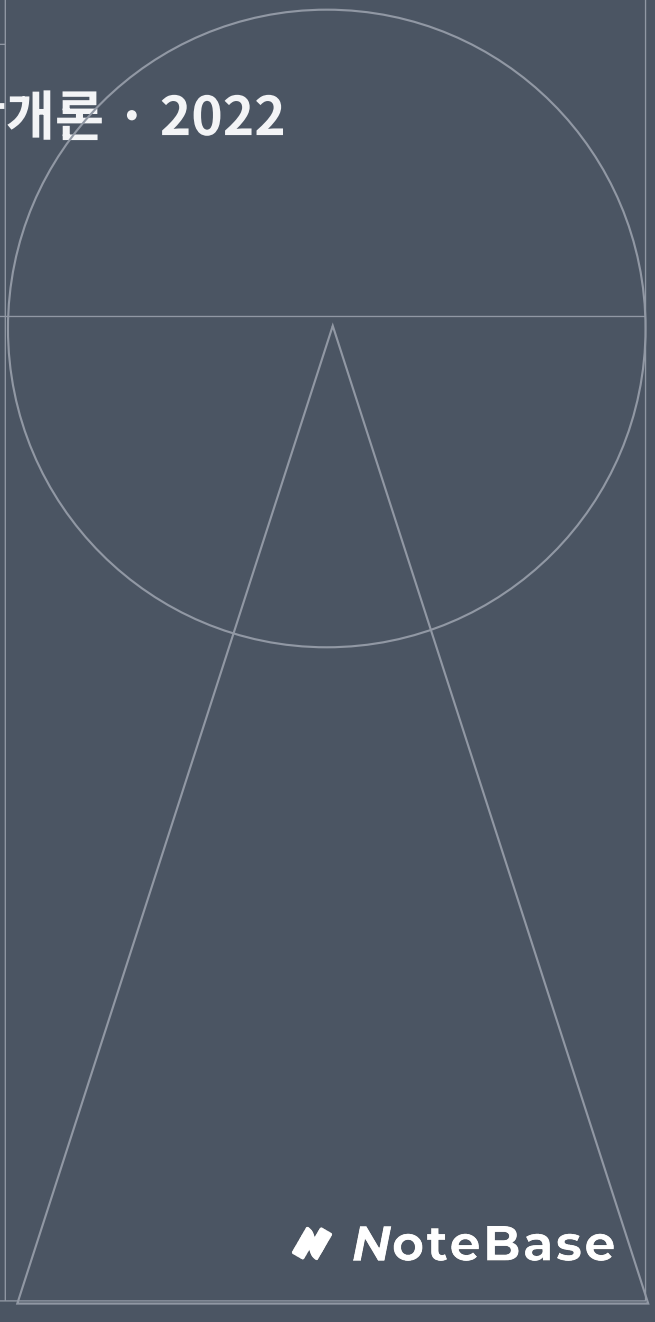
기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2022



NoteBase

기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2022

공무원/국회직/9급 · 정보학개론 · 2022

1. **정보서비스와 운영 기관의 연결이 옳지 않은 것은?**

   ① ScienceON - 한국과학기술정보연구원
   ② RISS - 한국교육학술정보원
   ③ KCI - 한국연구재단
   ④ OAK - 국립중앙도서관
   ⑤ DataON - 국회도서관

정답 5번

출제 포인트

국내 학술·과학기술 정보서비스의 기능과 운영 기관을 정확히 연결할 수 있는지를 묻는 문제이다.

핵심 개념

ScienceON은 과학기술정보와 연구 인프라를 연계하는 서비스이며 한국과학기술정보연구원(KISTI)이 운영한다.
RISS는 학술연구정보서비스이며 한국교육학술정보원(KERIS)이 운영한다.
KCI는 국내 학술지와 논문의 인용 정보를 제공하는 한국학술지인용색인이며 한국연구재단(NRF)이 운영한다.
OAK는 오픈액세스 지식정보의 수집·공유를 지원하는 국가리포지터리이며 국립중앙도서관이 운영한다.
DataON은 연구데이터의 수집·공유·활용을 지원하는 국가연구데이터플랫폼이며 한국과학기술정보연구원(KISTI)이 운영한다.

선지별 해설

① O — ScienceON은 한국과학기술정보연구원(KISTI)이 운영하는 과학기술 지식인프라 서비스이므로 연결이 옳다.
② O — RISS는 한국교육학술정보원(KERIS)이 제공하는 학술연구정보서비스이므로 연결이 옳다.
③ O — KCI는 한국연구재단(NRF)이 운영하는 한국학술지인용색인이므로 연결이 옳다.
④ O — OAK 국가리포지터리는 국립중앙도서관이 운영하므로 연결이 옳다.
⑤ X — 틀린 표현은 '국회도서관'이다. 올바르게는 'DataON - 한국과학기술정보연구원(KISTI)'이다.

NoteBase

오답 포인트

DataON을 국가 차원의 정보서비스라는 이유만으로 국회도서관과 연결하기 쉽다. 그러나 DataON의 핵심어는 ‘연구데이터’이고, 연구데이터 플랫폼을 운영하는 기관은 한국과학기술정보연구원이다.

다시 나오면 이렇게 판단

서비스명에서 핵심 기능을 먼저 찾고 운영 기관과 연결한다. 과학기술정보·연구데이터는 KISTI, 교육·학술연구정보는 KERIS, 학술지 인용색인은 NRF, 오픈액세스 국가리포지터리는 국립중앙도서관으로 구분한다.

암기 포인트

ScienceON·DataON은 KISTI, RISS는 KERIS, KCI는 NRF, OAK는 국립중앙도서관이다.

NoteBase

2. 인용색인에 대한 설명으로 옳지 않은 것은?

① 자기인용은 색인시스템에 잡음을 초래한다.
② 도서관에서 정기간행물 선정과 폐기를 결정하는 데 사용되기도 한다.
③ 인용색인을 이용하여 이용자들에게 SDI와 같은 맞춤형 서비스를 제공할 수 있다.
④ 웹 문서 내의 링크 역시 관련된 다른 웹 문서로 연결시켜주는 인용링크이며, 이 인용링크를 일종의 색인어로 사용할 수 있다.
⑤ 대표적인 인용색인 서비스인 Science Direct는 SCIE, SSCI, A&HCI를 포함한다.

정답 5번

출제 포인트

인용색인의 기능과 활용을 이해하고, 인용색인 데이터베이스와 원문 제공 플랫폼을 구별하는 문제이다.

핵심 개념

인용색인은 문헌이 다른 문헌을 인용한 관계를 색인하여 선행·후속 연구를 추적하게 하는 도구이다. ScienceDirect는 학술지 논문 등의 원문을 검색·열람하는 플랫폼이고, SCIE·SSCI·A&HCI는 Web of Science의 인용색인이다.

선지별 해설

① O — 자기인용은 저자가 자신의 이전 연구를 인용하는 것으로, 외부 연구자가 부여한 영향력과 무관하게 인용 횟수를 늘릴 수 있어 인용 분석에서 잡음이나 왜곡 요인이 된다.
② O — 인용 빈도와 이용자 수요를 함께 살피면 학술지의 상대적 활용도와 영향력을 판단할 수 있으므로 도서관의 정기간행물 선정 및 폐기 결정에 활용할 수 있다.
③ O — 이용자가 관심을 보인 핵심 문헌을 인용한 후속 문헌을 추적하면 관심 주제의 최신 문헌을 골라 제공할 수 있으므로 SDI와 같은 맞춤형 정보서비스에 활용할 수 있다.
④ O — 웹 문서의 링크는 한 문서가 다른 문서를 가리킨다는 점에서 문헌의 인용과 유사하다. 따라서 링크 관계나 링크에 쓰인 표현을 분석하여 웹 자원의 연결·검색을 위한 색인 정보로 활용할 수 있다.
⑤ X — 틀린 부분은 '대표적인 인용색인 서비스인 ScienceDirect는 SCIE, SSCI, A&HCI를 포함한다'이다. ScienceDirect는 Elsevier가 제공하는 학술 원문 검색·열람 플랫폼이며, SCIE·SSCI·A&HCI를 포함하는 대표적 인용색인 서비스는 Web of Science이다.

오답 포인트

ScienceDirect에서도 학술논문을 검색할 수 있다는 이유만으로 이를 인용색인 데이터베이스로 오인하기 쉽다. ⑤는 '원문을 제공하는 플랫폼'과 '문헌 간 인용 관계를 색인하는 서비스'를 바꾸어 놓은 선지이다.

NoteBase

다시 나오면 이렇게 판단

서비스 이름을 보면 먼저 주된 기능을 구분한다. 논문 원문 검색·열람이 중심이면 ScienceDirect이고, 인용 관계와 SCIE·SSCI·A&HCI가 핵심이면 Web of Science이다.

암기 포인트

ScienceDirect는 원문, Web of Science는 인용색인이다.

NoteBase

3. 특정 자료에서 일반적인 자료로 접근하기 위한 서지탐색 전략에 대한 설명으로 <보기>에서 옳은 것만을 모두 고르면?

