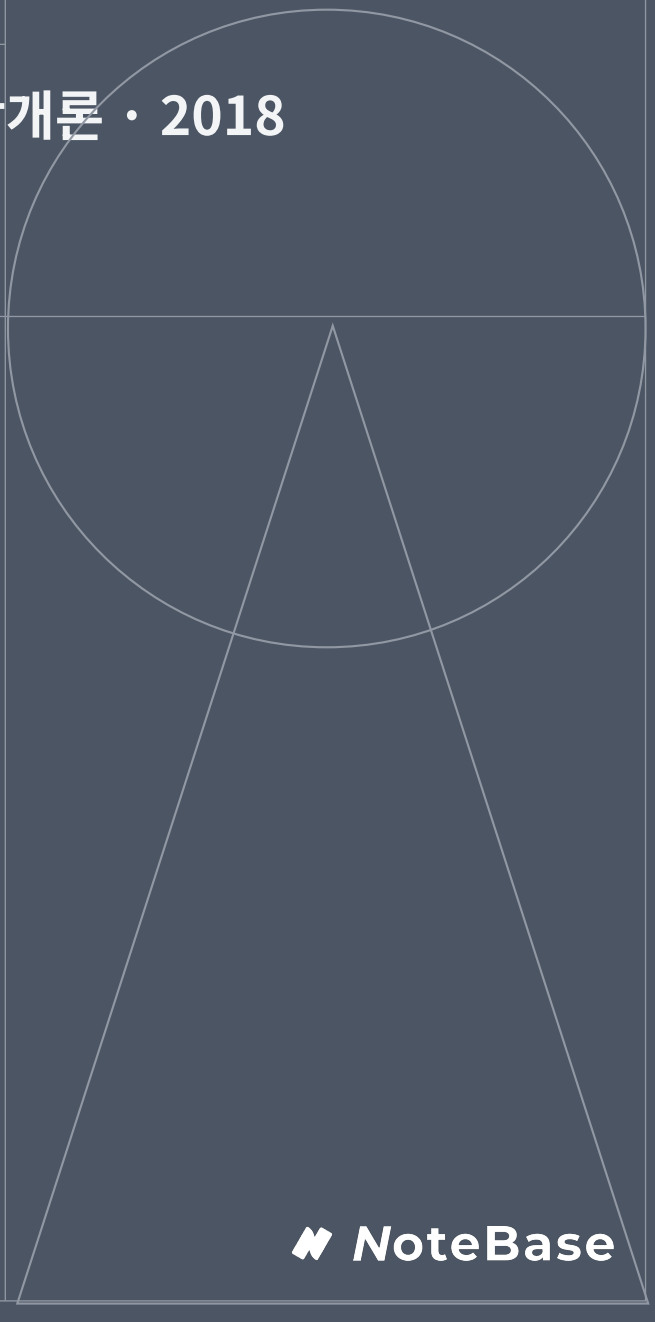
기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2018



NoteBase

기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2018

공무원/국회직/9급 · 정보학개론 · 2018

1. Shannon과 Weaver의 정보이론에서 언급된 불확실성과 엔트로피의 개념에 대한 설명으로 옳지 않은 것은?

① 어떤 사건이 일어날 확률이 작을 때 정보량은 더 크다고 할 수 있다.

② 정보는 엔트로피를 감소시키는 도구이다.

③ 불확실성의 정도를 나타내는 엔트로피가 커지면 불확실성이 높다.

④ 엔트로피는 정보이론에서 정보전달의 효율적 이행을 방해하는 요소를 의미한다.

⑤ 불확실성이 전혀 없을 때 실제 엔트로피는 최댓값인 1이다.

정답 5번

출제 포인트

이 문제는 Shannon과 Weaver의 정보이론에서 사건의 확률, 정보량, 불확실성, 엔트로피의 관계를 묻는다.

핵심 개념

정보량은 어떤 사건이 발생했다는 사실이 주는 놀라움의 크기이며, 발생 확률이 낮을수록 커진다. 엔트로피는 가능한 사건 중 무엇이 발생할지 모르는 평균적인 불확실성의 크기이다. 결과가 확실하면 엔트로피는 최솟값인 0이고, 가능한 결과들이 같은 확률로 나타날 때 엔트로피가 가장 커진다. 정보는 이러한 불확실성을 줄인다.

