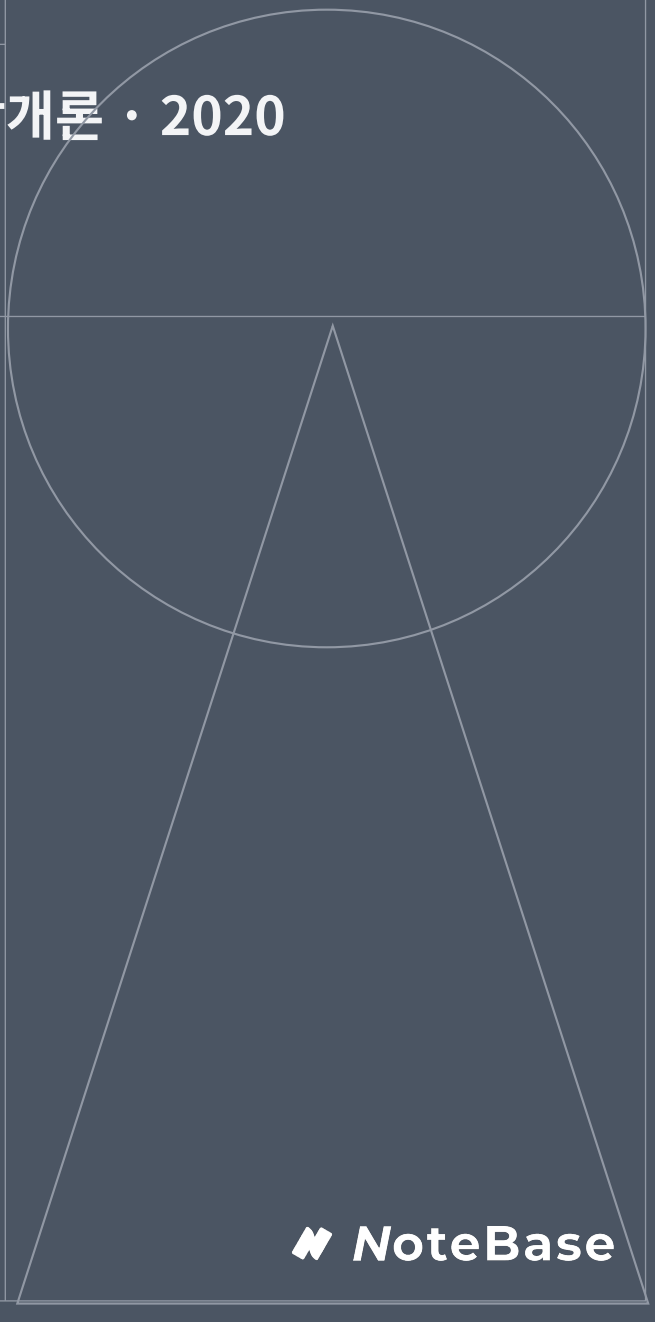
기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2020



NoteBase

기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2020

공무원/국회직/9급 · 정보학개론 · 2020

1. **브라우징을 통한 검색의 내용으로 옳지 않은 것은?**

① 탐색과 달리 구조화된 검색방법이 아니다.
② 브라우징은 훈련이나 연습없이 수행될 수 있다.
③ 검색하는 동안 우연한 발견을 할 수 있다.
④ 정보요구가 광범위하고 모호할 때 사용한다.
⑤ “from what to where” 행위로 간주될 수 있다.

정답 5번

출제 포인트

이 문제는 브라우징과 탐색의 차이, 특히 정보요구의 명확성·검색 과정의 구조성·출발점과 도착점의 방향을 묻는 문제이다.

핵심 개념

브라우징은 구체적인 목표나 정교한 검색식 없이 정보공간을 훑어보며 필요한 정보를 발견하는 비구조적·탐색적 검색행동이다. 이미 알고 있는 정보원이나 위치에서 출발해 그 안에서 무엇이 있는지를 발견하므로 ‘from where to what’의 행위로 설명된다. 반면 탐색은 찾으려는 대상이 비교적 분명한 상태에서 그것이 있는 위치를 찾아가는 구조화된 검색으로, ‘from what to where’의 행위에 가깝다.

선지별 해설

① O — 브라우징은 미리 정한 검색어와 절차에 따라 진행되는 구조화된 탐색과 달리, 정보공간을 훑으며 진행하는 비구조적 검색방법이다.
② O — 브라우징은 복잡한 검색식이나 전문적인 검색 훈련 없이도 정보원이나 화면을 둘러보며 수행할 수 있다.
③ O — 검색 도중 처음부터 찾으려던 것이 아닌 유용한 정보를 우연히 발견하는 우연성은 브라우징의 대표적 특징이다.
④ O — 원하는 정보가 정확히 정해지지 않고 정보요구가 광범위하거나 모호할 때 여러 정보를 살펴보는 브라우징이 적합하다.
⑤ X — 틀린 표현은 ‘from what to where’이다. 이는 찾을 대상이 무엇인지 알고 그 위치를 찾아가는 탐색의 방향에 해당한다. 브라우징은 정보원이나 위치에서 출발해 무엇이 있는지를 발견하는 ‘from where to what’의 행위로 간주된다.

NoteBase

오답 포인트

⑤는 what과 where의 순서만 바꾸어 브라우징과 탐색을 혼동하게 한다. 브라우징은 먼저 둘러볼 곳(where)이 있고 그 안에서 정보(what)를 발견하는 행동이라는 점이 함정의 핵심이다.

다시 나오면 이렇게 판단

목표가 모호하고 정보공간을 훑다가 우연히 발견하면 브라우징, 찾을 대상이 명확하고 그 소재를 찾아가면 탐색으로 판단한다. 방향은 브라우징이 where→what, 탐색이 what→where이다.

암기 포인트

브라우징은 '어디에서 무엇을'(where→what), 탐색은 '무엇을 어디에서'(what→where)이다.

NoteBase

2. <보기>에서 성격이 다른 학술DB끼리 짝지어진 것만을 고르면?

<보 기>

ㄱ. Web of Science - SCOPUS
ㄴ. PMC - BioMed Central
ㄷ. PQDT - DDOD
ㄹ. RefWorks - ScienceDirect
ㅁ. JCR - SAGE journals

① ㄱ, ㄴ
② ㄱ, ㄷ
③ ㄴ, ㅁ
④ ㄷ, ㄹ
⑤ ㄹ, ㅁ

정답 5번

출제 포인트

학술 데이터베이스와 학술정보 도구를 기능별로 구분하여, 두 항목의 성격이 서로 다른 짝을 찾는 문제이다.

핵심 개념

인용색인 데이터베이스는 논문의 서지정보와 인용 관계를 검색·분석하는 서비스이고, 원문 데이터베이스는 학술지 논문 등의 원문을 제공하는 서비스이다. 참고문헌 관리도구는 이용자가 수집한 서지정보와 인용문헌을 정리하는 도구이며, 인용분석 도구는 학술지의 인용 지표와 영향력을 평가하는 데 쓰인다.
ㄱ Web of Science와 SCOPUS: 모두 대표적인 인용색인 데이터베이스이다.
ㄴ PMC와 BioMed Central: 모두 생의학 분야의 오픈액세스 원문을 제공한다.
ㄷ PQDT와 DDOD: 모두 학위논문 정보를 제공하는 데이터베이스이다.
ㄹ RefWorks는 참고문헌 관리도구이고 ScienceDirect는 학술 원문 데이터베이스이므로 성격이 다르다.
ㅁ JCR은 학술지 인용분석 도구이고 SAGE journals는 학술 원문 제공 플랫폼이므로 성격이 다르다.

NoteBase

선지별 해설

① X
ㄱ과 ㄴ은 모두 같은 성격끼리 짝지어져 있으므로 조건에 맞지 않는다. 틀린 조합은 'ㄱ, ㄴ'이며, 성격이 다른 짝의 올바른 조합은 'ㄹ, ㅁ'이다.

② X
ㄱ은 같은 인용색인 데이터베이스끼리의 짝이고 ㄷ은 같은 학위논문 데이터베이스끼리의 짝이므로 조건에 맞지 않는다. 틀린 조합은 'ㄱ, ㄷ'이며, 올바르게는 'ㄹ, ㅁ'이다.

③ X
ㄴ은 둘 다 생의학 분야의 오픈액세스 원문을 제공하므로 성격이 같고, ㅁ만 성격이 다르다. 틀린 조합은 'ㄴ, ㅁ'이며, 올바르게는 'ㄹ, ㅁ'이다.

④ X
ㄹ은 RefWorks와 ScienceDirect의 기능이 달라 맞지만, ㄷ의 PQDT와 DDOD는 모두 학위논문 데이터베이스이므로 틀렸다. 틀린 항목은 'ㄷ'이며, 이를 'ㅁ'으로 고치면 'ㄹ, ㅁ'이 된다.

⑤ O
ㄹ은 참고문헌 관리도구와 원문 데이터베이스의 짝이고, ㅁ은 인용분석 도구와 원문 제공 플랫폼의 짝이다. 두 짝 모두 구성 항목의 성격이 서로 다르므로 정답이다.