<보 기>

ㄱ. 인용문헌 확대 탐색 전략(citation pearl growing)
ㄴ. 특정 패싯 우선 전략(most specific facet first)
ㄷ. 인용색인 활용(citation indexing)
ㄹ. 시소러스 훑어보기(thesaurus browsing)
ㅁ. 점진 분할 전략(successive fraction)

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

정답 1번

출제 포인트

특정한 문헌이나 좁은 검색 집합에서 출발하여 인용 관계 또는 탐색 패싯을 따라 더 넓고 일반적인 자료로 확장하는 서지탐색 전략을 구별하는 문제이다.

핵심 개념

인용문헌 확대 탐색 전략(citation pearl growing)은 이미 찾은 적합 문헌을 씨앗 문헌으로 삼아 참고문헌과 인용 문헌을 연쇄적으로 추적하며 관련 자료를 넓혀 가는 방법이다.
특정 패싯 우선 전략(most specific facet first)은 검색 패싯 가운데 가장 구체적이고 제한적인 패싯부터 검색한 뒤 다른 패싯을 조합하여 탐색 범위를 조정하는 방법이다.
인용색인 활용(citation indexing)은 특정 문헌의 인용·피인용 관계를 이용하여 그 문헌과 연결된 자료를 추적하는 방법이다.
시소러스 훑어보기(thesaurus browsing)는 통제어의 상위어·하위어·관련어 관계를 살펴 검색어를 찾는 어휘 탐색 방식이며, 특정 문헌에서 출발하는 확대 전략은 아니다.
점진 분할 전략(successive fraction)은 넓은 검색 집합에 조건을 차례로 추가하여 결과를 점점 좁히는 방식이므로 이 문제에서 묻는 방향과 반대이다. 따라서 옳은 것은 ㄱ, ㄴ, ㄷ이다.

NoteBase

선지별 해설

① O — ㄱ은 씨앗 문헌에서 인용 관계를 따라 확장하고, ㄴ은 가장 구체적인 패싯에서 탐색을 시작하며, ㄷ은 특정 문헌의 인용 관계로 관련 자료를 추적하므로 세 항목이 모두 해당한다.
② X — ㄹ을 포함하고 ㄴ을 제외한 구성이 틀렸다. 시소러스 훑어보기는 통제어 관계를 탐색하는 방식이고, 특정 패싯 우선 전략인 ㄴ이 해당한다. 올바른 조합은 ㄱ, ㄴ, ㄷ이다.
③ X — ㅁ을 포함하고 ㄴ을 제외한 구성이 틀렸다. 점진 분할 전략은 넓은 검색 결과를 조건 추가로 좁혀 가는 방식이고, 이 문항에는 특정 패싯 우선 전략인 ㄴ이 포함되어야 한다. 올바른 조합은 ㄱ, ㄴ, ㄷ이다.
④ X — ㄹ과 ㅁ을 포함하고 ㄱ과 ㄷ을 제외한 구성이 틀렸다. ㄹ은 어휘 관계 탐색이고 ㅁ은 넓은 집합을 좁히는 전략인 반면, 인용 관계를 이용해 특정 문헌에서 확장하는 ㄱ과 ㄷ이 해당한다. 올바른 조합은 ㄱ, ㄴ, ㄷ이다.
⑤ X — ㄹ과 ㅁ을 포함하고 ㄱ과 ㄴ을 제외한 구성이 틀렸다. ㄹ과 ㅁ은 이 문항의 특정 자료에서 일반 자료로 확장하는 전략에 해당하지 않으며, ㄱ과 ㄴ이 포함되어야 한다. 올바른 조합은 ㄱ, ㄴ, ㄷ이다.

오답 포인트

ㄹ의 상위어 탐색만 보고 '구체적 용어에서 일반적 용어로 이동한다'고 판단하기 쉽다. 그러나 시소러스 훑어보기의 출발점은 특정 문헌이 아니라 통제어 체계이다. 또한 ㅁ은 조건을 누적하여 넓은 결과를 좁히는 전략이므로 탐색 방향을 반대로 기억하면 틀리기 쉽다.

다시 나오면 이렇게 판단

출발점과 탐색 방향을 함께 확인한다. 이미 확보한 특정 문헌의 인용 관계를 따라가는 전략은 확대형이고, 가장 구체적인 패싯부터 시작하는 전략도 특정한 것에서 출발한다. 반대로 넓은 집합에 조건을 추가하는 점진 분할은 축소형이며, 시소러스 훑어보기는 문헌 출발 전략이 아니라 어휘 관계 탐색이다.

암기 포인트

특정 문헌에서 확장: 펄 그로잉·인용색인, 구체 패싯에서 시작: 특정 패싯 우선, 넓은 집합을 축소: 점진 분할이다.

NoteBase

4. 메타데이터 스키마의 세 가지 요소만을 모두 고르면?

<보 기>

ㄱ. 구문(syntax)
ㄴ. 자원(resource)
ㄷ. 내용규칙(content rules)
ㄹ. 의미구조(semantics)
ㅁ. 속성값(value)

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

정답 3번

출제 포인트

메타데이터 스키마를 구성하는 세 요소인 구문, 의미구조, 내용규칙을 실제 기술 대상인 자원 및 그 자원에 입력되는 속성값과 구별하는 문제이다.

핵심 개념

구문(syntax): 메타데이터 요소와 값을 어떤 구조와 형식으로 표현할지를 정한다.
의미구조(semantics): 각 메타데이터 요소가 무엇을 뜻하는지와 요소 사이의 의미 관계를 정한다.
내용규칙(content rules): 각 요소에 값을 어떤 원칙과 형식에 따라 선정·기입할지를 정한다.
따라서 메타데이터 스키마의 세 요소는 ㄱ 구문, ㄷ 내용규칙, ㄹ 의미구조이다. ㄴ 자원은 메타데이터가 기술하는 대상이고, ㅁ 속성값은 요소에 실제로 입력되는 값이므로 스키마의 세 구성 요소에 해당하지 않는다.

NoteBase

선지별 해설

① X — ㄴ 자원(resource)은 메타데이터가 기술하는 대상이지 스키마의 구성 요소가 아니다. ‘ㄱ, ㄴ, ㄷ’은 ‘ㄱ, ㄷ, ㄹ’로 고쳐야 한다.
② X — ㄴ 자원(resource)이 잘못 포함되고 ㄷ 내용규칙(content rules)이 빠졌다. ‘ㄱ, ㄴ, ㄹ’은 ‘ㄱ, ㄷ, ㄹ’로 고쳐야 한다.
③ O — ㄱ 구문(syntax), ㄷ 내용규칙(content rules), ㄹ 의미구조(semantics)는 메타데이터 스키마의 세 요소를 모두 정확히 제시한다.
④ X — ㄴ 자원(resource)과 ㅁ 속성값(value)은 스키마가 적용되어 기술·기입되는 대상이지 스키마의 세 요소가 아니다. ‘ㄴ, ㄷ, ㅁ’은 ‘ㄱ, ㄷ, ㄹ’로 고쳐야 한다.
⑤ X — ㅁ 속성값(value)은 개별 메타데이터 요소에 실제로 입력되는 값이며 스키마의 구성 요소가 아니다. ‘ㄷ, ㄹ, ㅁ’은 ‘ㄱ, ㄷ, ㄹ’로 고쳐야 한다.

오답 포인트

‘자원’과 ‘속성값’도 메타데이터를 설명할 때 자주 쓰이는 말이어서 스키마 자체의 요소처럼 보이는 것이 함정이다. 자원은 기술 대상이고 속성값은 실제 기입 내용이며, 스키마는 이를 표현하고 해석하고 작성하는 규칙의 틀이다.

다시 나오면 이렇게 판단

후보가 스키마라는 규칙의 틀을 정하는 항목인지, 그 틀이 적용되는 대상이나 실제 입력값인지를 먼저 구분한다. 표현 형식은 구문, 요소의 뜻은 의미구조, 값의 작성 원칙은 내용규칙으로 판단한다.

암기 포인트

메타데이터 스키마는 ‘어떻게 표현하는가(구문)－무엇을 뜻하는가(의미구조)－어떻게 값을 쓰는가(내용규칙)’의 세 축이다.

NoteBase

5. 지프(Zipf)의 제1법칙을 적용할 경우, 문헌에서 나타나는 첫 번째 순위의 단어출현빈도가 400일 때 다섯 번째 순위의 단어출현빈도는?

① 200 ② 133 ③ 100 ④ 80 ⑤ 66

정답 4번

출제 포인트

Zipf의 제1법칙에서 단어의 출현빈도와 순위가 반비례한다는 원리를 이용해 다섯 번째 순위의 출현빈도를 구하는 문제이다.

조건 정리

주어진 값: 첫 번째 순위의 단어출현빈도는 400이다.
구해야 할 값: 다섯 번째 순위의 단어출현빈도이다.
주의할 조건: 순위가 커질수록 빈도는 그 순위에 반비례하여 감소한다.

사용할 공식

Zipf의 제1법칙은 $f_r = \frac{f_1}{r}$이다. 여기서 f_1은 첫 번째 순위의 빈도, r은 순위, f_r은 r번째 순위의 빈도이다. 단어의 출현빈도가 순위에 반비례하므로 이 공식을 사용한다.

계산 과정

1. 첫 번째 순위의 빈도 $f_1 = 400$, 구하려는 순위 $r = 5$를 확인한다.
2. 공식에 대입하면 $f_5 = \frac{400}{5}$이다.
3. 따라서 $f_5 = 80$이다.