선지별 해설

① O — 발생 확률이 작은 사건일수록 실제로 발생했을 때 놀라움이 크므로 그 사건의 정보량도 크다.
② O — 정보를 얻으면 가능한 경우가 좁혀져 수신자의 불확실성, 즉 엔트로피가 감소한다.
③ O — 엔트로피는 불확실성의 크기를 나타내므로 엔트로피가 커질수록 어떤 결과가 나올지에 대한 불확실성도 높다.
④ O — 정보이론에서 엔트로피는 메시지 선택의 불확실성 또는 무질서의 정도를 나타내며, 그 값이 클수록 전달 과정에서 해소해야 할 불확실성이 커져 효율적인 정보전달을 어렵게 하는 요소로 볼 수 있다.
⑤ X — 틀린 표현은 '불확실성이 전혀 없을 때 실제 엔트로피는 최댓값인 1'이다. 불확실성이 전혀 없으면 결과가 확정되어 있으므로 엔트로피는 최댓값이 아니라 최솟값인 0이다.

NoteBase

오답 포인트

⑤는 불확실성과 엔트로피의 방향을 반대로 제시하고, 엔트로피의 최솟값 0을 최댓값 1로 바꾼 선지이다.
엔트로피 1은 두 결과가 같은 확률로 가능한 이진 상황에서의 최댓값일 수 있지만, 불확실성이 없는 상황의 값은 아니다.

다시 나오면 이렇게 판단

먼저 결과가 확정되어 있는지 확인한다. 결과가 확정되어 있으면 불확실성이 없으므로 엔트로피는 0이다. 반대로 가능한 결과의 확률이 고르게 분산될수록 예측이 어려워져 엔트로피가 커진다. 개별 사건의 정보량은 그 사건의 발생 확률과 반대로 움직인다.

암기 포인트

희귀한 사건은 정보량이 크고, 확실한 결과는 엔트로피가 0이다.

NoteBase

2. IFLA가 제안한 메타데이터 핵심요소 중 FRBR의 검색, 식별, 선택, 획득 기능에 모두 적용되는 요소는?

① 주제
② 날짜
③ 출판사
④ 자원명
⑤ 자원식별자

정답 4번

출제 포인트

IFLA가 제안한 핵심 메타데이터 요소를 FRBR의 네 가지 이용자 과업인 검색, 식별, 선택, 획득과 연결할 수 있는지를 묻는 문제이다.

핵심 개념

FRBR의 이용자 과업은 검색(조건에 맞는 자원을 찾는 것), 식별(찾은 자원이 원하는 자원인지 확인하는 것), 선택(필요에 적합한 자원을 고르는 것), 획득(자원에 접근하거나 입수하는 것)으로 구분된다.
자원명은 자원에 부여된 이름, 즉 표제에 해당한다. 자원명은 검색어와 식별 단서가 되고, 적합한 자원을 선택하는 기준이 되며, 최종적으로 획득할 대상을 지시하므로 네 기능에 모두 적용된다.

선지별 해설

① X — 주체는 여기서 자원의 주제(Subject)를 나타내는 요소로, 검색·식별·선택에는 적용되지만 획득에는 적용되지 않는다. 틀린 부분은 네 기능에 모두 적용된다고 보는 것이며, 올바르게는 획득을 제외한 세 기능에 적용된다.
② X — 날짜는 자원의 시점을 통해 식별·선택을 돕고 획득 대상의 판본 등을 구별하는 데 쓰이지만, 제시된 대응에서는 검색에 적용되지 않는다. 틀린 부분은 네 기능에 모두 적용된다고 보는 것이며, 올바르게는 식별·선택·획득에 적용된다.
③ X — 출판사는 발행 주체를 확인하여 자원을 식별·선택하고 입수처와 연결하는 데 쓰이지만, 제시된 대응에서는 검색에 적용되지 않는다. 틀린 부분은 네 기능에 모두 적용된다고 보는 것이며, 올바르게는 식별·선택·획득에 적용된다.
④ O — 자원명은 검색·식별·선택·획득의 네 기능에 모두 적용되는 핵심요소이다. 자원을 찾고, 구별하고, 적합성을 판단하고, 접근할 대상을 특정하는 전 과정의 기본 단서가 된다.
⑤ X — 자원식별자는 자원을 고유하게 식별하고 검색하며 획득 대상을 연결하는 데 쓰이지만, 제시된 대응에서는 선택에 적용되지 않는다. 틀린 부분은 네 기능에 모두 적용된다고 보는 것이며, 올바르게는 검색·식별·획득에 적용된다.