오답 포인트

서비스가 논문을 다룬다는 공통점만 보고 모두 같은 종류로 판단하면 안 된다. 특히 RefWorks는 논문 원문을 제공하는 데이터베이스가 아니라 참고문헌을 관리하는 도구이고, JCR은 원문 검색 서비스가 아니라 학술지 인용 지표를 분석하는 도구라는 점이 핵심 함정이다.

다시 나오면 이렇게 판단

각 서비스의 주된 기능을 '인용색인·원문 제공·학위논문·참고문헌 관리·인용분석' 중 하나로 먼저 분류한 뒤, 한 짝 안의 두 기능이 일치하는지 비교한다.

암기 포인트

RefWorks는 참고문헌 관리, ScienceDirect와 SAGE journals는 원문 제공, JCR은 학술지 인용분석이다.

NoteBase

3. 문헌 클러스터링에 대한 설명으로 옳지 않은 것은?

① 높은 연관성을 갖는 문서들을 낮은 연관성을 갖는 문서들과 분리해서 집단화하는 기법이다.

② 문서와 문서간의 관련성을 확인하기 위해 문서간의 유사도를 계산하여 상호 관련성 여부를 확인한다.

③ 계층적 클러스터링 방법은 문헌집단 내의 모든 대상물이 연결될때까지 대상물이나 클러스터 쌍들을 연속적으로 연결해가는 기법이다.

④ 비계층적 클러스터링은 처리시간 측면에서 계층적 기법에 비해 덜 효율적인 반면, 클러스터링 성능이 우수하다.

⑤ 문서 클러스터링에서 유사도 계산에 사용되는 기법으로는 다이스 계수, 자카드 계수, 코사인 계수 등이 있다.

정답 4번

출제 포인트

문헌 클러스터링의 기본 원리와 계층적·비계층적 클러스터링의 처리 특성을 구분하는 문제이다.

핵심 개념

문헌 클러스터링: 문헌 사이의 유사도 또는 관련성을 기준으로 비슷한 문헌들을 같은 집단으로 묶는 기법이다.
계층적 클러스터링: 개별 문헌이나 작은 군집을 단계적으로 합치거나 큰 군집을 나누어 계층 구조를 만드는 방식이다. 군집 수를 미리 정하지 않아도 되지만, 유사도 계산과 반복적인 결합 과정 때문에 대규모 자료에서는 계산 부담이 커질 수 있다.
비계층적 클러스터링: 보통 군집 수를 정한 뒤 문헌을 군집에 배정하고 재배치하는 방식이다. 일반적으로 계층적 방식보다 처리시간 측면에서 효율적이지만, 초기 중심이나 군집 수의 설정에 따라 결과가 달라질 수 있으므로 성능이 언제나 우수하다고 단정할 수 없다.

NoteBase

선지별 해설

① O — 문헌 클러스터링은 문헌 간 유사도나 관련성을 기준으로 유사성이 높은 문헌은 같은 집단에, 낮은 문헌은 다른 집단에 배치하는 기법이다.
② O — 문헌 사이의 관련성을 판별하려면 문헌을 표현한 벡터나 속성을 바탕으로 유사도를 계산하고, 그 값으로 상호 관련성 여부를 판단한다.
③ O — 응집형 계층적 클러스터링은 개별 문헌에서 시작하여 가장 가까운 문헌 또는 군집을 단계적으로 연결·병합하며, 최종적으로 전체가 하나의 계층 구조로 연결될 때까지 진행할 수 있다.
④ X — 틀린 표현은 '처리시간 측면에서 계층적 기법에 비해 덜 효율적인 반면, 클러스터링 성능이 우수하다'이다. 올바르게는 비계층적 클러스터링이 일반적으로 계층적 기법보다 처리시간 측면에서 더 효율적이며, 클러스터링 성능은 초기값·군집 수·자료 특성 등에 따라 달라지므로 항상 더 우수하다고 할 수 없다.
⑤ O — 이진 속성이나 벡터로 표현된 문헌의 유사도를 측정하는 방법에는 다이스 계수, 자카드 계수, 코사인 계수 등이 사용된다.

오답 포인트

④는 비계층적 클러스터링의 장단점을 서로 뒤바꿔 제시한다. 비계층적 방식은 대체로 계산 효율성이 장점이지만, 결과의 품질은 초기값과 군집 수 등에 영향을 받으므로 계층적 방식보다 항상 우수하다고 일반화할 수 없다.

다시 나오면 이렇게 판단

'계층적'은 단계적 병합·분할과 계층 구조, '비계층적'은 미리 정한 군집 수와 반복적 재배치를 핵심 기준으로 구분한다. 처리시간은 일반적으로 비계층적 방식이 더 효율적이며, 성능의 절대적 우열을 단정하는 표현은 의심한다.

암기 포인트

계층적은 단계별 연결, 비계층적은 빠른 재배치이다. 비계층적 방식은 대체로 빠르지만 언제나 더 정확한 것은 아니다.

NoteBase

4. 다음은 Tim Berners-Lee가 제안한 5-star Open Data의 내용이다. 괄호에 들어갈 ㉠, ㉡, ㉢의 내용을 순서대로 바르게 나열한 것은?

☆ 1단계 : 온라인상에서 활용 가능하도록 공개된 상태
☆☆ 2단계 : (㉠) 공개된 상태
☆☆☆ 3단계 : 개방형 데이터 형태로 공개된 상태
☆☆☆☆ 4단계 : (㉡) 공개된 상태
☆☆☆☆☆ 5단계 : (㉢) 공개된 상태

	㉠	㉡	㉢
①	기계가 읽을 수 있는 형태로	URI로 식별되는 형태로	Linked Data로
②	DOI가 부여된 형태로	Open API로	Linked Data로
③	URI로 식별되는 형태로	Linked Data로	Open API로
④	DOI가 부여된 형태로	기계가 읽을 수 있는 형태로	Open API로
⑤	기계가 읽을 수 있는 형태로	DOI가 부여된 형태로	Linked Data로

정답 1번

출제 포인트

Tim Berners-Lee의 5-star Open Data에서 단계가 높아질 때 추가되는 공개 조건을 구별하는 문제이다.

핵심 개념

1단계: 웹에서 이용할 수 있도록 공개된 데이터이다.
2단계: 컴퓨터가 처리할 수 있는 구조화된 형태의 데이터이다. 핵심은 기계 판독 가능성이다.
3단계: 특정 소프트웨어에 종속되지 않는 개방형 형식의 데이터이다.
4단계: 대상을 URI로 식별하여 웹에서 참조할 수 있는 데이터이다.
5단계: 다른 데이터와 연결하여 맥락을 제공하는 Linked Data이다.
따라서 빈칸은 ㉠ 기계가 읽을 수 있는 형태, ㉡ URI로 식별되는 형태, ㉢ Linked Data로 순서대로 채워진다.

NoteBase

선지별 해설

① O — 2단계의 기계 판독 가능 형태, 4단계의 URI 식별, 5단계의 Linked Data를 정확히 배열하였다.
② X — ㉠의 'DOI가 부여된 형태'와 ㉡의 'Open API'가 틀렸다. 올바르게는 ㉠은 기계가 읽을 수 있는 형태이고, ㉡은 URI로 식별되는 형태이다. ㉢의 Linked Data는 맞다.
③ X — 세 빈칸의 단계가 모두 맞지 않는다. 올바르게는 ㉠은 기계가 읽을 수 있는 형태, ㉡은 URI로 식별되는 형태, ㉢은 Linked Data이다. URI 식별은 4단계이고 Linked Data는 5단계이며, Open API는 5단계의 정의가 아니다.
④ X — 'DOI가 부여된 형태', '기계가 읽을 수 있는 형태', 'Open API'의 배치가 모두 틀렸다. 올바르게는 2단계가 기계 판독 가능 형태, 4단계가 URI 식별 형태, 5단계가 Linked Data이다.
⑤ X — ㉠의 '기계가 읽을 수 있는 형태'와 ㉢의 'Linked Data'는 맞지만, ㉡의 'DOI가 부여된 형태'가 틀렸다. 올바르게는 4단계인 ㉡에 URI로 식별되는 형태가 들어간다.

오답 포인트

DOI와 Open API도 데이터 공개·활용과 관련된 용어이므로 단계의 핵심 조건처럼 보이지만, 이 모형의 4단계 기준은 URI를 이용한 식별이고 5단계 기준은 다른 데이터와의 연결이다. 특히 ⑤는 2단계와 5단계를 맞게 제시하고 4단계만 DOI로 바꾼 함정이다.

다시 나오면 이렇게 판단

단계를 '웹 공개 → 기계 판독 → 개방형 형식 → URI 식별 → 데이터 연결'의 누적 순서로 대조한다. 선지에 DOI나 Open API가 제시되어도 이 다섯 핵심 조건 중 어느 단계의 정의인지부터 확인한다.

암기 포인트

별 2개는 기계, 별 3개는 개방형 형식, 별 4개는 URI, 별 5개는 연결이다.