선지 판단

① X — 200은 400 ÷ 2로 구한 두 번째 순위의 빈도이다. 다섯 번째 순위는 200이 아니라 80이다.
② X — 133은 400 ÷ 3을 정수로 나타낸 값에 가깝고 세 번째 순위에 해당한다. 다섯 번째 순위는 133이 아니라 80이다.
③ X — 100은 400 ÷ 4로 구한 네 번째 순위의 빈도이다. 다섯 번째 순위는 100이 아니라 80이다.
④ O — 400 ÷ 5 = 80이므로 다섯 번째 순위의 빈도와 일치한다.
⑤ X — 66은 400 ÷ 6을 정수로 나타낸 값에 가깝고 여섯 번째 순위에 해당한다. 다섯 번째 순위는 66이 아니라 80이다.
계산 결과가 80이므로 정답은 4번이다.

NoteBase

오답 포인트

선지의 200, 133, 100, 66은 각각 다른 순위로 나누었을 때 나오는 값이다. 첫 번째 순위의 빈도 400을 목표 순위인 5로 나누어야 한다.

다시 나오면 이렇게 판단

첫 순위의 빈도와 목표 순위를 확인한 뒤 $f_r = f_1/r$에 대입하고, 빈도와 순위의 곱 rf_r이 첫 순위의 빈도와 같은지 검산한다.

암기 포인트

Zipf의 제1법칙은 '순위가 r배이면 빈도는 $1/r$배'이다.

NoteBase

6. 통제어휘 검색에 대한 설명으로 옳지 않은 것은?

① 통제어휘집을 참고하여 동일한 개념을 가진 용어들을 하나의 색인어로 검색할 수 있다.
② 통제어휘 시스템이 주제별로 구축되어 있는 예시로 ERIC 시소러스, MeSH 등이 있다.
③ 찾고자 하는 탐색어가 비디스크립터인 경우에는 디스크립터가 출현한 문서도 검색될 수 있게 해준다.
④ 통제어휘의 사용은 색인어를 부여하는 데 일관성을 갖게 한다.
⑤ 통제어휘의 사용으로 색인어의 특정성과 융통성이 높아진다.

정답 5번

출제 포인트

통제어휘 검색의 기능과 장점을 자연어 검색의 특성과 구별하는 문제이다.

핵심 개념

통제어휘는 시소러스처럼 미리 정한 표준 용어와 용어 간 관계를 이용하여 문헌을 색인하고 검색하는 어휘 체계이다. 동의어를 대표어로 모으고 상위어·하위어·관련어를 연결하므로 검색의 일관성과 특정성이 높아진다. 반면 검색자가 자유롭게 표현을 바꾸거나 새 용어를 즉시 사용할 수 있는 융통성은 자연어 검색보다 낮다.

선지별 해설

① O — 시소러스는 동의어나 유사어를 우선어인 디스크립터로 통일한다. 따라서 표현이 달라도 동일한 개념의 문헌을 하나의 표준 색인어를 중심으로 검색할 수 있다.
② O — 통제어휘 시스템은 특정 주제 분야의 개념과 용어 관계를 체계화하여 구축된다. ERIC 시소러스는 교육 분야, MeSH는 의학·생명과학 분야의 대표적인 통제어휘이다.
③ O — 시소러스의 계층관계는 상위어와 하위어를 연결한다. 탐색어가 비디스크립터이면 우선어로 안내하고, 필요한 경우 해당 디스크립터와 그 하위 개념까지 확장하여 검색할 수 있으므로 관련 문헌의 누락을 줄일 수 있다.
④ O — 동일한 개념에 동일한 우선어를 부여하므로 색인 담당자나 문헌이 달라도 용어 사용의 일관성을 확보할 수 있다.
⑤ X — 틀린 표현은 '특정성과 융통성이 높아진다'이다. 통제어휘는 의미가 명확한 표준어를 사용하므로 검색의 특정성과 일관성은 높아지지만, 검색어를 자유롭게 선택하는 융통성은 낮아진다. 올바르게는 '통제어휘의 사용으로 색인어의 특정성과 일관성은 높아지지만 융통성은 낮아진다'이다.

오답 포인트

⑤는 통제어휘의 장점을 모두 높아지는 성질처럼 묶은 것이 함정이다. 통제어휘는 정해진 표준어를 쓰기 때문에 의미 통제에는 강하지만, 자연어처럼 새 표현이나 다양한 표현을 즉시 수용하는 데에는 제약이 있다.

NoteBase

다시 나오면 이렇게 판단

‘동의어 통제, 용어 관계, 일관성, 특정성’은 통제어휘의 장점으로 판단한다. ‘자유로운 표현, 신조어의 즉시 반영, 융통성’은 자연어 검색의 상대적 장점으로 구별한다.

암기 포인트

통제어휘는 특정성·일관성은 높이고, 표현의 융통성은 낮춘다.

NoteBase

7. 다음에서 설명하는 정보검색 모델과 가장 거리가 먼 것은?

> 정보콘텐츠의 내용 표현과 검색질의의 표현 간의 유사도를 평가하여 매칭여부를 결정하고 검색결과를 도출하는 모델

① 확률 모델
② 벡터공간 모델
③ 불리안 모델
④ 연계검색 모델
⑤ 이진독립검색 모델

정답 4번

출제 포인트

정보콘텐츠의 내용 표현과 검색질의의 표현을 비교하여 검색 결과를 만드는 내용 기반 검색 모델과, 문서 사이의 연결 관계를 이용하는 연계검색 모델을 구별하는 문제이다.

핵심 개념

내용 기반 검색 모델은 문서와 질의를 색인어 등의 표현으로 나타내고 두 표현의 일치 정도나 유사도, 관련 확률을 이용하여 검색한다. 벡터공간 모델은 문서와 질의를 벡터로 표현하여 유사도를 계산하는 대표적인 모델이다.
반면 연계검색 모델은 문서 내용과 질의 표현의 유사도보다 문서 사이의 링크·인용 등 연결 관계를 중심으로 검색하므로 제시된 설명과 가장 거리가 멀다.

NoteBase

선지별 해설

① X — 확률 모델은 문서가 질의에 관련될 확률을 추정하여 검색 결과를 제시하므로 문서 표현과 질의 표현을 비교하는 내용 기반 검색의 범위에 속한다. '가장 거리가 먼 모델'이라는 판단은 틀리며, 올바르게는 관련 확률을 기준으로 문서와 질의의 관련성을 판단하는 모델이다.
② X — 벡터공간 모델은 문서와 질의를 벡터로 표현하고 두 벡터의 유사도를 계산하여 순위를 정하므로 제시문에 가장 직접적으로 부합한다. '가장 거리가 먼 모델'이라는 판단은 틀리며, 올바르게는 제시문의 대표적인 모델이다.
③ X — 불리언 모델은 색인어와 AND·OR·NOT 같은 논리 연산자로 표현된 질의를 문서의 색인어 표현과 대조하여 일치 여부를 결정한다. '가장 거리가 먼 모델'이라는 판단은 틀리며, 올바르게는 유사도 순위화보다는 논리 조건의 일치 여부로 검색하는 내용 기반 모델이다.
④ O — 연계검색 모델은 문서 내용과 질의 표현의 유사도보다 문서 사이의 링크나 인용 같은 연결 구조를 활용한다. 따라서 제시된 내용 기반 유사도·일치도 검색과 가장 거리가 멀다.
⑤ X — 이진독립검색 모델은 문서에서 검색어가 나타나는지 여부를 이진적으로 표현하고, 검색어들의 독립을 가정하여 질의에 대한 관련 확률을 산출하는 확률적 검색 모델이다. '가장 거리가 먼 모델'이라는 판단은 틀리며, 올바르게는 문서와 질의의 용어 정보를 이용해 관련성을 판단하는 모델이다.

오답 포인트

제시문에 '유사도'가 있으므로 ②만 관련 있고 나머지는 모두 무관하다고 생각하기 쉽다. 그러나 질문은 제시문과 '가장 거리가 먼 것'을 묻는다. ①·③·⑤도 방식은 서로 다르지만 문서의 내용 표현과 질의 표현을 대조하며, ④만 문서 사이의 연결 구조를 핵심 정보로 삼는다.

다시 나오면 이렇게 판단

모델이 무엇을 비교하는지 먼저 확인한다. 문서 내용과 질의의 용어 표현을 비교하면 내용 기반 검색이고, 문서와 문서 사이의 링크·인용 관계를 이용하면 연계검색이다. '가장 거리가 먼 것'에서는 비교 대상 자체가 다른 모델을 고른다.

암기 포인트

내용↔질의는 내용 기반 검색, 문서↔문서의 링크·인용은 연계검색이다.

NoteBase

8. 인공지능(Artificial Intelligence)에 대한 설명으로 옳지 않은 것은?

① 인공지능은 인간의 지능을 모방하여 학습한 기계가 인간처럼 추론하고 문제를 해결하는 것을 목적으로 한다.

② 인공지능은 기계학습과 딥러닝을 포괄하는 상위개념이며, 데이터의 품질은 인공지능의 성능에 영향을 미친다.

③ 기계학습은 주로 지도학습, 비지도학습 및 강화학습으로 구분되며 모든 학습에는 결과값에 대한 보상을 주는 방식이 적용된다.