NoteBase

오답 포인트

자원식별자는 자원을 정확히 가리키므로 모든 기능에 쓰일 것처럼 보이지만, 이 문제의 기능 대응에서는 선택이 빠진다. 반면 자원명은 네 과업 모두에 표시되는 유일한 요소이다.

다시 나오면 이렇게 판단

각 요소가 실제로 유용한지를 막연히 판단하지 않고, IFLA의 기능 대응을 과업별로 대조한다. 네 칸이 모두 채워지는 요소는 자원명이며, 자원식별자는 선택, 날짜와 출판사는 검색, 주체는 획득이 각각 빠진다.

암기 포인트

자원명은 FRBR의 검색·식별·선택·획득을 모두 잇는 요소이다.

NoteBase

3. 다음은 디렉토리 분류체계에 대한 설명이다. ㉠~㉣에 들어갈 용어로 옳은 것은?

(㉠)은/는 (㉡)와/과 공유하는 특징이 많지만, (㉠)은/는 체계적인 관계를 나타낼 뿐이며 유전성에 바탕을 둔 관계는 아니라는 점에서 (㉡)와/과 차이가 있다. 각 단계 간에 특성을 공유하거나 유전성이 없기 때문에 (㉡)에 비해 표현력이 제한적이며, 개체의 특성을 추론하기 어렵다.
(㉢)은/는 하나의 개체를 독립적인 복수의 관점을 통해 나타낸다는 점에서 (㉣)와/과 유사성을 가지지만, 반드시 두 가지의 관점이 교차된 구조로 개체를 표현해야 한다는 점에서 차이가 있다.

	㉠	㉡	㉢	㉣
①	패싯분류	패러다임	계층구조	나무구조
②	계층구조	나무구조	패싯분류	패러다임
③	나무구조	계층구조	패러다임	패싯분류
④	패러다임	패싯분류	나무구조	계층구조
⑤	나무구조	패러다임	계층구조	패싯분류

정답 3번

출제 포인트

디렉터리 분류체계에서 나무구조와 계층구조, 패러다임과 패싯분류를 각각 구별하는 특징을 묻는 문제이다.

핵심 개념

나무구조: 요소 사이의 체계적인 관계를 나타내지만 단계 사이에 특성의 공유나 유전성이 없는 구조이다.
계층구조: 상하 단계가 특성을 공유하거나 유전하는 구조로, 나무구조보다 표현력이 크고 개체의 특성을 추론하기 쉽다.
패러다임: 하나의 개체를 독립적인 복수의 관점으로 나타내되, 반드시 두 관점이 교차하는 구조로 표현한다.
패싯분류: 하나의 개체를 서로 독립적인 여러 관점으로 나타내며, 반드시 두 관점의 교차 구조일 필요는 없다.
따라서 첫 문단은 ㄱ 나무구조·ㄴ 계층구조이고, 둘째 문단은 ㄷ 패러다임·ㄹ 패싯분류이다.

NoteBase

선지별 해설

① X — ㄱ을 '패싯분류', ㄴ을 '패러다임', ㄷ을 '계층구조', ㄹ을 '나무구조'로 배치한 부분이 모두 틀렸다. 올바르게는 ㄱ 나무구조, ㄴ 계층구조, ㄷ 패러다임, ㄹ 패싯분류이다.
② X — ㄱ을 '계층구조', ㄴ을 '나무구조', ㄷ을 '패싯분류', ㄹ을 '패러다임'으로 배치하여 각 쌍의 위치를 뒤바꾸었다. 올바르게는 ㄱ 나무구조, ㄴ 계층구조, ㄷ 패러다임, ㄹ 패싯분류이다.
③ O — ㄱ 나무구조, ㄴ 계층구조, ㄷ 패러다임, ㄹ 패싯분류로 모든 빈칸이 설명과 일치한다. 나무구조는 유전성이 없다는 점에서 계층구조와 구별되고, 패러다임은 두 관점의 교차가 필수라는 점에서 패싯분류와 구별된다.
④ X — ㄱ을 '패러다임', ㄴ을 '패싯분류', ㄷ을 '나무구조', ㄹ을 '계층구조'로 배치한 부분이 모두 틀렸다. 올바르게는 ㄱ 나무구조, ㄴ 계층구조, ㄷ 패러다임, ㄹ 패싯분류이다.
⑤ X — ㄱ '나무구조'와 ㄹ '패싯분류'는 맞지만, ㄴ '패러다임'과 ㄷ '계층구조'가 틀렸다. 올바르게는 ㄴ 계층구조, ㄷ 패러다임이다.

오답 포인트

'체계적인 관계'만 보고 ㄱ을 계층구조로 고르거나, '복수의 관점'만 보고 ㄷ을 패싯분류로 고르기 쉽다. 첫 문단에서는 '유전성이 아님', 둘째 문단에서는 '반드시 두 관점이 교차함'이 개념을 확정하는 표현이다.

다시 나오면 이렇게 판단

단계 사이의 특성 공유·유전성이 없으면 나무구조, 있으면 계층구조로 판단한다. 독립적인 복수 관점 가운데 반드시 두 관점의 교차를 요구하면 패러다임, 그러한 교차가 필수가 아니면 패싯분류로 판단한다.

암기 포인트

나무는 관계만, 계층은 유전까지, 패러다임은 두 관점의 교차, 패싯은 독립적인 여러 관점이다.

NoteBase

4. 디지털도서관 환경에서 통합검색방법에 대한 설명으로 옳지 않은 것은?

① 내·외부 장서의 메타데이터를 로컬 DB에 통합하여 검색하는 메타통합검색 방식으로 가능하다.
② 수확 방식은 OCLC의 종합목록과 같이 서지 데이터를 수집하는 경우에 많이 사용하는 방법이다.
③ 통합 메타데이터 구축 방법은 수확방식, 웹로봇 수집방식, 업로드 방식, 반입방식이 있다.
④ 통합검색은 메타데이터를 통합하지 않고도 구현할 수 있다.
⑤ 메타통합검색 방식은 동작이 단순하며, 검색행위가 하나의 디지털도서관에 집중된다는 장점이 있다.

정답 2번

출제 포인트

이 문제는 디지털도서관 통합검색의 구현 방식과 통합 메타데이터 구축 방법을 구분하는 문제이다.

핵심 개념

메타통합검색은 여러 디지털도서관의 메타데이터를 하나의 로컬 데이터베이스에 모아 검색하는 방식이다. 반면 분산통합검색은 메타데이터를 한곳에 통합하지 않고 각 기관의 검색 시스템에 질의를 보내 결과를 합친다.
통합 메타데이터의 구축 방법에는 수확 방식, 웹로봇 수집방식, 업로드 방식, 반입방식이 있다. 수확 방식은 표준 프로토콜 등을 이용해 분산된 저장소의 메타데이터를 자동으로 가져오는 방식이고, OCLC 종합목록처럼 참여기관이 서지 레코드를 중앙 시스템에 제출하는 경우에는 업로드 방식이 대표적이다.

선지별 해설

① O — 내·외부 장서의 메타데이터를 로컬 데이터베이스에 통합한 뒤 한 번에 검색하는 것은 메타통합검색 방식이다.
② X — 틀린 표현은 '수확 방식은 OCLC의 종합목록과 같이 서지 데이터를 수집하는 경우에 많이 사용한다'이다. OCLC 종합목록과 같이 참여기관이 서지 레코드를 중앙 시스템에 보내 구축하는 경우에는 업로드 방식이 대표적이다. 수확 방식은 분산된 저장소가 제공하는 메타데이터를 수확기가 자동으로 가져올 때 사용한다.
③ O — 통합 메타데이터는 수확 방식, 웹로봇 수집방식, 업로드 방식, 반입방식 등으로 구축할 수 있다.
④ O — 분산통합검색은 메타데이터를 중앙 데이터베이스에 미리 통합하지 않고도 각 기관에 질의를 보내 결과를 모으므로 통합검색을 구현할 수 있다.
⑤ O — 메타통합검색은 검색 시 여러 기관에 실시간으로 질의하지 않고 미리 구축한 하나의 통합 데이터베이스를 검색하므로 동작이 단순하고 검색 처리가 한 디지털도서관에 집중된다.

NoteBase

오답 포인트

②는 '수확'이라는 말 때문에 모든 서지 데이터 수집에 적용되는 것처럼 보이게 만든 선지이다. 자동으로 가져오는 수확 방식과 참여기관이 중앙 시스템에 레코드를 보내는 업로드 방식을 구분해야 한다.

다시 나오면 이렇게 판단

메타데이터가 중앙에 미리 모이면 메타통합검색, 중앙에 모으지 않고 각 기관에 질의하면 분산통합검색으로 판단한다. 구축 방식은 '자동으로 가져옴'이면 수확, '제공기관이 중앙에 보냄'이면 업로드로 구분한다.

암기 포인트

수확은 가져오기, 업로드는 보내기, 분산검색은 각 기관에 묻기이다.

NoteBase

5. <보기> 중 재현율과 정확률에 대한 설명으로 옳은 것을 모두 고르면?

<보 기>

ㄱ. 재현율과 정확률은 가장 기본적이면서 보편적인 검색성능 척도이다.

ㄴ. 재현율 50%는 검색된 문헌 10건 가운데 5건이 적합문헌임을 의미한다.

ㄷ. 재현율은 검색의 완전성을, 정확률은 검색의 정확성을 측정하는 척도라 할 수 있다.

ㄹ. 대체적으로 검색된 문헌 수가 증가할수록 재현율이 증가하게 된다.

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

정답 4번

출제 포인트

이 문제는 정보검색 성능의 기본 척도인 재현율과 정확률의 의미와 차이를 묻는 문제이다.

핵심 개념

재현율: 전체 적합문헌 가운데 실제로 검색된 적합문헌의 비율이다. 검색의 완전성, 즉 적합문헌을 얼마나 빠뜨리지 않았는지를 나타낸다.
정확률: 검색된 전체 문헌 가운데 적합문헌의 비율이다. 검색 결과의 정확성, 즉 검색된 문헌에 부적합문헌이 얼마나 적은지를 나타낸다.
재현율은 '전체 적합문헌'을 분모로 하고, 정확률은 '검색된 전체 문헌'을 분모로 한다.

NoteBase

선지별 해설

① X — ㄱ은 옳지만 ㄴ은 틀리므로 옳은 설명을 모두 고른 조합이 아니다. ㄴ의 '재현율 50%'는 정확률 50%로 고쳐야 하며, 올바른 조합은 ㄱ, ㄷ, ㄹ이다.
② X — ㄱ과 ㄹ은 옳지만 옳은 ㄷ이 빠져 있다. 'ㄱ, ㄹ'은 'ㄱ, ㄷ, ㄹ'로 고쳐야 한다.
③ X — ㄱ과 ㄹ은 옳지만 틀린 ㄴ이 포함되고 옳은 ㄷ이 빠져 있다. 'ㄱ, ㄴ, ㄹ'은 'ㄱ, ㄷ, ㄹ'로 고쳐야 한다.
④ O — ㄱ, ㄷ, ㄹ은 모두 옳다. 재현율과 정확률은 기본적인 검색성능 척도이고, 재현율은 완전성을, 정확률은 정확성을 측정하며, 일반적으로 검색 범위를 넓혀 검색 문헌 수가 증가하면 재현율도 증가한다.
⑤ X — ㄴ이 틀렸으므로 네 설명을 모두 고른 조합은 정답이 아니다. 검색된 문헌 10건 중 적합문헌이 5건이라는 조건은 정확률 50%를 뜻한다. 'ㄱ, ㄴ, ㄷ, ㄹ'은 'ㄱ, ㄷ, ㄹ'로 고쳐야 한다.

오답 포인트

ㄴ은 비율이 50%라는 숫자만 맞아 보이게 만든 함정이다. '검색된 문헌 10건 중 적합문헌 5건'은 검색 결과 안에서 적합문헌의 비율을 구한 것이므로 재현율이 아니라 정확률이다. 또한 이 문항은 옳은 설명을 모두 고르는 조합형이므로, 옳은 ㄷ을 빠뜨린 ②도 정답이 될 수 없다.

다시 나오면 이렇게 판단

분모를 먼저 확인한다. 전체 적합문헌이 기준이면 재현율이고, 검색된 전체 문헌이 기준이면 정확률이다.
조합형에서는 각 진술의 O/X를 먼저 확정한 뒤 옳은 진술이 모두 포함된 선지를 고른다.

암기 포인트

재현율은 적합문헌을 빠짐없이 찾는 완전성, 정확률은 검색 결과가 얼마나 알맞은지를 보는 정확성이다.

NoteBase

6. RDF에 대한 설명으로 옳지 않은 것은?

① 동일한 의미를 갖지만 메타데이터나 자원 형식에 따라 다르게 표현되는 데이터 요소들을 분리할 수 있다.

② 자원 검색, 자원 편목, 지능 소프트웨어 에이전트, 내용순위 부여 및 평가, 디지털 서명, 지적 재산권 보호 등의 분야에 응용될 수 있다.

③ RDF의 기술에 필요한 세 가지 요소는 자원, 속성 유형, 속성 값이다.

④ 메타데이터의 기술과 교환을 위한 구조로서, 웹상의 메타데이터를 지원하는 데 필요한 구조를 정의하기 위해 W3C에서 제안한 표준이다.

⑤ RDF를 이용하여 자원을 기술하려는 경우, 자원의 성격 및 특성에 따라 그 기술 내용을 자유롭게 확장할 수 있다.

정답 1번

출제 포인트

RDF의 목적, 기본 구성 요소, 활용 분야와 확장성을 정확히 이해하는지를 묻는 문제이다.

핵심 개념

RDF(Resource Description Framework)는 웹 자원에 관한 정보를 공통된 형식으로 표현하고 교환하기 위한 W3C 표준 데이터 모델이다. 하나의 문장은 주어·술어·목적어의 트리플로 표현되며, 이는 각각 자원·속성 유형·속성 값에 대응한다. 서로 다른 메타데이터 체계의 정보를 연결하고 통합할 수 있고, 필요한 어휘를 추가하여 자원 기술 내용을 확장할 수 있다.

NoteBase

선지별 해설

① X
틀린 표현은 '데이터 요소들을 분리할 수 있다'이다. RDF는 메타데이터나 자원 형식이 서로 다른 정보를 공통된 트리플 구조로 표현하여 교환·연결·통합하기 위한 모델이다. 올바르게는 '서로 다른 형식으로 표현된 데이터 요소들을 공통 모델로 표현하고 통합할 수 있다'이다.

② O
RDF는 자원 검색과 편목, 콘텐츠의 내용순위 부여 및 평가, 디지털 서명과 지적 재산권 정보 표현 등 다양한 자원 기술 분야에 활용될 수 있다. 구조화된 의미 정보를 기계가 처리할 수 있으므로 지능형 소프트웨어 에이전트에도 응용될 수 있다.

③ O
RDF 문장을 이루는 세 요소는 자원, 속성 유형, 속성 값이다. 트리플 용어로는 각각 주어, 술어, 목적어에 대응한다.

④ O
RDF는 웹 자원의 메타데이터를 기술하고 서로 교환할 수 있도록 W3C가 표준화한 구조이다. 특정 문서 형식 자체가 아니라 자원 정보를 표현하는 공통 데이터 모델이라는 점이 핵심이다.

⑤ O
RDF는 자원의 성격과 특성에 맞는 속성 어휘를 사용하거나 추가할 수 있으므로 자원 기술 내용을 유연하게 확장할 수 있다.

오답 포인트

①의 '동일한 의미'라는 표현에 이끌리면 옳은 설명처럼 보이지만, 함정은 '분리'에 있다. RDF가 지향하는 것은 형식별 분리가 아니라 공통 구조를 통한 표현, 교환, 연결과 통합이다.

다시 나오면 이렇게 판단

RDF 관련 선지는 '자원-속성 유형-속성 값'의 트리플 구조, W3C 표준, 메타데이터의 교환·통합, 어휘의 확장성을 기준으로 판단한다. 형식 차이에 따라 정보를 고립시키거나 분리한다는 설명은 RDF의 목적과 반대이다.

암기 포인트

RDF는 자원 정보를 트리플로 표현하여 분리하는 체계가 아니라 연결·교환·통합하는 체계이다.

NoteBase

7. <보기>의 내용은 XML로 인코딩한 MODS 레코드의 일부이다. ㉠, ㉡에 들어갈 MODS 상위요소로 옳은 것은?

<보 기>

```
<titleInfo>
    <title>Abraham Lincoln Research Site</title>
</titleInfo>
<typeOfResource>text</typeOfResource>
<(    ㉠    )>
    <form authority="marcform">electronic</form>
    <internetMediaType>text/html</internetMediaType>
    <digitalOrigin>born digital</digitalOrigin>
</(    ㉠    )>
<(    ㉡    )>
    <topic>Abraham Lincoln, abraham lincoln, assassination,
    springfield, honest abe, ann rutledge, booth, john
    wilkes booth, civil war</topic>
</(    ㉡    )>
```

	㉠	㉡
①	format	subject
②	format	classification
③	physicalDescription	subject
④	physicalDescription	classification
⑤	originInfo	note

정답 3번

출제 포인트

이 문제는 MODS 레코드에서 하위요소의 의미를 보고 알맞은 상위요소를 찾는 문제이다. `<form>`, `<internetMediaType>`, `<digitalOrigin>`은 `<physicalDescription>`에 속하고, `<topic>`은 `<subject>`에 속한다.

NoteBase

핵심 개념

physicalDescription: 자료의 물리적·디지털 형태에 관한 정보를 묶는 상위요소이다. 이 문항에서는 자료의 형태를 나타내는 `<form>`, 인터넷 매체 유형을 나타내는 `<internetMediaType>`, 디지털 자료의 생성 기원을 나타내는 `<digitalOrigin>`을 포함한다.
subject: 자료의 주제 접근점을 묶는 상위요소이다. 이 문항의 `<topic>`은 주제어를 나타내므로 `<subject>` 안에 들어간다.
따라서 ㉠은 physicalDescription이고 ㉡은 subject이다.

선지별 해설

① X — ㉡을 subject로 본 것은 맞지만, ㉠을 format으로 본 부분이 틀리다. 올바르게는 ㉠이 physicalDescription이고 ㉡이 subject이다.
② X — ㉠의 format과 ㉡의 classification이 모두 틀리다. 올바르게는 ㉠이 physicalDescription이고 ㉡이 subject이다.
③ O — physicalDescription은 `<form>`, `<internetMediaType>`, `<digitalOrigin>`을 묶고, subject는 `<topic>`을 포함하므로 두 상위요소가 모두 정확하다.
④ X — ㉠의 physicalDescription은 맞지만 ㉡의 classification이 틀리다. `<topic>`의 상위요소는 classification이 아니라 subject이다.
⑤ X — ㉠의 originInfo와 ㉡의 note가 모두 틀리다. 올바르게는 ㉠이 physicalDescription이고 ㉡이 subject이다.

오답 포인트

`<form>`이라는 하위요소 이름만 보고 상위요소도 format이라고 고르기 쉽다. 그러나 빈칸 하나가 세 하위요소를 함께 감싸므로 `<form>`만이 아니라 `<internetMediaType>`과 `<digitalOrigin>`까지 공통으로 포괄하는 physicalDescription을 찾아야 한다. 또한 `<topic>`은 분류기호가 아니라 주제어이므로 classification이 아닌 subject에 속한다.

다시 나오면 이렇게 판단

XML의 빈칸이 여러 태그를 감싸면 각 태그를 하나씩 보지 않고 공통 의미를 먼저 찾는다. 형태·매체유형·디지털 생성 기원은 physicalDescription, 주제어인 topic은 subject로 판단한다.

암기 포인트

form·internetMediaType·digitalOrigin → physicalDescription, topic → subject이다.

NoteBase

8. 정보검색시스템의 이용자 인터페이스 사용성 평가에 대한 설명으로 옳지 않은 것은?

① 시스템을 사용하기에 얼마나 편리한지를 평가하는 방법이다.
② 안구추적이나 생각말하기 등의 기법은 평가방법 중 감정법에 해당한다.
③ ISO 표준에서 정의된 사용성 평가의 요소는 이해가능성, 학습성, 운영성 등이 있다.
④ 평가방법은 일반적으로 질문법, 감정법, 검증법 등으로 구분된다.
⑤ 인간과 컴퓨터의 상호작용(HCI)을 구성하고 있는 핵심 원리이며, 부가적인 고려사항이 아닌 필수적으로 갖추어야 할 시스템의 핵심 요소로 받아들여지는 개념이다.

정답 2번

출제 포인트

사용성 평가의 목적과 평가 요소를 이해하고, 질문법·감정법·검증법의 구분 및 관찰 기법의 사례를 판별하는 문제이다.

핵심 개념

사용성 평가는 이용자가 시스템을 얼마나 효과적이고 효율적이며 만족스럽게 사용할 수 있는지를 확인하는 활동이다. 질문법은 이용자에게 질문하여 의견을 수집하고, 감정법은 전문가가 정해진 원칙이나 기준에 따라 인터페이스를 점검하며, 검증법은 실제 이용자가 과업을 수행하는 과정과 결과를 시험한다. 안구추적과 생각말하기는 이용자의 시선이나 사고 과정을 관찰하는 기법이다.

선지별 해설

① O — 사용성 평가는 이용자가 시스템을 얼마나 편리하게 사용할 수 있는지를 평가하는 방법이다. 편리성뿐 아니라 과업 수행의 효과성과 효율성, 만족도 등도 함께 살핀다.
② X — 틀린 표현은 '안구추적이나 생각말하기 등의 기법은 감정법에 해당한다'이다. 안구추적은 이용자의 시선 이동을, 생각말하기는 이용자가 과업을 수행하며 말하는 사고 과정을 관찰하므로 관찰법에 해당한다. 올바르게는 '안구추적이나 생각말하기 등의 기법은 관찰법에 해당한다'이다.
③ O — ISO의 사용성 관련 품질 특성에는 이해가능성, 학습성, 운영성 등이 포함된다. 이는 사용자가 시스템의 기능을 이해하고, 사용법을 익히며, 실제로 조작할 수 있는지를 나타내는 요소이다.
④ O — 사용성 평가방법은 일반적으로 질문법, 감정법, 검증법 등으로 구분할 수 있다. 각각 이용자 의견 수집, 전문가의 기준 기반 점검, 실제 이용자의 과업 수행 시험에 초점을 둔다.
⑤ O — HCI는 인간과 컴퓨터 사이의 상호작용을 연구하고 설계하는 핵심 영역이다. 사용자의 특성과 사용 맥락을 고려한 상호작용 설계는 시스템이 갖추어야 할 필수 요소로 본다.

NoteBase

오답 포인트

②는 기법의 이름보다 분류를 바꾸어 놓은 선지이다. ‘생각을 말한다’는 표현 때문에 질문법이나 감정법으로 오인하기 쉽지만, 평가자가 이용자의 행동과 발화를 지켜보며 자료를 얻으므로 관찰법이다.

다시 나오면 이렇게 판단

평가 자료를 누가 어떤 방식으로 얻는지 확인한다. 이용자에게 묻는 방식은 질문법, 전문가가 원칙에 따라 점검하는 방식은 감정법, 실제 과업 수행을 시험하는 방식은 검증법이며, 시선·행동·발화를 지켜보는 안구추적과 생각말하기는 관찰 기법으로 판단한다.

암기 포인트

안구추적·생각말하기는 감정법이 아니라 관찰법이다.

NoteBase

9. <보기>는 임의의 시소러스의 예이다. 이에 대한 설명으로 옳은 것을 모두 고르면?

<보 기>

정보탐색		정보활용	
UF	온라인 탐색	UF	정보응용
BT	학습활동	NT	정보검색
RT	개념형성		정보배포
	정보활용	BT	정보처리
	정보검색	RT	정보학

ㄱ. '온라인 탐색'과 '정보응용'은 비디스크립터이다.
ㄴ. '정보검색'은 '정보탐색'의 관련어이며, '정보활용'의 하위어이다.
ㄷ. '학습활동'은 '정보탐색'의 상위어이며, '정보배포'는 '정보활용'의 하위어이다.

① ㄴ
② ㄷ
③ ㄱ, ㄷ
④ ㄴ, ㄷ
⑤ ㄱ, ㄴ, ㄷ

정답 5번

출제 포인트

시소러스에서 UF, BT, NT, RT가 나타내는 용어 관계를 읽고 디스크립터와 비디스크립터, 상위어·하위어·관련어를 판별하는 문제이다.

핵심 개념

디스크립터는 색인과 검색에 사용하도록 채택된 우선어이고, 비디스크립터는 우선어로 안내되는 비우선어이다. UF(Used For)는 해당 디스크립터 대신 사용하지 않는 비디스크립터를 나타낸다. 따라서 '정보탐색'의 UF인 '온라인 탐색'과 '정보활용'의 UF인 '정보응용'은 비디스크립터이다.
BT(Broader Term)는 상위어, NT(Narrower Term)는 하위어, RT(Related Term)는 계층관계가 아닌 관련어를 뜻한다. 같은 관계는 반대쪽에서 서로 대응하므로 BT와 NT는 역관계이고 RT는 상호 관련관계이다.

NoteBase

선지별 해설

① X — ㄴ만 옳다고 한 범위가 잘못되었다. ㄱ과 ㄷ도 옳으므로 올바른 조합은 ㄱ, ㄴ, ㄷ이다.
② X — ㄷ만 옳다고 한 범위가 잘못되었다. ