤니케이션한다.
- ⑤ 통합검색을 위해 수확검색 방식을 사용하고 있으며, 기술 장벽이 낮고 저장소 간 정보를 쉽게 공유할 수 있다.

14. 계량서지 지표에 대한 설명으로 옳지 않은 것은?

- ① CNCI(Category Normalized Citation Impact)는 실제 인용 횟수를 동일한 문서 유형, 출판 연도, 주제 영역에 속하는 문서의 예상 인용률로 나누어 정규화한 지수이다.
- ② Eigenfactor는 하나의 학술지가 다른 학술지에 비해 상대적으로 어느 정도의 영향력을 가지는지 평가하는 데 사용된다.
- ③ JIF(Journal Impact Factor)는 일정 기간 특정 학술지에 게재 된 논문의 평균 피인용 횟수를 의미한다.
- ④ SJR(SCImago Journal Rankings)은 SCOPUS 내 정보를 활용한 학술지 영향력 지수이다.
- ⑤ SNIP(Source Normalized Impact per Paper)는 학술 논문이 소셜 미디어, 언론 보도 등에 노출된 정도를 기반으로 계산되는 사회적 영향력 지수이다.

15. 다음은 단어빈도를 제시한 문헌과 용어의 행렬이다. 주어진 불리언 질의에 따라 검색되는 문헌을 모두 고르면?

문헌 \ 용어	data	media	literacy	reading
d1	3	4	2	4
d2	2	2	1	0
d3	1	0	3	0
d4	0	1	0	2

불리언 질의

(data AND media) OR (literacy AND (NOT reading))

- ① d1, d2
- ② d2, d3
- ③ d3, d4
- ④ d1, d2, d3
- ⑤ d1, d2, d4

16. 정보 탐색전략에 대한 설명으로 옳지 않은 것은?

- ① 블록 만들기 방법을 사용하여 복잡한 탐색 과제를 단순하게 분해할 수 있다.
- ② 특정 패킷 우선탐색 방법은 다수의 개념을 포함하는 탐색에 적용되므로 초보자에게 유용하다.
- ③ 눈덩이 굴리기 방법은 탐색 결과가 양적으로 증가하는 것을 의미하며 재현율 향상을 목적으로 한다.
- ④ 점진 분할 방법은 넓은 개념으로 탐색을 시작한 후에 다양한 제한 기법을 사용하여 탐색 범위를 좁혀가는 방법이다.
- ⑤ 신속/편의 지향법은 탐색기능을 사용하지 않고 탐색하려는 용어 하나를 입력하거나 스페이스로 구분한 몇 개의 용어를 한꺼번에 입력하여 검색하는 방법이다.

17. 벡터공간모델에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. 문헌과 질의에 나타난 용어를 벡터의 차원으로 표현한다.

ㄴ. 불리언모델이 가지는 이진 가중치에 대한 대안으로 나온 검색모델이다.

ㄷ. 셸튼(Salton)이 대수론(algebra theory)을 이론적 근거로 하여 만든 모델이다.

ㄹ. 벡터 간 사이각의 코사인 값이 클수록 벡터 간 유사도가 크다는 것을 나타낸다.

ㅁ. 문헌의 색인어와 질의어가 각각 가중치를 가지며, 질의와 부분적으로 일치하는 문헌을 검색할 수 있다.

- ① ㄱ, ㄴ, ㄹ
- ② ㄱ, ㄹ, ㅁ
- ③ ㄴ, ㄷ, ㄹ
- ④ ㄱ, ㄴ, ㄷ, ㅁ
- ⑤ ㄱ, ㄴ, ㄷ, ㄹ, ㅁ

18. 기계학습 방법에 대한 설명으로 옳지 않은 것은?

- ① 비지도학습은 입력 값만을 활용하여 학습하는 방식이다.
- ② 지도학습은 입력 값과 출력 값을 함께 활용하여 학습하는 방식이다.
- ③ 딥러닝은 인공신경망 구조를 활용하여 데이터 표현을 학습한다.
- ④ 클러스터링은 동질적인 데이터를 하나의 그룹으로 묶는 대표적인 지도학습 방법이다.
- ⑤ 강화학습은 현재의 상태에서 누적 보상이 최대가 되는 행동을 하도록 하는 학습 방법이다.

19. 정보학의 이론과 기술 발전에 영향을 준 학자의 주요 업적에 대한 설명으로 옳지 않은 것은?

- ① 토브(Taube)는 유니탐에 불리언 논리를 결합하여 정보검색에 적용하는 기초를 마련하였다.
- ② 스팍 존스(Spärck Jones)는 단어의 가중치 방법으로 역문헌빈도를 이용하는 방식을 제안하였다.
- ③ 스몰(Small)은 텍스트에 출현한 단어의 통계적 특성에 의해 색인어를 선정하는 방식을 처음 제시하였다.
- ④ 셸튼(Salton)은 SMART Information Retrieval System을 이용하여 벡터공간모델, 적합성 피드백 등을 실험하였다.
- ⑤ 가필드(Garfield)는 인용색인 기반의 학술논문 정보서비스 기관인 ISI(Institute for Scientific Information)를 설립하였으며, 계량 서지학과 계량과학의 창시자 중 한 사람이다.

20. 다음은 동일한 문헌 집단을 대상으로 질의 Q를 사용하여 A 검색 시스템과 B 검색 시스템의 검색 성능을 평가한 결과이다. 질의 Q에 대한 문헌 집단 내의 총 적합 문헌 수가 20건이라고 할 때, ㉠과 ㉡의 값으로 옳은 것은?

구분	검색 결과	재현율	정확률
A 검색 시스템	15	(㉠)%	60%
B 검색 시스템	32	40%	(㉡)%

- ㉠

㉡
- ① 35 35
- ② 45 25
- ③ 45 35
- ④ 55 25
- ⑤ 55 35