④ 딥러닝은 인간의 뇌신경망을 모방한 인공신경망의 한 종류로, 현재는 음성·이미지 인식 등의 영역에서 높은 성능을 보인다.

⑤ 국내외 도서관에서는 인공지능 기술에 기반한 챗봇으로 참고면담을 진행하는 정보서비스를 제공하고 있다.

정답 3번

출제 포인트

인공지능·기계학습·딥러닝의 포함 관계와 기계학습의 주요 학습 유형을 구별하는 문제이다.

핵심 개념

인공지능은 인간의 지능적 활동을 컴퓨터가 수행하도록 하는 넓은 개념이다. 기계학습은 데이터를 통해 규칙과 패턴을 학습하는 인공지능의 한 분야이고, 딥러닝은 여러 층의 인공신경망을 이용하는 기계학습의 한 분야이다. 따라서 포함 관계는 인공지능 > 기계학습 > 딥러닝이다. 기계학습의 대표 유형은 정답이 표시된 데이터로 배우는 지도학습, 정답 표시 없이 구조를 찾는 비지도학습, 행동 결과의 보상을 바탕으로 배우는 강화학습이다.

선지별 해설

① O — 인공지능은 인간의 지능적 활동을 모방하여 추론·학습·문제 해결 등을 수행하는 기계를 구현하는 것을 목적으로 한다.

② O — 인공지능은 기계학습을 포함하고, 기계학습은 딥러닝을 포함하는 상위 개념이다. 학습에 쓰이는 데이터의 품질이 낮으면 잘못된 패턴을 학습할 수 있으므로 인공지능의 성능에도 영향을 미친다.

③ X — 틀린 표현은 '모든 학습에는 결과값에 대한 보상을 주는 방식이 적용된다'이다. 보상을 이용하는 방식은 강화학습의 특징이며, 지도학습은 정답이 표시된 데이터로 학습하고 비지도학습은 정답 표시 없이 데이터의 구조나 패턴을 찾는다.

④ O — 딥러닝은 다층 인공신경망을 이용하는 기계학습의 한 종류이며, 음성·이미지 인식과 같은 분야에서 높은 성능을 보인다.

⑤ O — 국내외 도서관은 인공지능 기반 챗봇을 활용하여 이용자의 질문에 자동으로 응답하고 참고면담을 지원하는 정보서비스를 제공한다.

NoteBase

오답 포인트

③은 강화학습의 특징인 '보상'을 기계학습의 모든 유형에 공통으로 적용되는 방식처럼 일반화한 것이 함정이다. 지도학습의 정답 데이터, 비지도학습의 패턴 탐색, 강화학습의 보상을 구분해야 한다.

다시 나오면 이렇게 판단

먼저 인공지능 > 기계학습 > 딥러닝의 포함 관계를 확인한 뒤, 학습 방식은 '정답 데이터=지도학습, 정답 표시 없음=비지도학습, 보상=강화학습'으로 대응시킨다. '모든 학습'처럼 한 유형의 특징을 전체에 확대하는 표현이 있는지도 확인한다.

암기 포인트

AI 안에 ML, ML 안에 DL이 있으며, 보상은 모든 학습이 아니라 강화학습의 핵심이다.

NoteBase

9. <보기>에서 통합검색을 위한 프로토콜의 개수는?

<보 기>

○ SUSHI	○ OAI-PMH
○ Z39.50	○ SHERPA
○ LDAP	○ SRU

① 2개
② 3개
③ 4개
④ 5개
⑤ 6개

정답 2번

출제 포인트

제시된 용어 가운데 통합검색에 이용되는 검색·검색결과 회수 프로토콜을 골라 그 개수를 판단하는 문제이다.

핵심 개념

통합검색 프로토콜은 서로 다른 정보원이나 디렉터리에 검색 요청을 보내고 결과를 회수하는 데 쓰이는 통신 규약이다.
Z39.50은 원격 데이터베이스의 검색과 레코드 회수를 위한 클라이언트·서버 프로토콜이고, LDAP은 디렉터리 정보의 접근과 검색을 위한 프로토콜이며, SRU는 URL과 CQL을 이용해 검색·회수 요청을 처리하는 웹 기반 프로토콜이다. 따라서 해당 항목은 Z39.50, LDAP, SRU의 3개이다.
SUSHI는 전자자원 이용통계를 자동 수집하는 프로토콜이고, OAI-PMH는 저장소의 메타데이터를 수확하는 프로토콜이다. SHERPA는 오픈액세스 정책·저장소 정보를 제공하는 서비스 계열의 명칭이므로 통합검색 프로토콜에 해당하지 않는다.

NoteBase

선지별 해설

① X - 통합검색 프로토콜은 2개가 아니라 Z39.50, LDAP, SRU의 3개이다. 틀린 표현 '2개'는 '3개'로 고쳐야 한다.
② O - Z39.50, LDAP, SRU가 해당하므로 모두 3개이다.
③ X - 4개로 세려면 SUSHI, OAI-PMH, SHERPA 중 하나를 잘못 포함하게 된다. 이들은 이 문항의 통합검색 프로토콜이 아니므로 '4개'는 '3개'로 고쳐야 한다.
④ X - 5개가 아니라 3개이다. SUSHI는 이용통계 수집, OAI-PMH는 메타데이터 수확에 쓰이고 SHERPA는 서비스 계열의 명칭이므로 제외한다.
⑤ X - 여섯 항목 전부가 통합검색 프로토콜인 것은 아니다. SUSHI, OAI-PMH, SHERPA를 제외하므로 '6개'는 '3개'로 고쳐야 한다.

오답 포인트

SUSHI와 OAI-PMH도 이름에 프로토콜의 성격이 드러나므로 모두 통합검색용이라고 착각하기 쉽다. 그러나 SUSHI의 대상은 이용통계이고 OAI-PMH의 핵심 동작은 메타데이터 수확이다. '검색·회수'와 '통계·메타데이터 수집'을 구분해야 한다.

다시 나오면 이렇게 판단

각 용어의 처리 대상을 먼저 확인한다. 데이터베이스나 디렉터리에 질의를 보내 결과를 찾고 회수하면 통합검색 쪽으로, 이용통계나 메타데이터를 일괄 수집하면 수확 쪽으로 분류한다.

암기 포인트

검색 3종은 Z39.50·LDAP·SRU, 수집 2종은 SUSHI·OAI-PMH이다. SHERPA는 서비스 명칭이다.

NoteBase

10. 계량서지학에 관한 설명으로 옳지 않은 것은?

① 연구기관들의 연구성과를 비교 평가하는 데 유용하다.
② 로트카(Lotka)는 도서관에서 최소한의 학술지를 선택할 수 있는 경제적인 방법을 제시해준다.
③ 기존 방법론을 보완한 것으로 소셜 미디어에서의 연구 영향력을 측정할 수 있는 알트매트릭스(altmetrics)가 있다.
④ 출판물의 인용데이터를 활용한 대표적인 사례로 인용색인 데이터베이스 분석이 있다.
⑤ SCOPUS에 등재된 모든 학술지의 SJR 평균값은 1이다.

정답 2번

출제 포인트

계량서지학의 주요 법칙과 지표가 무엇을 측정하고 어디에 활용되는지를 구별하는 문제이다.

핵심 개념

계량서지학은 출판물과 인용 등의 수량적 데이터를 분석하여 학술정보의 생산·분포·이용 양상을 연구하는 분야이다.
로트카 법칙은 저자의 논문 생산성 분포를 설명하고, 브래드퍼드 법칙은 특정 주제의 논문이 학술지에 분산되는 양상을 설명하여 핵심 학술지를 가려내는 데 활용된다.
알트메트릭스는 소셜 미디어, 뉴스, 온라인 저장·공유 등에서 나타나는 관심과 확산을 측정하여 전통적 인용지표를 보완한다.

선지별 해설

① O — 논문 수, 인용 수, 영향력 지표 등의 계량 자료를 이용하면 연구기관별 연구성과를 비교·평가하는 데 활용할 수 있다.
② X — 틀린 부분은 '로트카(Lotka)는 도서관에서 최소한의 학술지를 선택할 수 있는 경제적인 방법을 제시한다'이다. 로트카 법칙은 저자의 논문 생산성 분포를 설명한다. 학술지의 분산 정도를 분석하여 핵심 학술지를 선정하는 데 활용되는 것은 브래드퍼드 법칙이다.
③ O — 알트메트릭스는 소셜 미디어 등 온라인 환경에서 나타나는 연구성과의 관심과 확산을 측정하므로, 인용 중심의 기존 영향력 지표를 보완한다.
④ O — 인용색인 데이터베이스 분석은 출판물 사이의 인용 관계를 수량화하여 연구 영향과 지식 구조를 파악하는 대표적인 계량서지학 활용 사례이다.
⑤ O — SJR은 Scopus 자료를 바탕으로 인용한 학술지의 위신까지 가중하여 학술지의 평균 논문당 위신을 나타내며, 전체 평균이 1이 되도록 정규화된 지표이다.

NoteBase

오답 포인트

②는 '학술지 선택'이라는 활용 목적 때문에 로트카 법칙과 브래드퍼드 법칙을 바꾸어 제시한 선지이다. 저자의 생산성을 다루면 로트카, 주제별 학술지의 분산과 핵심 학술지를 다루면 브래드퍼드로 구별한다.