NoteBase

5. 다음은 XML구문을 이용하여 DC(Dublin Core) 레코드를 작성한 사례의 일부이다. 이에 대한 설명으로 옳지 않은 것은?

```
<?xml version ="1.0"?>
<metadata
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:dcterms="http://purl.org/dc/terms/">
  <dc:title> UKOLN </dc:title>
  <dcterms:alternative> UK Office for Library and
  Information Networking </dcterms:alternative>
  <dc:subject scheme="dcterms:DDC"> 062 </dc:subject>
  <dc:description> UKOLN is a national focus expertise
  in digital information management.                    (가)
  ...(중략)...</dc:description>
  ...(중략)...
</metadata>
```

① xml 선언이 있다.
② 네임스페이스를 사용하지 않은 DC 레코드 작성 사례이다.
③ 세분화요소 <alternative>를 사용하여 표제의 이형을 기술하고 있다.
④ 인코딩 스킴을 사용하여 레코드 작성의 일관성을 확보하고 있다.
⑤ (가)부분에 밑줄 친 <description>은 MODS의 상위요소인 <abstract>, <tableOfContents> 등과 대응될 수 있다.

정답 2번

출제 포인트

XML로 표현한 Dublin Core 레코드에서 XML 선언, 네임스페이스, 세부화요소, 인코딩 스킴, 메타데이터 요소 간 대응 관계를 식별하는 문제이다.

핵심 개념

네임스페이스는 서로 다른 메타데이터 어휘의 요소명을 구별하기 위해 접두어와 URI를 연결하는 장치이다. 이 사례에서는 xmlns:dc가 DC 요소 집합을, xmlns:dcterms가 DCMI Terms를 선언한다.
dcterms:alternative는 표제의 대체형을 나타내는 세부화요소이고, dcterms:DDC와 같은 인코딩 스킴은 주제값이 어떤 분류 체계에 따라 작성되었는지를 밝힌다.

NoteBase

선지별 해설

① O — 첫 줄의 <?xml version="1.0"?>은 XML 버전을 명시한 XML 선언이다.
② X — 틀린 표현은 '네임스페이스를 사용하지 않은'이다. 실제로는 metadata 시작 태그에서 xmlns:dc와 xmlns:dcterms를 선언하고, dc:title·dc:subject·dc:description 및 dcterms:alternative와 같은 접두어가 붙은 요소를 사용한 네임스페이스 적용 사례이다.
③ O — dcterms:alternative는 title을 세부화하여 표제의 이형, 즉 대체 표제를 기술한다. 사례에서는 'UKOLN'의 확장된 명칭을 대체 표제로 제시한다.
④ O — dc:subject의 scheme="dcterms:DDC"는 값 '062'가 DDC라는 인코딩 스킴에 따른 분류기호임을 명시하므로 값의 해석과 기술 방식의 일관성을 높인다.
⑤ O — DC의 description은 자원의 내용에 대한 설명을 담는 요소이므로 MODS로 변환할 때 내용 요약을 담는 abstract나 목차 정보를 담는 tableOfContents 등과 의미에 따라 대응될 수 있다.

오답 포인트

②는 별도의 namespace라는 단어가 본문에 없다는 점을 함정으로 삼는다. XML에서는 xmlns:접두어="URI" 형식 자체가 네임스페이스 선언이므로, xmlns:dc와 xmlns:dcterms가 보이면 네임스페이스를 사용한 것이다.

다시 나오면 이렇게 판단

먼저 xmlns 선언과 접두어 사용 여부를 확인하고, 그다음 요소명은 의미를, scheme 속성은 값의 표현 체계나 통제 어휘를 나타내는지 구분한다. 서로 다른 메타데이터 스키마의 요소 대응은 이름의 문자적 일치보다 실제 의미와 기능을 기준으로 판단한다.

암기 포인트

xmlns는 네임스페이스, alternative는 대체 표제, scheme은 값의 표현 체계이다.

NoteBase

6. 클라우드 컴퓨팅에 대한 설명으로 옳지 않은 것은?

① Peer-to-Peer 기반의 개방형 시스템 운영 모형을 추구한다.
② 모바일 환경에 결합되면서 시너지를 창출할 수 있다.
③ 도서관 협력망에 효과적으로 적용이 가능하다.
④ IT 자원의 주문형 아웃소싱 서비스라고 할 수 있다.
⑤ 개별 기관의 전산 운영 비용 절감이 가능하다.

정답 1번

출제 포인트

클라우드 컴퓨팅의 운영 방식과 활용 효과를 구분하고, 클라우드와 Peer-to-Peer 방식의 특징을 혼동하지 않는지를 묻는 문제이다.

핵심 개념

클라우드 컴퓨팅: 서버·저장장치·소프트웨어 등의 IT 자원을 네트워크를 통해 필요한 만큼 서비스로 제공받는 방식이다. 핵심 키워드는 주문형 이용, 자원 공유, 확장성, 비용 절감이다.
Peer-to-Peer: 중앙 서버에만 의존하지 않고 참여자들의 컴퓨터가 서로 자원을 주고받는 분산형 통신 방식이다. 클라우드 컴퓨팅의 일반적인 운영 모형과는 구별된다.

선지별 해설

① X — 'Peer-to-Peer 기반의 개방형 시스템 운영 모형'이라는 표현이 틀렸다. 클라우드 컴퓨팅은 일반적으로 서비스 제공자의 데이터센터에 집약된 컴퓨팅 자원을 네트워크를 통해 주문형 서비스로 제공하는 모형이다.
② O — 클라우드는 모바일 기기가 장소에 구애받지 않고 네트워크를 통해 데이터와 응용 서비스에 접근하게 하므로 모바일 환경과 결합하여 시너지를 낼 수 있다.
③ O — 여러 도서관이 클라우드 기반의 시스템과 자원을 공동으로 이용하면 중복 구축을 줄이고 협력망의 자원 공유와 공동 서비스를 효율화할 수 있다.
④ O — 필요한 서버·저장공간·소프트웨어 등의 IT 자원을 외부 제공자로부터 필요한 때에 서비스로 이용하므로 주문형 아웃소싱 서비스라고 볼 수 있다.
⑤ O — 개별 기관이 자체 전산 설비를 각각 구축·유지하는 부담을 줄이고 사용량에 따라 자원을 이용할 수 있으므로 전산 운영 비용을 절감할 수 있다.

오답 포인트

①의 '개방형'이라는 말 때문에 인터넷 기반 서비스라는 점과 혼동하기 쉽다. 그러나 틀린 핵심은 'Peer-to-Peer 기반'이다. 클라우드의 자원 공유는 제공자가 관리하는 컴퓨팅 자원을 여러 이용자가 서비스 형태로 공유하는 것이며, 이용자 컴퓨터끼리 직접 자원을 교환하는 Peer-to-Peer 구조를 뜻하지 않는다.

NoteBase

다시 나오면 이렇게 판단

‘필요할 때 네트워크로 IT 자원을 서비스받는다’면 클라우드의 특징으로 판단한다. ‘참여자 단말끼리 직접 자원을 주고받는다’면 Peer-to-Peer의 특징이므로 두 구조를 구분한다.

암기 포인트

클라우드는 제공자 자원의 주문형 서비스이고, Peer-to-Peer는 참여자 간 직접 자원 공유이다.

NoteBase

7. 도서관 디스커버리 서비스의 특징을 설명한 것으로 옳지 않은 것은?

① 구글이나 아마존 검색창과 같은 단일 검색창을 이용한 검색이 가능하다.

② 검색결과를 다양한 패싯(주제, 출판년 등)에 따라 재정렬 또는 내비게이션할 수 있다.

③ 메타데이터를 미리 수집하여 통합색인을 구축하지 않고 실시간으로 다양한 정보를 제공하는 메타서치서비스를 제공한다.

④ 시각적으로 풍부한 디스플레이, 검색어 추천, 연관 자원 추천 등의 기능을 가지고 있다.

⑤ 도서관 소장자료, 전자저널 논문, 오픈액세스 자원 등 다양한 정보원에 대해 통합색인을 구축하여 서비스한다.

정답 3번

출제 포인트

도서관 디스커버리 서비스의 핵심인 단일 검색창, 통합색인, 패싯 탐색, 풍부한 화면 구성과 추천 기능을 구별하는 문제이다.

핵심 개념

디스커버리 서비스는 도서관 소장자료, 전자저널 논문, 오픈액세스 자료 등 여러 정보원의 메타데이터를 미리 수집하여 하나의 통합색인을 만들고, 이용자가 단일 검색창에서 이를 한꺼번에 검색하도록 하는 서비스이다. 통합색인은 미리 모은 메타데이터를 하나의 색인에서 검색하는 방식이고, 메타검색은 검색할 때마다 여러 정보원에 질의를 보내 결과를 모으는 방식이다.

NoteBase

선지별 해설

① O
구글이나 아마존처럼 하나의 검색창에서 다양한 도서관 자원을 통합 검색할 수 있으므로 옳다.

② O
검색 결과를 주제, 출판년, 자료유형, 출판사 등의 패싯으로 좁히거나 이동할 수 있으므로 옳다.

③ X
틀린 표현은 '통합색인을 구축하지 않고 실시간으로 정보를 제공하는 메타서치서비스'이다. 디스커버리 서비스는 여러 정보원의 메타데이터를 미리 수집하여 통합색인을 구축한 뒤 이를 검색하는 방식이다. 실시간으로 여러 정보원에 각각 질의하는 메타검색 방식과 구별된다.

④ O
검색 결과를 시각적으로 풍부하게 제시하고, 검색어 추천이나 연관 자원 추천 등 이용자의 탐색을 돕는 기능을 제공하므로 옳다.

⑤ O
도서관 소장자료뿐 아니라 전자저널 논문과 오픈액세스 자원 등 이질적인 정보원의 메타데이터를 통합색인으로 구축하여 서비스하므로 옳다.

오답 포인트

③은 '여러 정보원을 한 번에 검색한다'는 공통점 때문에 디스커버리 서비스와 메타검색을 혼동하게 만든다. 판단의 핵심은 검색 시점의 실시간 분산 질의가 아니라, 메타데이터를 사전에 수집한 통합색인의 존재이다.

다시 나오면 이렇게 판단

단일 검색창, 사전 수집된 통합색인, 패싯을 통한 결과 정제, 추천·연관 자원 기능이 제시되면 디스커버리 서비스의 특징으로 판단한다. 반대로 검색할 때마다 여러 정보원에 실시간으로 질의한다고 하면 메타검색의 특징으로 판단한다.

암기 포인트

디스커버리는 '미리 모아 한 번에 찾는 통합색인', 메타검색은 '그때그때 여러 곳에 묻는 실시간 분산 검색'이다.

NoteBase

8. 적합성 피드백에 의한 질의확장에 대한 설명으로 옳지 않은 것은?

① 질의확장은 벡터공간, 확률검색 등에서 사용할 수 있다.
② 로치오(Rocchio) 공식을 사용한 질의수정에서는 적합문헌에 포함된 용어가 부적합문헌에 포함된 용어보다 효과적이다.
③ 질의확장 결과, 적합하다고 판정된 문헌들의 내용과 유사한 질의가 작성된다.
④ 전체 문헌집단을 대상으로 하여 추가 질의어를 선정한다.
⑤ 원칙적으로 적합문헌에 출현한 용어는 질의어로 추가하고 부적합문헌에 출현한 용어는 질의어에서 제외시킨다.

정답 4번

출제 포인트

적합성 피드백에 의한 질의확장이 어떤 문헌과 용어를 근거로 질의 벡터를 수정하는지를 묻는 문제이다.

핵심 개념

적합성 피드백은 초기 검색 결과에 대해 이용자가 적합·부적합을 판정하면, 그 판정 정보를 이용하여 기존 질의를 수정하는 방법이다. 질의확장에서는 적합문헌의 용어는 질의와의 관련성을 높이는 방향으로 반영하고, 부적합문헌의 용어는 관련성을 낮추는 방향으로 반영한다. 로치오 공식은 질의 벡터를 적합문헌 벡터의 중심 쪽으로 이동시키고 부적합문헌 벡터의 중심에서는 멀어지게 하는 대표적인 벡터공간 기반 방법이다.

선지별 해설

① O — 질의확장은 벡터공간모형의 로치오 방식뿐 아니라 용어의 적합확률을 반영하는 확률검색 방식에서도 사용할 수 있다.
② O — 로치오 질의수정에서는 적합문헌의 용어를 양의 방향으로, 부적합문헌의 용어를 음의 방향으로 반영하며 일반적으로 적합문헌에서 얻은 용어에 더 큰 긍정적 효과를 부여한다.
③ O — 수정된 질의는 적합하다고 판정된 문헌들의 대표 벡터 쪽으로 이동하므로, 그 문헌들의 내용과 유사한 방향의 질의가 작성된다.
④ X — 틀린 표현은 '전체 문헌집단을 대상으로'이다. 추가 질의어는 전체 문헌집단이 아니라 이용자가 적합 또는 부적합으로 판정한 검색 문헌의 용어와 가중치를 근거로 선정한다.
⑤ O — 적합문헌에 나타난 용어는 질의어의 가중치를 높이거나 새 질의어로 반영하고, 부적합문헌에 나타난 용어는 가중치를 낮추거나 제외하는 것이 적합성 피드백의 기본 방향이다.

NoteBase

오답 포인트

④의 '전체 문헌집단'은 질의확장의 후보 용어가 검색 시스템의 모든 문헌에서 직접 선정되는 것처럼 범위를 과도하게 넓힌 표현이다. 이 문제에서 피드백의 직접 근거는 이용자가 적합성과 부적합성을 판정한 검색 결과 문헌이다.

다시 나오면 이렇게 판단

적합성 피드백 문제에서는 '적합문헌 쪽으로 이동, 부적합문헌에서 멀어짐'을 기준으로 판단한다. 추가 질의어의 근거를 전체 문헌집단이라고 하면 틀린 설명으로 본다.

암기 포인트

로치오 질의수정은 적합문헌은 더하고, 부적합문헌은 빼며, 전체 문헌집단을 직접 피드백 대상으로 삼지 않는다.

NoteBase

9. DC의 기본 요소와 세분화 요소를 바르게 연결한 것은?

① Description - Medium
② Coverage - Access Rights
③ Rights - Extent
④ Date - Temporal
⑤ Identifier - Bibliographic Citation

정답 5번

출제 포인트

이 문제는 더블린 코어(DC)의 기본 요소와 그 의미를 더 구체화하는 세분화 요소의 연결 관계를 묻는다.

핵심 개념

기본 요소는 자원을 기술하는 상위 범주이고, 세분화 요소는 기본 요소의 의미를 더 구체적으로 한정한다.
Format은 자원의 파일 형식·물리적 형태와 관련되며 Medium과 Extent로 세분화된다.
Coverage는 자원의 공간적·시간적 범위를 나타내며 Spatial과 Temporal로 세분화된다.
Rights는 자원의 이용 권리와 관련되며 Access Rights로 세분화된다.
Identifier는 자원을 식별하는 참조값이며 Bibliographic Citation으로 세분화된다.

선지별 해설

① X — Medium은 Description이 아니라 Format의 세분화 요소이다. 올바른 연결은 Format - Medium이다.
② X — Access Rights는 Coverage가 아니라 Rights의 세분화 요소이다. 올바른 연결은 Rights - Access Rights이다.
③ X — Extent는 Rights가 아니라 Format의 세분화 요소이다. 올바른 연결은 Format - Extent이다.
④ X — Temporal은 Date가 아니라 Coverage의 세분화 요소이다. 올바른 연결은 Coverage - Temporal이다.
⑤ O — Bibliographic Citation은 자원을 식별하는 서지 인용 정보를 뜻하므로 Identifier의 세분화 요소이다.

오답 포인트

Medium과 Extent는 형식·물리적 특성을 나타내므로 Format에 속하고, Temporal은 시간이라는 말 때문에 Date와 혼동하기 쉽지만 자원의 시간적 범위를 나타내므로 Coverage에 속한다.

다시 나오면 이렇게 판단

세분화 요소가 무엇을 구체화하는지 의미로 판단한다. 매체·크기는 Format, 공간·시간 범위는 Coverage, 접근 권한은 Rights, 서지 인용을 통한 식별은 Identifier에 연결한다.

NoteBase

암기 포인트

Format-Medium · Extent, Coverage-Spatial · Temporal, Rights-Access Rights, Identifier-Bibliographic Citation으로 묶는다.

NoteBase

10. 신경망분류기의 성능평가 척도로 사용되는 분류재현율과 분류정확률을 구하는 공식으로 각각 옳은 것은?

	옳은 범주	틀린 범주
범주에 분류	a	b
범주에 비분류	c	d

	분류재현율	분류정확률
①	a/a+b	a/a+c
②	a/a+c	a/a+b
③	a+b/a+c	a+c/a+d
④	a+b/c+d	a+b/a+c
⑤	a/a+b+c+d	a/a+b

정답 2번

출제 포인트

혼동행렬에서 분류재현율과 분류정확률의 분모가 각각 실제로 옳은 범주의 전체 수와 해당 범주로 분류한 전체 수라는 점을 구분하는 문제이다.

조건 정리

a: 옳은 범주에 속하면서 그 범주로 올바르게 분류한 수
b: 틀린 범주에 속하지만 그 범주로 잘못 분류한 수
c: 옳은 범주에 속하지만 그 범주로 분류하지 못한 수
d: 틀린 범주에 속하면서 그 범주로 분류하지 않은 수
실제로 옳은 범주의 전체 수는 $a + c$이고, 해당 범주로 분류한 전체 수는 $a + b$이다.

사용할 공식

분류재현율은 실제로 옳은 범주에 속한 것 중 올바르게 찾아낸 비율이므로 $\frac{a}{a+c}$이다.
분류정확률은 해당 범주로 분류한 것 중 실제로 그 범주가 옳은 비율이므로 $\frac{a}{a+b}$이다.

NoteBase

계산 과정

1. 분류재현율의 분모에는 옳은 범주의 열 전체인 $a + c$를 놓고, 그중 올바르게 분류한 a를 분자로 놓는다.
따라서 분류재현율은 $\frac{a}{a+c}$이다.
2. 분류정확률의 분모에는 범주에 분류한 행 전체인 $a + b$를 놓고, 그중 올바르게 분류한 a를 분자로 놓는다.
따라서 분류정확률은 $\frac{a}{a+b}$이다.
3. 두 공식을 순서대로 제시한 선지는 ②이다.

선지 판단

① X — 분류재현율의 $\frac{a}{a+b}$와 분류정확률의 $\frac{a}{a+c}$는 두 분모를 서로 바꾼 식이다. 올바르게는 각각 $\frac{a}{a+c}$, $\frac{a}{a+b}$이다.
② O — 분류재현율이 $\frac{a}{a+c}$, 분류정확률이 $\frac{a}{a+b}$로 두 공식이 모두 옳다.
③ X — $\frac{a+b}{a+c}$와 $\frac{a+c}{a+d}$는 어느 것도 해당 정의에 맞지 않는다. 올바르게는 각각 $\frac{a}{a+c}$, $\frac{a}{a+b}$이다.
④ X — $\frac{a+b}{c+d}$와 $\frac{a+b}{a+c}$는 분자와 분모가 각 지표의 기준 집합을 나타내지 못한다. 올바르게는 각각 $\frac{a}{a+c}$, $\frac{a}{a+b}$이다.
⑤ X — 분류정확률 $\frac{a}{a+b}$는 옳지만, 분류재현율을 전체 표본 수 $a + b + c + d$로 나눈 $\frac{a}{a+b+c+d}$가 틀렸다.
분류재현율은 올바르게 $\frac{a}{a+c}$이다.

오답 포인트

핵심 함정은 a를 공통 분자로 사용하면서 재현율과 정확률의 분모를 뒤바꾸는 것이다. 재현율은 실제 옳은 범주를 기준으로 하므로 열 합계 $a + c$, 정확률은 범주에 분류한 결과를 기준으로 하므로 행 합계 $a + b$를 분모로 한다.

다시 나오면 이렇게 판단

먼저 지표가 실제 범주를 기준으로 묻는지, 분류 결과를 기준으로 묻는지 확인한다. 실제 범주 기준이면 재현율 $\frac{a}{a+c}$, 분류 결과 기준이면 정확률 $\frac{a}{a+b}$을 적용한다.

암기 포인트

재현율은 실제 옳은 범주 중 맞힌 비율 $\frac{a}{a+c}$, 정확률은 범주에 넣은 것 중 맞힌 비율 $\frac{a}{a+b}$이다.

NoteBase

11. 택소노미(Taxonomy)에 대한 설명으로 옳은 것은?

① 분류된 정보는 다른 이용자의 피드백에 의해 수정이 가능하다.
② 학문분류를 기반으로 하는 도서관의 전통적 분류법과 무관하다.
③ 소셜 북마크인 딜리셔스(Delicious)도 해당한다.
④ 전문가에 의한 주기적 갱신과 관리가 이루어진다.
⑤ 비계층적인 구조로 신생 주제의 삽입이 용이하다.

정답 4번

출제 포인트

택소노미와 폭소노미의 분류 주체, 계층 구조, 통제 방식 및 대표 사례를 구별하는 문제이다.

핵심 개념

택소노미(Taxonomy)는 전문가가 통제된 용어와 정해진 기준에 따라 개념을 계층적으로 조직하고 관리하는 분류 체계이다.
폭소노미(Folksonomy)는 다수 이용자가 자유롭게 태그를 붙여 만들어 가는 비계층적·사회적 분류 방식이다.
택소노미는 전문가·통제·계층 구조가 핵심이고, 폭소노미는 이용자·자유 태깅·비계층 구조가 핵심이다.

NoteBase

선지별 해설

① X
틀린 표현은 '다른 이용자의 피드백에 의해 수정이 가능하다'이다. 이용자들이 태그를 붙이고 상호작용하면서 분류를 형성·변화시키는 것은 폭소노미의 성격이다. 택소노미의 분류 정보는 정해진 기준과 통제된 절차에 따라 전문가가 관리한다고 고쳐야 한다.

② X
틀린 표현은 '도서관의 전통적 분류법과 무관하다'이다. 학문 분야를 기준으로 지식을 계층적으로 배열하는 전통적 도서관 분류법은 택소노미와 관련된 대표적인 분류 방식이다. 따라서 전통적 분류법과 관련이 있다고 고쳐야 한다.

③ X
틀린 표현은 '소셜 북마크인 딜리셔스(Delicious)도 해당한다'이다. 딜리셔스는 이용자가 자유롭게 태그를 부여하고 공유하는 폭소노미의 대표 사례이다. 따라서 딜리셔스는 택소노미가 아니라 폭소노미에 해당한다고 고쳐야 한다.

④ O
택소노미는 전문가가 분류 체계와 통제어를 설계하고, 변화한 지식과 용어를 반영하여 주기적으로 갱신·관리하는 방식이다.

⑤ X
틀린 표현은 '비계층적인 구조'이다. 택소노미는 상위 개념과 하위 개념의 관계가 드러나는 계층적 구조를 사용하므로 새 주제도 기존 분류 체계와의 관계를 정하여 편입해야 한다. 비계층적 구조에서 이용자가 새 태그를 쉽게 추가하는 것은 폭소노미의 성격이라고 고쳐야 한다.

오답 포인트

①·③·⑤는 이용자 참여, 소셜 태깅, 비계층 구조라는 폭소노미의 특징을 택소노미의 특징처럼 제시한 선지이다. 특히 ②의 '무관하다'처럼 관계를 전면 부정하는 표현도 주의해야 한다.

다시 나오면 이렇게 판단

분류의 주체와 구조를 먼저 확인한다. 전문가가 통제어를 사용하여 계층적으로 관리하면 택소노미이고, 이용자가 자유롭게 태그를 붙여 비계층적으로 형성하면 폭소노미이다.

암기 포인트

택소노미는 전문가·통제·계층, 폭소노미는 이용자·태그·비계층이다.

NoteBase

12. 동시인용 분석기법에 대한 설명으로 옳지 않은 것은?

① 동시인용 데이터를 인용하여 형성된 저자군집이나 문헌군집은 현재 진행되고 있는 연구자들의 연구경향을 나타낸다.

② 두 편의 문헌 또는 두 명의 저자가 제3의 문헌이나 저자에 의해 동시에 인용될 때 이 두 편의 문헌 혹은 저자는 밀접한 관계를 가지는 것으로 본다.

③ 동시인용법은 문헌이 발표된 후 함께 인용되는 초기 문헌들을 연결시키는 것으로 문헌들의 외부적 관계에 근거한 동적연결이라 할 수 있다.

④ 비슷하거나 보충적인 사실을 거듭 인용하는 경우, 동시인용빈도 수를 증가시킨다.

⑤ 동시인용 분석기법을 최초로 제안한 인물은 Small이다.

정답 1번

출제 포인트

동시인용의 형성 방식과 동적 성격을 이해하고, 이를 서지결합과 구별하는 문제이다.

핵심 개념

동시인용은 두 문헌이나 두 저자가 후속 문헌에 함께 인용되는 관계이다. 후속 연구가 새로 발표될 때마다 동시인용 빈도가 달라질 수 있으므로 동적인 관계이며, 축적된 학문 구조와 연구 전선의 변화를 파악하는 데 활용된다.

서지결합은 두 문헌이 동일한 선행 문헌을 함께 인용할 때 형성되는 관계이다. 두 문헌이 발표되는 시점에 참고문헌이 정해지므로 관계가 고정적이며, 발표 당시 연구자들의 현재적 연구 관심이나 경향을 파악하는 데 적합하다.

NoteBase

선지별 해설

① X — 동시인용으로 형성된 저자군집이나 문헌군집은 후속 문헌의 인용이 누적되어 드러나는 학문 분야의 지적 구조나 연구 전선을 나타낸다. 틀린 표현은 '현재 진행되고 있는 연구자들의 연구경향'이다. 올바르게는 현재적 연구경향은 발표 시점의 공통 참고문헌을 바탕으로 하는 서지결합이 더 잘 나타내며, 동시인용 군집은 누적된 인용 관계에 따른 지적 구조와 연구 동향을 나타낸다고 해야 한다.
② O — 두 문헌 또는 두 저자가 제3의 문헌에 함께 인용되면 동시인용 관계가 성립하며, 동시인용 빈도가 높을수록 두 대상의 주제적 관련성이 큰 것으로 본다.
③ O — 동시인용은 문헌 발표 후 후속 문헌이 두 문헌을 함께 인용하면서 형성된다. 후속 인용이 계속 추가될 수 있으므로 외부 인용에 근거한 동적 연결이다.
④ O — 내용이 비슷하거나 상호 보완적인 문헌은 후속 연구에서 함께 근거로 제시될 가능성이 크므로 동시인용 빈도가 증가한다.
⑤ O — 동시인용 분석은 Small이 최초로 제안한 기법으로 설명된다.

오답 포인트

①의 '현재 진행되고 있는 연구자들의 연구경향'은 서지결합의 성격과 혼동하게 만드는 표현이다. 동시인용은 과거에 발표된 문헌들이 이후 연구에서 함께 인용되면서 관계가 계속 변하고, 서지결합은 새 문헌들이 공통으로 인용한 선행 문헌을 기준으로 발표 시점에 관계가 정해진다.

다시 나오면 이렇게 판단

'후속 문헌이 두 문헌을 함께 인용한다'면 동시인용, '두 문헌이 같은 선행 문헌을 인용한다'면 서지결합으로 판단한다. 동시인용은 발표 후에도 빈도가 변하는 동적 관계이고, 서지결합은 발표 시점에 정해지는 고정적 관계이다.

암기 포인트

동시인용은 남이 둘을 함께 인용하는 동적 관계이고, 서지결합은 둘이 같은 문헌을 인용하는 고정적 관계이다.

NoteBase

13. MODS에 대한 설명으로 옳지 않은 것은?

① 디지털 자원을 기술하기 위한 표준 메타데이터 스키마이다.
② 문자 기반 태그를 사용하여 보다 이용자 편의적이며 사용이 편리하다.
③ 미국의회도서관(LC)에 의해 개발되었다.
④ MODS는 DC, MARC 등 다른 메타데이터 포맷을 보완하기 위해 제안되어, 구조적이고 복잡하며 세밀한 형식 필드로 구성된다.
⑤ 기존 도서관의 표준인 MARC와의 양방향 변환 및 상호운용성이 뛰어나다.

정답 4번

출제 포인트

MODS의 형식적 특징과 MARC·DC와의 관계를 구분하는 문제이다. 특히 MODS가 MARC보다 단순하면서 DC보다 풍부한 중간 수준의 메타데이터 스키마라는 점이 핵심이다.

핵심 개념

MODS(Metadata Object Description Schema)는 미국의회도서관이 개발한 XML 기반의 서지 메타데이터 스키마이다. MARC의 주요 의미 요소를 계승하면서 태그 이름을 문자로 표현해 이해하기 쉽게 만들었으며, 표현 수준은 단순한 DC보다 풍부하고 세밀하지만 MARC보다는 단순하다.

선지별 해설

① O — MODS는 디지털 자원을 포함한 정보자원의 서지적 특성을 기술하기 위한 표준 메타데이터 스키마이다.
② O — MODS는 XML의 문자 기반 요소명과 태그를 사용하므로 숫자 중심의 MARC 필드보다 의미를 파악하고 다루기 쉽다.
③ O — MODS는 미국의회도서관(Library of Congress, LC)이 개발·관리하는 스키마이다.
④ X — 틀린 표현은 '구조적이고 복잡하며 세밀한 형식 필드로 구성된다'이다. MODS는 DC보다 풍부한 구조를 제공하지만 MARC의 복잡성을 줄이기 위해 제안되었으므로, 올바르게는 'DC보다 상세하고 MARC보다 단순한 XML 기반 요소로 구성된다'이다.
⑤ O — MODS는 MARC의 주요 의미 요소를 바탕으로 설계되어 MARC와의 상호 변환을 지원하며, 기존 MARC 데이터와 연계하기에 유리하다.

오답 포인트

④는 'DC를 보완한다'는 앞부분 때문에 맞는 설명처럼 보이지만, MODS를 MARC처럼 복잡하고 세밀한 형식이라고 한 뒷부분이 틀렸다. MODS의 위치는 DC와 MARC의 중간이며, MARC의 복잡성을 완화한 스키마이다.

NoteBase

다시 나오면 이렇게 판단

MODS 관련 선지는 ‘XML 기반’, ‘미국의회도서관 개발’, ‘DC보다 풍부함’, ‘MARC보다 단순함’, ‘MARC와의 연계·변환 지원’의 다섯 기준으로 판단한다. MODS가 MARC보다 더 복잡하다고 하면 틀린 설명이다.

암기 포인트

MODS는 XML 기반이며, 상세성은 DC보다 높고 복잡성은 MARC보다 낮다.

NoteBase

14. Zipf의 제1법칙에 대한 설명으로 옳지 않은 것은?

① 문헌에 출현하는 단어의 빈도와 순위의 곱은 거의 일정하다는 법칙이다.
② 최소노력의 법칙으로도 불린다.
③ 문헌에 나타나는 단어들의 횟수를 순위로 매겨보면 순위가 내려갈수록 출현빈도가 기하급수적으로 줄어든다.
④ 저빈도 단어에는 정확하게 적용되지 않는 단점이 있다.
⑤ 한 문헌에 한 번만 출현한 단어의 수와 n번 출현한 단어 수의 비율은 일정하다.

정답 5번

출제 포인트

Zipf의 제1법칙이 설명하는 단어의 빈도와 순위의 관계를 이해하고, 이를 다른 빈도 법칙과 구별하는 문제이다.

핵심 개념

Zipf의 제1법칙은 문헌에 나타난 단어를 출현빈도가 높은 순서대로 배열할 때 단어의 빈도가 그 순위에 대체로 반비례한다는 경험 법칙이다. 따라서 빈도와 순위의 곱은 거의 일정하게 나타난다. 다만 실제 언어 자료에서는 특히 저빈도 영역에서 정확히 들어맞지 않을 수 있다.

선지별 해설

① O — 순위를 r, 출현빈도를 f라고 하면 제1법칙의 핵심은 f가 r에 대체로 반비례하여 빈도와 순위의 곱이 거의 일정하다는 것이다.
② O — Zipf는 언어 사용자가 들이는 노력과 전달 효과 사이의 균형이라는 관점에서 이 현상을 설명하므로 최소노력의 법칙이라고도 한다.
③ O — 여기서 '기하급수적으로 줄어든다'는 순위가 내려갈수록 출현빈도가 급격히 감소한다는 정성적 의미로 사용된 시험식 표현이다. 다만 수학적으로 엄밀한 관계는 지수적 감소가 아니라 출현빈도가 순위에 대체로 반비례하는 것이다.
④ O — Zipf의 법칙은 실제 자료에서 나타나는 근사적 경험 법칙이므로 모든 단어에 정확히 적용되지는 않으며, 저빈도 단어에서는 편차가 나타날 수 있다.
⑤ X — 틀린 표현은 '비율은 일정하다'이다. 한 번 출현한 단어의 수와 n번 출현한 단어의 수 사이의 비율은 n에 따라 달라지므로 일정하지 않다. 올바르게는 제1법칙에서 거의 일정한 것은 각 단어의 '출현빈도와 순위의 곱'이라고 해야 한다.

NoteBase

오답 포인트

⑤는 '한 번 출현한 단어의 수'와 'n번 출현한 단어의 수'를 끌어와 제1법칙의 일정한 곱 관계와 혼동하게 한다. 제1법칙의 두 변수는 단어의 출현빈도와 빈도순 순위이다.

다시 나오면 이렇게 판단

선지에서 '빈도와 순위'를 연결하면 두 값의 곱이 거의 일정한지 확인한다. 서로 다른 출현횟수를 가진 단어 종류 수의 비율을 무조건 일정하다고 하면 제1법칙의 관계가 아니다.

암기 포인트

Zipf 제1법칙은 빈도와 순위가 반비례하고, 빈도×순위가 거의 일정하다는 법칙이다.

NoteBase

15. 다음은 동일한 용어관계를 갖는 주제명표목표, 시소러스, 전거데이터베이스의 사례이다. ㉠, ㉡, ㉢에 들어갈 용어 또는 기호로 바르게 연결된 것은?

주제명표목표	시소러스	전거데이터베이스
국립중앙도서관 ㉠ 조선총독부 도서관 XX 국립도서관	국립중앙도서관 UF 조선총독부 도서관 RT 국립도서관	100 국립중앙 도서관 400 조선총독부 도서관 ㉢ 국립도서관
	조선총독부도서관 ㉡ 국립중앙 도서관	

	㉠	㉡	㉢
①	See	NT	500
②	See	USE	600
③	X	USE	500
④	X	NT	600
⑤	→	USE	500

정답 3번

출제 포인트

주제명표목표, 시소러스, 전거데이터베이스에서 우선어·비우선어 관계와 관련어 관계를 나타내는 기호 및 필드 번호를 서로 대응시키는 문제이다.

자료 읽는 순서

1. '국립중앙도서관'과 '조선총독부도서관'의 관계가 우선어와 비우선어의 관계임을 확인한다.
2. 주제명표목표의 X는 해당 표목 아래에 제시되는 비우선 표목을, 시소러스의 UF는 우선어에서 비우선어를 가리키는 관계를 나타낸다. 반대로 비우선어 표목에서는 USE로 우선어를 지시한다.
3. '국립도서관'은 관련어이므로 주제명표목표의 XX, 시소러스의 RT, 전거데이터베이스의 500 필드가 서로 대응한다.

NoteBase

핵심 근거

표의 첫 행에서 '조선총독부도서관'은 시소러스의 UF와 전거데이터베이스의 400에 대응하므로 ㉠은 X이다. 아래 행에서는 비우선어인 '조선총독부도서관'이 우선어인 '국립중앙도서관'을 지시하므로 ㉡은 USE이다. '국립도서관'은 시소러스에서 RT로 표시된 관련어이므로 전거데이터베이스에서는 500에 해당하여 ㉢은 500이다.

선지별 해설

① X — ㉠의 See와 ㉡의 NT가 틀렸다. 표에서 UF 및 400과 대응하는 ㉠은 X이고, 비우선어에서 우선어를 지시하는 ㉡은 USE이다. 올바른 조합은 X－USE－500이다.
② X — ㉡의 USE는 맞지만 ㉠의 See와 ㉢의 600이 틀렸다. 관련어 RT와 대응하는 전거데이터베이스 필드는 500이다. 올바른 조합은 X－USE－500이다.
③ O — ㉠ X는 UF·400과 같은 비우선어 관계를, ㉡ USE는 비우선어에서 우선어로의 지시를, ㉢ 500은 RT·XX와 같은 관련어 관계를 나타낸다. 세 빈칸이 모두 표의 관계와 일치한다.
④ X — ㉠ X는 맞지만 ㉡의 NT와 ㉢의 600이 틀렸다. ㉡은 USE이고, RT에 대응하는 ㉢은 500이다. 올바른 조합은 X－USE－500이다.
⑤ X — ㉡ USE와 ㉢ 500은 맞지만 ㉠의 화살표가 틀렸다. UF 및 400과 대응하여 '조선총독부도서관'을 비우선 표목으로 제시하는 기호는 X이다. 올바른 조합은 X－USE－500이다.

오답 포인트

UF와 USE의 방향을 혼동하기 쉽다. UF는 우선어 '국립중앙도서관' 아래에서 비우선어 '조선총독부도서관'을 제시하고, USE는 비우선어에서 사용할 우선어를 지시한다. 또한 RT는 상하위 관계인 NT가 아니라 관련어 관계이며 전거데이터베이스의 500에 대응한다.

다시 나오면 이렇게 판단

먼저 두 용어의 관계가 동의어 관계인지 관련어 관계인지 구분한다. 우선어에서 비우선어를 제시하는 관계는 X－UF－400으로, 비우선어에서 우선어를 지시하는 표시는 USE로, 관련어 관계는 XX－RT－500으로 연결한다.

암기 포인트

비우선어는 X－UF－400, 역방향 지시는 USE, 관련어는 XX－RT－500이다.

NoteBase

16. Cranfield Ⅰ 실험과 결과에 대한 설명으로 옳은 것은?

① 미국에서 실시된 정보검색의 성능평가에 대한 실험이다.
② 유니텀시스템이 최고의 재현율을 나타내었다.
③ 실험에 사용된 색인시스템은 DDC, 패싯분류법, 유니텀시스템이다.
④ 색인작성 시간이 증가하면 반드시 재현율이 증가한다.
⑤ 색인자의 경험과 주제지식은 성능 개선과 상당한 연관관계가 있다.

정답 2번

출제 포인트

Cranfield I 실험의 실시 장소, 비교 대상이 된 색인시스템, 재현율 결과, 색인작성 시간과 색인자의 개인적 특성이 검색 성능에 미친 영향을 묻는 문제이다.

핵심 개념

Cranfield I은 영국 Cranfield 항공대학에서 수행된 정보검색 실험으로, 서로 다른 색인시스템의 검색 성능을 비교한 연구이다.
비교 대상은 UDC, 알파벳순 주제명 색인, 패싯분류법, 유니텀시스템의 네 가지이며, 주된 비교 결과에서는 유니텀시스템의 재현율이 가장 높았다.
재현율은 전체 적합문헌 가운데 검색된 적합문헌의 비율이다.

NoteBase

선지별 해설

① X
틀린 표현은 '미국에서 실시된'이다. Cranfield I은 미국이 아니라 영국 Cranfield 항공대학에서 실시되었다. 다만 미국 국립과학재단의 지원을 받았다는 점 때문에 실시 국가를 미국으로 혼동할 수 있다.

② O
주된 비교 결과에서 유니텀시스템은 비교된 네 색인시스템 가운데 가장 높은 재현율을 나타냈다. 따라서 옳은 설명이다.

③ X
틀린 부분은 'DDC'와 세 시스템만 열거한 점이다. 실제 비교 대상은 DDC가 아니라 UDC였으며, UDC·알파벳순 주제명 색인·패싯분류법·유니텀시스템의 네 가지였다.

④ X
틀린 표현은 '반드시 재현율이 증가한다'이다. Cranfield I에서는 색인작성 시간을 늘린다고 재현율이 필연적으로 높아지지 않았고, 짧은 시간의 색인작성도 효율적인 결과를 보였다. 올바르게는 색인작성 시간의 증가와 재현율의 증가가 반드시 비례하는 것은 아니라고 해야 한다.

⑤ X
틀린 표현은 '상당한 연관관계가 있다'이다. 실험에서는 색인자의 색인 경험이나 주제지식과 검색 성능 사이에 유의미한 관련성이 확인되지 않았다. 올바르게는 이러한 개인적 특성이 성능 개선에 큰 영향을 미치지 않았다고 해야 한다.

오답 포인트

①은 미국 국립과학재단의 지원 사실을 실험 장소와 바꾸어 놓았고, ③은 UDC를 DDC로 바꾸면서 알파벳순 주제명 색인을 누락했다. ④의 '반드시'와 ⑤의 '상당한'처럼 결과를 단정하거나 과장하는 표현도 핵심 함정이다.

다시 나오면 이렇게 판단

Cranfield I의 기본 사실을 '영국에서 실시, 네 시스템 비교, UDC 사용, 유니텀의 높은 재현율'로 묶어 판단한다. 색인작성 시간이나 색인자의 경험·주제지식이 늘면 성능도 반드시 좋아진다는 선지는 실험 결과와 어긋나는지 확인한다.

암기 포인트

Cranfield I: 영국·UDC·주제명·패싯·유니텀, 재현율은 유니텀이 최고이다.

NoteBase

17. 지능형 IoT기반의 도서관 서비스를 구현하는데 활용되는 기술에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

<보 기>

ㄱ. QR코드: 무선주파수를 이용한 비접촉 인식 시스템으로 도서관에서 대출, 반납, 장서 점검 등에 활용될 수 있음

ㄴ. Beacon: 저전력 블루투스(BLE)를 통한 차세대 스마트폰 근거리통신 기술로 도서관 내 알림 및 좌석 지정 서비스 등으로 활용될 수 있음

ㄷ. NFC: 13.56MHz 주파수 대역을 사용하는 비접촉식 근거리 무선통신 모듈로 도서관에서 모바일 셀프 대출 서비스 등을 구현할 수 있음

ㄹ. O2O: 현실 환경에 가상 정보를 추가해 새로운 현실감을 제공하는 공간을 만들어내는 기술로 도서관에서 동화체험 서비스 등을 구현할 수 있음

① ㄱ, ㄴ
② ㄱ, ㄷ
③ ㄴ, ㄷ
④ ㄴ, ㄹ
⑤ ㄷ, ㄹ

정답 3번

출제 포인트

지능형 IoT 도서관 서비스에 활용되는 QR 코드, 비콘, NFC, O2O의 작동 방식과 적용 사례를 구별하는 문제이다.

핵심 개념

QR 코드: 카메라나 광학 스캐너로 읽는 2차원 바코드이며, 무선주파수를 이용하는 비접촉 인식 기술은 아니다.
비콘: BLE 신호를 주기적으로 송출하여 가까운 스마트폰의 위치 기반 알림이나 안내 서비스를 지원한다.
NFC: 13.56*MHz* 대역을 사용하는 비접촉식 근거리 무선통신 기술이다.
O2O: 온라인과 오프라인의 서비스·거래를 연결하는 방식이다. 현실 화면에 가상 정보를 겹쳐 보여 주는 기술은 증강현실(AR)이다.

NoteBase

선지별 해설

① X — ㄱ은 틀리고 ㄴ은 옳으므로 옳은 보기 전체가 아니다. QR 코드는 무선주파수 기반 인식 시스템이 아니라 카메라 등으로 읽는 광학식 2차원 바코드이다. 올바른 조합은 ㄴ, ㄷ이다.
② X — ㄱ은 틀리고 ㄷ은 옳다. ㄱ의 '무선주파수를 이용한 비접촉 인식 시스템'은 RFID에 해당하며, QR 코드는 광학적으로 인식한다. 올바른 조합은 ㄴ, ㄷ이다.
③ O — ㄴ과 ㄷ이 모두 옳다. 비콘은 BLE를 이용해 근거리 알림·위치 기반 서비스를 제공할 수 있고, NFC는 13.56*MHz* 대역의 근거리 무선통신으로 모바일 셀프 대출 등에 활용할 수 있다.
④ X — ㄴ은 옳지만 ㄹ은 틀리므로 옳은 보기 전체가 아니다. ㄹ이 설명한 '현실 환경에 가상 정보를 추가'하는 기술은 O2O가 아니라 증강현실(AR)이다. 올바른 조합은 ㄴ, ㄷ이다.
⑤ X — ㄷ은 옳지만 ㄹ은 틀리므로 옳은 보기 전체가 아니다. O2O는 온라인과 오프라인 서비스를 연결하는 방식이며, 현실에 가상 정보를 중첩하는 기술은 증강현실(AR)이다. 올바른 조합은 ㄴ, ㄷ이다.

오답 포인트

ㄱ에서는 '비접촉'이라는 말만 보고 QR 코드를 RFID와 혼동하기 쉽다. ㄹ에서는 디지털 기술이 현실 공간과 결합된다는 공통점 때문에 O2O와 증강현실을 혼동하기 쉽다. 인식 매체와 서비스 연결 방식을 기준으로 구별해야 한다.

다시 나오면 이렇게 판단

카메라로 코드를 읽으면 QR 코드, BLE 신호로 근거리 위치·알림을 제공하면 비콘, 13.56*MHz* 근거리 통신이면 NFC로 판단한다. 온라인과 오프라인의 연결은 O2O이고, 현실 화면에 가상 정보를 겹치는 것은 증강현실이다.

암기 포인트

QR은 광학 인식, 비콘은 BLE 알림, NFC는 13.56*MHz* 근거리 통신, O2O는 온·오프라인 연결, 가상 정보 중첩은 AR이다.

NoteBase

18. Edmunson이 제시한 텍스트 요약 기법에 해당하지 않는 것은?

① Key 방법
② Section 방법
③ Title 방법
④ Location 방법
⑤ Cue 방법

정답 2번

출제 포인트

Edmondson이 제시한 자동 텍스트 요약의 문장 선정 방법 네 가지를 구별하는 문제이다.

핵심 개념

Edmondson의 텍스트 요약 기법은 원문에서 요약에 포함할 문장의 중요도를 평가하는 방법이다. 대표적인 네 방법은 Key 방법, Cue 방법, Title 방법, Location 방법이다. Key 방법은 핵심어의 출현을, Cue 방법은 중요성을 드러내는 단서어를, Title 방법은 제목·표제어와의 관련성을, Location 방법은 문장의 위치를 이용한다.

선지별 해설

① O — Key 방법은 문서에서 자주 나타나는 내용어 등 핵심어를 포함한 문장을 중요하게 평가하는 방법이므로 해당한다.
② X — Section 방법은 Edmondson이 제시한 네 가지 방법에 포함되지 않는다. 틀린 표현은 'Section 방법'이며, 올바르게는 'Location 방법'이다.
③ O — Title 방법은 제목이나 표제에 사용된 단어와 관련된 문장에 높은 중요도를 부여하는 방법이므로 해당한다.
④ O — Location 방법은 문서나 문단의 처음·끝 등 문장이 놓인 위치를 이용해 중요도를 판단하는 방법이므로 해당한다.
⑤ O — Cue 방법은 '중요한', '결론적으로'처럼 문장의 중요성을 암시하는 단서어를 이용하는 방법이므로 해당한다.

오답 포인트

Section은 문서의 구획을 뜻하므로 요약 방법처럼 보이지만, Edmondson의 분류에서 위치 정보를 이용하는 방법의 명칭은 Section이 아니라 Location이다. 또한 부정형 질문이므로 네 가지 정식 방법이 아닌 하나를 골라야 한다.

NoteBase

다시 나오면 이렇게 판단

Edmondson의 네 방법을 '무슨 정보를 기준으로 문장을 고르는가'에 따라 연결한다. 핵심어는 Key, 단서어는 Cue, 제목은 Title, 문장 위치는 Location이며 이 네 명칭 밖의 보기는 제외한다.

암기 포인트

Edmondson의 네 방법은 Key·Cue·Title·Location이며, Section은 포함되지 않는다.

NoteBase

19. 다음의 관계형 데이터베이스에서 정보학개론의 저자를 검색하기 위한 SQL 구문으로 옳은 것은?

Table name : BIB

AU	TI	PU	YEAR
김대한	정보학개론	한국도서관협회	2010
김민국	디지털도서관 운영론	홍릉과학	2008
김국회	문헌정보학개론	태일사	2013
김도서	정보검색론	구미무역	2019

① select AU from BIB where TI = '정보학개론'
② retrieve AU from BIB where TI = '정보학개론'
③ select AU in BIB where TI = '정보학개론'
④ retrieve '김대한' where AU from BIB
⑤ select TI = '정보학개론' from BIB

정답 1번

출제 포인트

관계형 데이터베이스에서 특정 조건을 만족하는 행의 원하는 열을 조회하는 SELECT문의 기본 구문을 묻는 문제이다.

핵심 개념

SELECT문은 테이블에서 필요한 데이터를 조회하는 SQL 명령문이다. 기본 형식은 SELECT 조회할 열 FROM 테이블명 WHERE 조건이다. 이 문제에서는 제목(TI)이 '정보학개론'인 행에서 저자(AU)를 찾아야 하므로 select AU from BIB where TI = '정보학개론'이 된다.

NoteBase

선지별 해설

① O — SELECT 뒤에 조회할 열 AU, FROM 뒤에 테이블 BIB, WHERE 뒤에 조건 TI = '정보학개론'을 올바른 순서로 썼다.
② X — 틀린 표현은 retrieve AU이다. 표준 SQL의 조회 명령어는 retrieve가 아니라 SELECT이므로 select AU from BIB where TI = '정보학개론'으로 고쳐야 한다.
③ X — 틀린 표현은 in BIB이다. 조회 대상 테이블은 IN이 아니라 FROM으로 지정하므로 select AU from BIB where TI = '정보학개론'으로 고쳐야 한다.
④ X — retrieve, where AU from BIB의 키워드와 배열 순서가 잘못되었다. 김대한이라는 저자를 조건으로 검색하려면 select AU from BIB where AU = '김대한'과 같이 SELECT-FROM-WHERE 순서로 써야 한다. 이 문제처럼 정보학개론의 저자를 구하려면 select AU from BIB where TI = '정보학개론'이 맞다.
⑤ X — 틀린 표현은 select TI = '정보학개론'이다. SELECT에는 조회할 열을 쓰고 검색 조건은 WHERE에 써야 하므로 select AU from BIB where TI = '정보학개론'으로 고쳐야 한다.

오답 포인트

SELECT는 결과로 받을 열, FROM은 데이터를 찾을 테이블, WHERE는 행을 거르는 조건을 담당한다. ③은 FROM을 IN으로 바꾸고, ⑤는 조건식을 SELECT 자리에 넣어 이 역할을 뒤섞은 선지이다.

다시 나오면 이렇게 판단

‘무엇을 보여 줄 것인가 → 어느 테이블에서 찾을 것인가 → 어떤 조건으로 거를 것인가’의 순서로 SELECT 열, FROM 테이블, WHERE 조건을 대응한다.

암기 포인트

조회는 SELECT 열 FROM 테이블 WHERE 조건이다.

NoteBase

20. 다음은 Noy와 McGuinness가 개발한 온톨로지 구축 과정인 'Ontology Development 101' 방법론 7단계의 일부이다. ㉠에 들어갈 과업으로 옳은 것은?

3단계	온톨로지의 중요한 용어들을 수집하여 나열한다.
4단계	㉠
5단계	클래스의 속성인 슬롯을 정의한다.

① 온톨로지의 재사용 여부를 검토하고 결정한다.
② 온톨로지의 도메인과 범위를 결정한다.
③ 슬롯의 패싯을 정의한다.
④ 클래스와 클래스의 계층구조를 결정한다.
⑤ 인스턴스를 생성한다.

정답 4번

출제 포인트

Noy와 McGuinness의 'Ontology Development 101'에서 온톨로지를 구축하는 7단계의 순서와 각 단계의 과업을 구별하는 문제이다.

핵심 개념

1단계는 온톨로지의 도메인과 범위를 결정하고, 2단계는 기존 온톨로지의 재사용 가능성을 검토한다. 3단계는 중요한 용어를 열거하고, 4단계는 클래스와 클래스의 계층구조를 정의한다. 5단계는 클래스의 속성인 슬롯을 정의하고, 6단계는 슬롯의 패싯을 정의하며, 7단계는 인스턴스를 생성한다.

선지별 해설

① X — 틀린 표현은 '4단계에서 온톨로지의 재사용 여부를 검토하고 결정한다'이다. 기존 온톨로지의 재사용 가능성을 검토하는 과업은 2단계이다.
② X — 틀린 표현은 '4단계에서 온톨로지의 도메인과 범위를 결정한다'이다. 도메인과 범위를 결정하는 과업은 1단계이다.
③ X — 틀린 표현은 '4단계에서 슬롯의 패싯을 정의한다'이다. 슬롯의 패싯을 정의하는 과업은 6단계이다.
④ O — 4단계의 과업은 클래스를 정의하고 클래스의 계층구조를 결정하는 것이다.
⑤ X — 틀린 표현은 '4단계에서 인스턴스를 생성한다'이다. 인스턴스를 생성하는 과업은 마지막인 7단계이다.

NoteBase

오답 포인트

각 선지는 모두 온톨로지 구축 과정에 실제로 포함되는 과업이지만 단계 번호만 바꾸어 놓았다. 특히 클래스의 속성인 슬롯을 정의하는 5단계와 슬롯의 세부 제약인 패싯을 정의하는 6단계를 클래스 계층을 정하는 4단계와 혼동하기 쉽다.

다시 나오면 이렇게 판단

구축 대상을 정한 뒤 기존 자원을 검토하고, 용어를 모은 다음 클래스의 뼈대를 세운다고 판단한다. 그 후 클래스의 속성과 그 속성의 세부 제약을 정하고 마지막에 실제 인스턴스를 만든다.

암기 포인트

범위 → 재사용 → 용어 → 클래스·계층 → 슬롯 → 패싯 → 인스턴스의 순서이다.

NoteBase