다시 나오면 이렇게 판단

선지의 분석 단위를 먼저 확인한다. 저자별 논문 생산량이면 로트카 법칙, 특정 주제 논문의 학술지별 분산과 핵심 학술지 선정이면 브래드퍼드 법칙, 온라인 관심과 확산이면 알트메트릭스로 판단한다.

암기 포인트

로트카는 저자 생산성, 브래드퍼드는 학술지 분산, 알트메트릭스는 온라인 관심이다.

NoteBase

11. 코로나 관련 정책에 대한 국내외 자료를 검색하기 위한 검색식으로 가장 적합한 것은?

① 코로나 AND 정책
② COVID AND policy
③ 코로나 OR COVID OR 정책 OR policy
④ (코로나 OR COVID) AND (정책 OR policy)
⑤ (코로나 AND COVID) OR (정책 AND policy)

정답 4번

출제 포인트

불리언 검색에서 동의어는 OR로 넓히고, 서로 다른 핵심 개념은 AND로 결합하는 원리를 묻는 문제이다.

핵심 개념

OR 검색은 여러 검색어 가운데 하나 이상이 포함된 자료를 찾으므로 동의어·유사어를 묶는 데 사용한다. AND 검색은 모든 검색어가 함께 포함된 자료를 찾으므로 서로 다른 개념을 결합하는 데 사용한다. 따라서 코로나의 한글·영문 표현은 (코로나 OR COVID), 정책의 한글·영문 표현은 (정책 OR policy)로 묶고, 두 개념군 사이는 AND로 연결한다.

선지별 해설

① X — '코로나 AND 정책'은 코로나와 정책을 모두 포함한 자료를 찾지만 영문 표현인 COVID와 policy를 포함하지 않아 국내외 자료를 폭넓게 검색하기 어렵다. 올바르게는 각 개념의 영문 동의어까지 포함하여 '(코로나 OR COVID) AND (정책 OR policy)'로 검색한다.
② X — 'COVID AND policy'는 영문 표현만 사용하여 한글로 작성된 관련 자료를 놓칠 수 있다. 올바르게는 한글과 영문 동의어를 각각 OR로 묶은 '(코로나 OR COVID) AND (정책 OR policy)'이다.
③ X — 모든 검색어를 OR로 연결하면 코로나 또는 정책 중 어느 한 개념만 포함된 자료도 검색되어 주제와 무관한 결과가 지나치게 많아진다. 올바르게는 동의어끼리는 OR, 코로나 개념군과 정책 개념군 사이는 AND로 연결한다.
④ O — '(코로나 OR COVID)'는 코로나의 한글·영문 표현을 포괄하고, '(정책 OR policy)'는 정책의 한글·영문 표현을 포괄한다. 두 개념군을 AND로 연결하므로 코로나 관련 정책 자료를 국내외 표현에 걸쳐 적절히 검색한다.
⑤ X — '코로나 AND COVID'와 '정책 AND policy'는 동의어가 모두 동시에 등장하도록 지나치게 좁히며, 두 묶음 사이의 OR 때문에 코로나만 다룬 자료나 정책만 다룬 자료도 검색될 수 있다. 올바르게는 동의어끼리 OR로 묶고 두 개념군을 AND로 결합한 '(코로나 OR COVID) AND (정책 OR policy)'이다.

NoteBase

오답 포인트

③은 검색어가 많아 보여 포괄적으로 느껴지지만, 전체를 OR로 연결하면 코로나와 정책이 함께 다뤄진 자료로 한정되지 않는다. ⑤는 괄호를 사용했지만 동의어 사이에 AND를 넣고 개념군 사이에 OR를 넣어 연산자의 역할을 반대로 적용한 함정이다.

다시 나오면 이렇게 판단

먼저 같은 개념의 다른 표현을 찾아 OR로 묶고, 반드시 함께 나타나야 하는 서로 다른 개념군은 AND로 연결한다. 괄호 안은 동의어 확장, 괄호 사이는 주제 결합으로 판단한다.

암기 포인트

동의어는 OR, 서로 다른 핵심 개념은 AND이다.

NoteBase

12. 다음은 RDF/XML 구문을 이용한 레코드의 일부이다. 이에 대한 설명으로 옳지 않은 것은?

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dcterms="http://purl.org/dc/terms/"
    xmlns:ex="http://example.org/taxonomy/"
    xmlns:dcam="http://purl.org/dc/dcam/">

  <rdf:Description rdf:about="http://example.org/123">
      <dcterms:subject>
      <rdf:Description rdf=about="http://example.org/subject32">
      <dcam:memberOf
      rdf:resource="http://example.org/taxonomy/ExampleSubjects"/>
      <rdf:value xml:lang="en">Biology</rdf:value>
      <rdf:value xml:lang="sv">Biologi</rdf:value>
      <rdf:value
      rdf:datatype="http://example.org/taxonomy/SubjectEncoding">
      EA32</rdf:value>
      </rdf:Description>
      </dcterms:subject>
  </rdf:Description>
</rdf:RDF>
```

① RDF는 웹상의 다양한 정보자원을 기술하고 데이터모델을 표현할 수 있도록 한다.

② XML 형태로 작성된 문서는 별다른 변환의 과정 없이 RDF로 직접 이용할 수 있다.

③ XML 네임스페이스를 사용하여 정보자원의 속성을 명시하고 있다.

④ 더블린코어의 메타데이터 요소를 표현해주는 구문규칙 형식을 적용하고 있다.

⑤ 'xml:lang' 속성을 사용하여 정보자원의 내용을 표현해주는 언어를 나타내고 있다.

정답 5번

출제 포인트

RDF/XML에서 네임스페이스, 더블린코어 메타데이터 요소, 언어 태그, 데이터형이 각각 무엇을 나타내는지 구별하는 문제이다.

NoteBase

핵심 개념

RDF는 자원을 주어-술어-목적어 관계로 기술하는 데이터 모델이고, RDF/XML은 그 RDF 데이터를 XML 문법으로 직렬화한 형식이다.
XML 네임스페이스는 rdf, dcterms, dcam처럼 어휘의 URI를 접두어에 연결하여 요소와 속성의 의미를 식별한다.
xml:lang은 그 속성이 붙은 문자열 리터럴의 자연언어를 표시한다. 반면 정보자원 자체의 내용 언어를 기술하려면 dcterms:language 같은 메타데이터 속성을 사용한다.

선지별 해설

① O — RDF는 웹상의 자원을 URI로 식별하고 자원 사이의 관계를 구조화하여 데이터 모델로 표현할 수 있게 한다.
② O — 제시문은 RDF 데이터를 XML 문법으로 직렬화한 RDF/XML 문서이므로 별도의 형식 변환 없이 RDF 파서가 RDF 데이터로 이용할 수 있다. 이는 모든 일반 XML 문서가 자동으로 RDF가 된다는 뜻은 아니다.
③ O — xmlns:rdf, xmlns:dcterms, xmlns:ex, xmlns:dcam 선언은 접두어를 각각의 URI에 연결한다. 이에 따라 dcterms:subject 같은 요소가 어느 어휘의 속성인지 명확해진다.
④ O — dcterms:subject는 더블린코어 메타데이터 어휘의 주제 요소이다. 제시문은 이 요소 안에서 해당 자원의 주제를 기술한다.
⑤ X — 틀린 표현은 ‘정보자원의 내용을 표현해주는 언어’이다. 제시문의 xml:lang="en"과 xml:lang="sv"는 각각 리터럴 값 ‘Biology’와 ‘Biologi’의 언어가 영어와 스웨덴어임을 표시한다. 올바르게는 ‘xml:lang 속성은 해당 문자열 리터럴의 언어를 나타낸다’이다. 정보자원 자체의 내용 언어는 dcterms:language 같은 속성으로 기술한다.

오답 포인트

⑤의 ‘정보자원의 내용 언어’와 ‘리터럴 문자열의 언어’를 혼동하기 쉽다. xml:lang은 바로 붙어 있는 값의 언어 태그이며, 자원 전체의 주된 언어를 선언하는 메타데이터 요소가 아니다.

다시 나오면 이렇게 판단

속성이 어디에 붙어 있는지를 먼저 확인한다. rdf:resource는 다른 자원의 URI, rdf:datatype은 리터럴의 데이터형, xml:lang은 리터럴의 자연언어를 나타낸다고 구분한다.

암기 포인트

xml:lang은 자원 전체의 언어가 아니라 바로 그 문자열 값의 언어 태그이다.

NoteBase

13. 정보행태 모델과 이를 주장한 학자의 연결이 옳지 않은 것은?

① ISP Model - Kuhlthau
② ELIS Model - Savolainen
③ Neutral Query Model - Taylor
④ Berrypicking Model - Bates
⑤ Sense Making Theory - Dervin

정답 3번

출제 포인트

정보추구행태를 설명하는 주요 모형과 이를 제안한 학자를 정확히 연결하는 문제이다.

핵심 정리

Kuhlthau의 ISP 모형: 정보탐색 과정에서 이용자의 감정·사고·행동이 단계적으로 변화하며, 특히 탐색 초기의 불확실성을 강조한다.
Savolainen의 ELIS 모형: 직업이나 학업 외의 일상생활에서 사람들이 생활양식을 유지하고 일상을 통제하기 위해 정보를 찾는 행태를 설명한다.
Dervin과 Dewdney의 중립적 질문: 이용자의 상황·정보 격차·정보의 예상 용도를 개방적 질문으로 파악하는 참고면담 기법이며, Dervin의 Sense-Making 관점에 기초한다.
Bates의 Berrypicking 모형: 탐색 과정에서 새 정보에 따라 질문이 계속 바뀌고 여러 탐색기법과 정보원을 옮겨 다니며 정보를 조금씩 모은다고 본다.
Dervin의 Sense-Making 이론: 사람이 특정 상황에서 마주친 인지적 격차를 정보로 메우고 앞으로 나아가는 과정을 설명한다.

NoteBase

선지별 해설

① O — ISP Model은 Kuhlthau가 제시한 정보탐색과정 모형이다. 이용자가 탐색 단계마다 겪는 불확실성과 감정·사고·행동의 변화를 함께 다룬다.
② O — ELIS Model은 Savolainen이 제시한 일상생활 정보추구 모형이다. 생활양식과 일상생활의 통제를 중심으로 일상적 정보추구를 설명한다.
③ X — 틀린 부분은 ‘Neutral Query Model - Taylor’라는 연결이다. 중립적 질문은 Dervin의 Sense-Making에 기초하여 Dervin과 Dewdney가 제시한 것이므로, 올바르게는 ‘Neutral Questioning - Dervin과 Dewdney’로 연결된다. Taylor는 정보요구를 내장적 요구, 의식적 요구, 형식화된 요구, 타협된 요구의 네 수준으로 설명한 학자이다.
④ O — Berrypicking Model은 Bates가 제시하였다. 탐색자가 여러 정보원에서 정보를 조금씩 수집하며 탐색 도중 질문과 탐색 방향을 수정한다는 모형이다.
⑤ O — Sense-Making Theory는 Dervin이 제시하였다. 개인이 상황 속에서 마주친 격차를 정보로 연결하여 의미를 구성하는 과정에 초점을 둔다.

출제자가 노린 함정

Taylor가 참고면담의 질문 협상과 정보요구의 표현 과정을 연구했다는 점 때문에 ③의 연결이 그럴듯해 보인다. 그러나 중립적 질문은 Dervin의 Sense-Making에 토대를 둔 Dervin·Dewdney의 기법이고, Taylor의 대표 연결어는 정보요구의 네 수준이다.

다시 나오면 이렇게 판단

모형 이름만 보지 않고 대표 개념과 학자를 함께 연결한다. ISP는 불확실성·Kuhlthau, ELIS는 일상생활·Savolainen, Berrypicking은 탐색 변화·Bates, Sense-Making은 상황-격차-이용·Dervin으로 판단한다. 중립적 질문은 Dervin과 Dewdney, 정보요구의 네 수준은 Taylor로 구분한다.

암기 포인트

Kuhlthau-ISP, Savolainen-ELIS, Bates-Berrypicking, Dervin-Sense-Making·중립적 질문, Taylor-정보요구 4단계로 묶어 기억한다.

NoteBase

14. DC 메타데이터 요소 중 지적재산권(Intellectual Property)에 해당되는 요소만을 바르게 나열한 것은?

① Creator, Identifier, Coverage
② Creator, Contributor, Publisher
③ Rights, Contributor, Identifier
④ Rights, Source, Publisher
⑤ Coverage, Source, Identifier

정답 2번

출제 포인트

DC 메타데이터 15개 요소를 내용 요소, 지적재산권 요소, 인스턴스화 요소로 구분할 수 있는지를 묻는 문제이다.

핵심 개념

지적재산권 요소는 자원의 제작 및 기여·발행 주체와 권리 정보를 나타내며 Creator, Contributor, Publisher, Rights가 해당한다.
내용 요소는 자원의 내용과 범위를 기술하며 Title, Subject, Description, Type, Source, Relation, Coverage가 해당한다.
인스턴스화 요소는 자원의 특정 표현형을 식별·관리하며 Date, Format, Identifier, Language가 해당한다.

선지별 해설

① X — Creator는 지적재산권 요소이지만 Identifier는 인스턴스화 요소이고 Coverage는 내용 요소이다. 틀린 표현은 Identifier와 Coverage이며, 지적재산권 요소만 나열하려면 이들을 Contributor, Publisher 또는 Rights로 바꾸어야 한다.
② O — Creator, Contributor, Publisher는 각각 자원의 제작자, 기여자, 발행 주체를 나타내므로 모두 지적재산권 요소이다.
③ X — Rights와 Contributor는 지적재산권 요소이지만 Identifier는 자원을 고유하게 식별하는 인스턴스화 요소이다. 틀린 표현은 Identifier이며, 이를 Creator 또는 Publisher로 바꾸면 모두 지적재산권 요소가 된다.
④ X — Rights와 Publisher는 지적재산권 요소이지만 Source는 현재 자원이 유래한 원천 자원을 나타내는 내용 요소이다. 틀린 표현은 Source이며, 이를 Creator 또는 Contributor로 바꾸면 모두 지적재산권 요소가 된다.
⑤ X — Coverage와 Source는 내용 요소이고 Identifier는 인스턴스화 요소이므로 지적재산권 요소가 하나도 없다. 올바르게는 Creator, Contributor, Publisher, Rights 중에서만 구성해야 한다.

NoteBase

오답 포인트

Rights라는 이름만 보고 권리 관련 선지를 고르거나, 자원의 출처를 뜻하는 Source와 발행 주체를 뜻하는 Publisher를 같은 범주로 보는 것이 함정이다. Rights가 포함되어 있어도 Source나 Identifier 같은 다른 범주의 요소가 섞이면 정답이 아니다.

다시 나오면 이렇게 판단

사람·기관과 권리를 나타내는 Creator, Contributor, Publisher, Rights를 먼저 한 묶음으로 분류한 뒤, 선지의 모든 요소가 이 네 가지 안에 들어가는지 확인한다. 한 요소라도 내용 요소나 인스턴스화 요소이면 제외한다.

암기 포인트

DC 지적재산권 4요소는 제작자(Creator)·기여자(Contributor)·발행자(Publisher)·권리(Rights)이다.

NoteBase

15. NISO(National Information Standards Organization)가 제시한 메타데이터의 상호운용성을 달성하는 방법이 아닌 것은?

① 데이터 리포지터리 지원하기
② 표준 메타데이터 스킴 사용하기
③ OAI의 메타데이터 형식과 수확 프로토콜 지원하기
④ '메타검색기'라는 분산통합 검색을 위한 프로토콜 지원하기
⑤ 하나의 스킴에서 사용한 요소를 다른 스킴의 요소로 매핑하기

정답 1번

출제 포인트

메타데이터의 상호운용성을 직접 높이는 표준화·프로토콜·요소 매핑 방법과 단순한 데이터 저장 기반의 지원을 구별하는 문제이다.

핵심 개념

메타데이터 상호운용성은 서로 다른 시스템이나 메타데이터 스킴이 메타데이터를 교환하고 그 의미를 일관되게 해석하여 활용할 수 있는 능력이다. 이를 위해 공통 스킴을 사용하거나, 교환·수확·검색 프로토콜을 지원하거나, 서로 다른 스킴의 요소 사이에 대응 관계를 설정한다. 데이터 리포지터리는 데이터를 저장·관리·제공하는 기반이지만, 그 자체가 메타데이터의 구조·의미·교환 방식을 맞추는 방법은 아니다.

선지별 해설

① X — 데이터 리포지터리 지원은 데이터의 저장·관리·접근을 위한 기반을 제공하는 것이므로 메타데이터 스킴 간 구조나 의미를 직접 맞추는 상호운용성 달성 방법이 아니다. 틀린 표현인 '데이터 리포지터리 지원하기'는 '표준 메타데이터 스킴이나 상호운용 프로토콜을 지원하기'로 고치면 적절하다.
② O — 표준 메타데이터 스킴을 사용하면 여러 시스템이 같은 요소 구조와 의미를 공유할 수 있어 상호운용성이 높아진다.
③ O — OAI의 메타데이터 형식과 수확 프로토콜 지원은 여러 저장소의 메타데이터를 표준화된 방식으로 노출하고 수집하게 하므로 상호운용성을 높인다.
④ O — 메타검색을 위한 프로토콜은 서로 다른 검색 시스템에 질의를 전달하고 결과를 통합할 수 있게 하므로 분산된 시스템 사이의 상호운용을 지원한다.
⑤ O — 한 스킴의 요소를 다른 스킴의 대응 요소에 매핑하는 크로스워크는 서로 다른 메타데이터 구조와 의미를 연결하는 대표적인 상호운용성 방법이다.

NoteBase

오답 포인트

①의 '리포지터리'가 메타데이터와 함께 자주 언급되므로 상호운용성 방법처럼 보이는 점이 함정이다.
리포지터리는 저장·관리 기반이고, 상호운용성은 공통 스킴, 교환·검색 프로토콜, 요소 매핑처럼 시스템 사이의 의미와 교환 방식을 연결하는 데 초점이 있다.

다시 나오면 이렇게 판단

선지가 서로 다른 시스템의 메타데이터 구조·의미·교환·검색 방식을 직접 연결하는지 판단한다. 표준 스킴, 프로토콜, 요소 매핑은 상호운용성 방법이고, 단순한 저장소 구축이나 저장 지원만으로는 직접적인 상호운용성 방법이 되지 않는다.

암기 포인트

상호운용성은 표준 스킴·프로토콜·매핑, 리포지터리는 저장·관리 기반이다.

NoteBase

16. 디지털도서관에 얼마나 많은 이용자가 접속했는지, 가장 많이 열람한 자료는 무엇인지 등을 파악하는 데 가장 유용한 평가 방법은?

① eye tracking
② thinking aloud
③ log file analysis
④ heuristic evaluation
⑤ performance measure

정답 3번

출제 포인트

이 문제는 디지털도서관의 실제 이용량과 자료별 열람 빈도를 파악하는 방법을 사용성 평가 기법과 구별하는 문제이다.

핵심 개념

로그 파일 분석은 시스템에 자동으로 남은 접속, 페이지 요청, 자료 열람 등의 기록을 분석하여 이용자 수와 이용 빈도를 정량적으로 파악하는 방법이다.
아이 트래킹은 이용자의 시선 위치와 이동을 추적하고, 사고 구술법은 과업 수행 중 떠오르는 생각을 말하게 하며, 발견적 평가는 전문가가 사용성 원칙에 따라 인터페이스를 점검한다. 수행도 측정은 과업 완료 시간, 성공률, 오류 수 등 이용자의 과업 수행 결과를 측정한다.

선지별 해설

① X — eye tracking은 화면에서 이용자의 시선이 머문 위치와 이동 경로를 측정하는 방법이므로, 전체 접속자 수나 자료별 열람 횟수를 직접 집계하는 방법이 아니다. 틀린 표현은 'eye tracking'이며, 올바르게는 시선의 위치·이동·주시 시간을 파악하는 방법이다.
② X — thinking aloud는 이용자가 과업을 수행하면서 생각과 판단 과정을 말하도록 하여 사용상의 문제를 찾는 방법이다. 틀린 표현은 'thinking aloud'이며, 올바르게는 이용자의 사고 과정과 불편 요인을 파악하는 질적 평가 방법이다.
③ O — log file analysis는 시스템에 축적된 접속 및 자료 요청 기록을 분석하므로 접속 이용자 수와 가장 많이 열람된 자료를 파악하는 데 가장 유용하다.
④ X — heuristic evaluation은 사용성 전문가가 정해진 휴리스틱, 즉 사용성 원칙에 비추어 인터페이스의 문제점을 점검하는 방법이다. 틀린 표현은 'heuristic evaluation'이며, 올바르게는 전문가가 사용성 원칙 위반 여부를 평가하는 방법이다.
⑤ X — performance measure는 이용자가 주어진 과업을 얼마나 빠르고 정확하게 수행하는지 과업 완료 시간, 성공률, 오류 수 등으로 측정하는 방법이다. 틀린 표현은 'performance measure'이며, 올바르게는 이용자의 과업 수행 효율과 효과성을 측정하는 방법이다.

NoteBase

오답 포인트

함정은 '이용자 평가'라는 넓은 표현만 보고 모든 사용성 평가 기법을 비슷하게 보는 데 있다. 이 문항의 핵심 단서인 '얼마나 많은 이용자가 접속했는지'와 '가장 많이 열람한 자료'는 실제 이용 기록의 횟수 집계를 요구하므로 로그 파일 분석과 연결된다.

다시 나오면 이렇게 판단

접속 수, 방문 횟수, 클릭 수, 자료별 열람 빈도처럼 시스템에 남는 실제 행동의 양을 묻는 경우에는 로그 파일 분석을 선택한다. 시선은 eye tracking, 생각 과정은 thinking aloud, 전문가 원칙 점검은 heuristic evaluation, 과업 시간·성공률·오류 수는 performance measure로 구별한다.

암기 포인트

접속·클릭·열람 횟수는 로그, 시선은 아이 트래킹, 생각은 사고 구술, 원칙 점검은 발견적 평가, 과업 성과는 수행도 측정이다.

NoteBase

17. 다음에서 설명하는 것은?

○ 미국의 MIT대학과 HP연구소가 개발한 기관 리포지터리
○ 오픈소스 소프트웨어로 공개됨
○ 이를 기반으로 우리나라에서는 OAK, KOASAS가 개발됨

① EPrints
② DSpace
③ arXiv
④ ROAR
⑤ Open DOAR

정답 2번

출제 포인트

기관 리포지터리 소프트웨어인 DSpace의 개발 주체와 공개 방식, 그리고 이를 기반으로 한 국내 리포지터리 사례를 묻는 문제이다.

핵심 개념

DSpace는 대학·연구기관이 생산한 디지털 학술자료를 수집·보존·공개하기 위한 오픈소스 기관 리포지터리 소프트웨어이다. 미국 MIT와 HP 연구소가 공동 개발했으며, 국내의 OAK와 KAIST의 KOASAS 등도 DSpace 기반 사례에 해당한다.
EPrints는 별도의 오픈소스 리포지터리 소프트웨어이고, arXiv는 학문 분야 중심의 전자논문 아카이브이다. ROAR와 OpenDOAR는 리포지터리 자체를 구축하는 소프트웨어가 아니라 리포지터리 정보를 모아 제공하는 등록·디렉터리 서비스이다.

NoteBase

선지별 해설

① X — EPrints는 오픈소스 기관 리포지터리 소프트웨어이지만, 제시된 'MIT와 HP 연구소의 공동 개발' 및 'OAK·KOASAS의 기반'에 해당하는 것은 DSpace이다. 틀린 표현 'EPrints'는 'DSpace'로 고쳐야 한다.
② O — DSpace는 MIT와 HP 연구소가 공동 개발해 오픈소스로 공개한 기관 리포지터리 소프트웨어이며, 제시된 국내 사례와도 일치한다.
③ X — arXiv는 물리학·수학·컴퓨터과학 등 여러 학문 분야의 논문을 축적·공개하는 주제 분야 중심 전자논문 아카이브이다. 틀린 표현 'arXiv'는 'DSpace'로 고쳐야 한다.
④ X — ROAR는 오픈액세스 리포지터리의 현황과 정보를 등록·검색할 수 있게 하는 레지스트리이다. 틀린 표현 'ROAR'는 'DSpace'로 고쳐야 한다.
⑤ X — OpenDOAR는 전 세계 오픈액세스 리포지터리 정보를 수집해 제공하는 디렉터리이다. 틀린 표현 'OpenDOAR'는 'DSpace'로 고쳐야 한다.

오답 포인트

EPrints도 기관 리포지터리용 오픈소스 소프트웨어라는 점이 가장 큰 함정이다. 그러나 'MIT·HP 공동 개발'이라는 개발 주체를 기준으로 보면 DSpace로 확정된다. 또한 ROAR와 OpenDOAR는 리포지터리를 구축하는 소프트웨어가 아니라 리포지터리 정보를 모아 제공하는 서비스이다.

다시 나오면 이렇게 판단

개발 주체, 서비스 기능, 국내 적용 사례를 차례로 대조한다. 'MIT와 HP', '기관 리포지터리 소프트웨어', 'OAK·KOASAS 기반'이 함께 제시되면 DSpace로 판단한다.

암기 포인트

MIT·HP가 만든 기관 리포지터리는 DSpace, 리포지터리 목록 서비스는 ROAR·OpenDOAR이다.

NoteBase

18. 서지결합분석에 관한 설명으로 옳지 않은 것은?

① 논문 간 간접 인용을 통한 관련도를 알 수 있다.
② 인용색인 데이터베이스에서 관련 레코드를 검색하는 데 활용될 수 있다.
③ 주제적으로 관련 있는 문헌 집단을 형성하여 학문의 지적구조를 제시할 수 있다.
④ 서지에서 추출한 단어 간 연관성을 산출하여 학문분야를 파악하는 것을 말한다.
⑤ 서지결합은 케슬러가 처음 제안하였다.

정답 4번

출제 포인트

서지결합의 기준이 공통으로 인용한 문헌인지, 서지에서 추출한 단어의 연관성인지 구별하는 문제이다.

핵심 개념

서지결합은 두 문헌이 동일한 선행 문헌을 함께 인용할수록 두 문헌의 주제적 관련성이 높다고 보는 계량서지학적 분석 방법이다. 결합 강도는 두 문헌의 참고문헌 목록에 공통으로 포함된 인용 문헌의 수를 바탕으로 판단한다.
서지결합은 '공통으로 인용한 문헌'을 기준으로 하고, 동시단어분석은 제목·초록·키워드 등에서 추출한 '단어의 동시 출현과 연관성'을 기준으로 한다.

선지별 해설

① O — 두 논문이 서로를 직접 인용하지 않더라도 동일한 선행 문헌을 인용하면 공통 참고문헌을 매개로 두 논문의 관련도를 파악할 수 있다.
② O — 인용색인 데이터베이스에는 문헌의 인용·피인용 관계를 추적할 수 있는 레코드가 구축되어 있으므로, 공통으로 인용된 문헌을 검색하여 서지결합 분석에 활용할 수 있다.
③ O — 공통 인용 문헌을 많이 가진 문헌들을 묶으면 주제적으로 관련된 문헌 집단을 형성할 수 있고, 이를 통해 학문 분야의 지적 구조를 파악할 수 있다.
④ X — 틀린 표현은 '서지에서 추출한 단어 간 연관성을 산출'한다는 부분이다. 이는 동시단어분석에 대한 설명이며, 서지결합은 두 문헌의 참고문헌에서 공통으로 인용한 문헌을 찾아 문헌 간 관련성을 산출하는 방법이다.
⑤ O — 서지결합은 케슬러가 처음 제안한 방법이다.

오답 포인트

④의 '서지에서 추출한 단어'라는 표현 때문에 서지정보를 이용하는 분석으로 보일 수 있다. 그러나 분석 단위가 단어이면 동시단어분석이고, 공통 참고문헌이면 서지결합이다.

NoteBase

다시 나오면 이렇게 판단

선지에서 무엇의 관계를 세는지 확인한다. 두 문헌이 함께 인용한 선행 문헌을 세면 서지결합이고, 서지의 제목·초록·키워드에서 함께 나타난 단어를 세면 동시단어분석이다.

암기 포인트

서지결합은 공통 참고문헌, 동시단어분석은 공통 출현 단어이다.

NoteBase

19. 다음에서 설명하는 것은?

> ○ 메타데이터 요소 세트에 수록된 개별적인 요소를 식별하는 것이다.
> ○ 특정 메타데이터 스킴에서 사용하는 메타데이터 요소세트가 다양한 커뮤니티에서 범용적으로 사용될 수 있도록 한다.
> ○ 메타데이터 구축자가 자신이 원하는 요소를 선택하여 사용할 수 있는 장점이 있다.

① Namespace
② Dublin Core
③ XML
④ MODS
⑤ RDF

정답 1번

출제 포인트

메타데이터 요소의 이름을 고유하게 식별하고 서로 다른 스키마의 요소를 함께 사용할 수 있게 하는 Namespace의 기능을 묻는 문제이다.

핵심 개념

Namespace는 이름의 집합을 고유한 식별자와 연결하여, 서로 다른 메타데이터 스키마에서 같은 이름을 사용하더라도 요소를 구별하게 하는 체계이다. 이를 이용하면 여러 공동체가 만든 요소 집합을 한 응용 환경에서 함께 사용하면서 이름 충돌을 피할 수 있다.
Dublin Core와 MODS는 메타데이터 요소 집합·스키마이고, XML은 구조화 문서의 마크업 언어이며, RDF는 자원과 그 관계를 기술하는 데이터 모델이다.

NoteBase

선지별 해설

① O — Namespace는 메타데이터 요소가 어느 요소 집합에 속하는지를 고유하게 식별하고, 서로 다른 요소 집합을 조합하여 사용할 수 있게 한다. 제시문의 세 설명에 해당한다.
② X — 틀린 부분은 'Dublin Core'이다. Dublin Core는 제목, 작성자, 주제 등 자원을 기술하기 위한 범용 메타데이터 요소 집합이며, 제시문처럼 요소 이름을 식별하고 여러 스키마의 요소를 함께 쓰게 하는 장치는 Namespace이다.
③ X — 틀린 부분은 'XML'이다. XML은 사용자가 태그를 정의하여 구조화된 문서를 표현하는 마크업 언어이다. XML 문서에서 서로 다른 어휘의 이름 충돌을 막고 요소를 구별하는 구체적 장치는 XML Namespace이다.
④ X — 틀린 부분은 'MODS'이다. MODS는 도서관 서지정보를 비교적 상세하게 기술하기 위한 메타데이터 스키마이다. 서로 다른 메타데이터 요소 집합의 이름을 식별하고 결합하는 체계는 Namespace이다.
⑤ X — 틀린 부분은 'RDF'이다. RDF는 자원에 관한 정보를 주어-술어-목적어의 삼중 구조로 표현하는 데이터 모델이다. 제시문이 설명하는 요소 이름의 고유 식별과 여러 스키마의 결합 기능은 Namespace의 기능이다.

오답 포인트

③의 XML은 Namespace를 사용할 수 있는 문법적 환경이어서 혼동하기 쉽다. 그러나 제시문은 마크업 언어 자체가 아니라 요소 이름을 고유하게 구별하고 여러 어휘를 함께 쓰게 하는 이름공간의 기능을 설명한다.

다시 나오면 이렇게 판단

'요소 이름의 고유 식별', '서로 다른 스키마의 요소 결합', '이름 충돌 방지'가 함께 제시되면 Namespace로 판단한다. 요소 집합 자체를 묻는다면 Dublin Core나 MODS, 문서 표현 언어를 묻는다면 XML, 자원 관계의 삼중 구조를 묻는다면 RDF로 구별한다.

암기 포인트

Namespace는 이름을 구별하고 어휘를 결합한다. Dublin Core·MODS는 요소 집합, XML은 마크업 언어, RDF는 삼중 구조 데이터 모델이다.

NoteBase

20. 오픈액세스에 대한 설명과 가장 거리가 먼 것은?

① 학술정보 구독비용 부담이 증가함에 따라 학술정보를 보다 자유롭게 이용할 수 있도록 하기 위해 등장한 운동이다.
② BOAI 선언과 Bethesda 선언은 오픈액세스 관련 주요 가이드라인을 제공했다.
③ 공적자금이 투입되어 산출된 연구 결과물은 국민 모두가 사용할 권리가 있는 국가적 자산이라는 인식에서 비롯되었다.
④ PubMed는 오픈액세스를 실천한 사례이다.
⑤ 구독 기반 학술지와 다른 방식의 출판비용모델을 확보해야 한다.

정답 3번

출제 포인트

오픈액세스의 등장 배경, 주요 선언, 실제 사례, 비용 모델을 구분하고 오픈액세스를 국가적 자산에 대한 국민의 권리라는 주장과 혼동하지 않는지를 묻는 문제이다.

핵심 개념

오픈액세스는 학술 연구 성과를 인터넷에서 누구나 무료로 접근하고, 허용된 조건에 따라 이용할 수 있게 하는 학술정보 유통 방식이다. 핵심은 접근 장벽의 제거와 이용 허용이며, 연구 결과를 국가 소유의 자산으로 보거나 모든 국민에게 당연한 사용권이 있다고 보는 개념은 아니다. 구독료로 비용을 회수하는 전통적 모델과 달리 저자처리비용, 기관·학회 지원, 공공 지원 등 다양한 비용 조달 방식이 활용될 수 있다.

선지별 해설

① O — 학술지 구독료 상승으로 도서관과 연구자의 접근 부담이 커진 이른바 학술지 위기가 오픈액세스 운동의 주요 등장 배경이다. 학술정보를 비용 장벽 없이 이용하게 하려는 취지와 부합한다.
② O — BOAI 선언과 Bethesda 선언은 오픈액세스의 의미와 실천 원칙을 제시한 대표적인 초기 선언이다. 무료 접근뿐 아니라 학술 성과의 이용·배포 범위에 관한 방향도 제시하였다.
③ X — 틀린 표현은 '연구 결과물은 국민 모두가 사용할 권리가 있는 국가적 자산'이라는 부분이다. 오픈액세스는 연구 결과의 소유권을 국가에 귀속시키거나 국민에게 당연한 사용권을 부여한다는 인식에서 비롯된 것이 아니다. 올바르게는 학술지 가격 상승에 따른 접근 위기와 디지털 네트워크의 발달을 배경으로, 저작권자가 허용한 조건 아래 학술정보의 접근·이용 장벽을 낮추려는 운동이라고 해야 한다.
④ O — PubMed는 의생명과학 문헌의 서지정보와 초록을 누구나 무료로 검색·열람하게 한다는 점에서 학술정보 접근성을 높인 사례이다. 다만 논문 전문까지 공개하는 대표적 저장소는 PubMed Central이므로 둘을 구분한다.
⑤ O — 독자 구독료를 받지 않고 공개하려면 출판·심사·보존 비용을 충당할 대체 모델이 필요하다. 저자처리비용, 기관 또는 학회의 지원, 공공 재원 등이 그 예이다.

NoteBase

오답 포인트

③은 '공공재원으로 생산된 연구 성과는 대중에게 공개되어야 한다'는 정책적 주장과 오픈액세스의 일반적 정의를 섞은 선지이다. 공개 필요성을 강조할 수는 있지만, 이를 곧바로 '국가적 자산'이나 '국민 모두의 당연한 사용권'으로 표현하면 소유권·저작권 문제까지 잘못 단정하게 된다. 또한 ④에서는 무료 서지정보 서비스인 PubMed와 무료 전문 저장소인 PubMed Central을 혼동하기 쉽다.

다시 나오면 이렇게 판단

오픈액세스 선지는 '온라인 무료 접근', '이용 장벽 완화', '대표 선언', '대체 비용 모델'과 연결되면 타당성을 검토한다. 반대로 국가 소유, 국민의 당연한 사용권처럼 소유권이나 권리 귀속을 일괄 단정하면 오픈액세스의 정의를 넘어선 표현인지 판단한다.

암기 포인트

오픈액세스는 국가 소유가 아니라 접근 장벽의 제거이다. BOAI·Bethesda는 선언, PubMed는 무료 검색, 비용은 구독료 외의 모델로 충당한다.

NoteBase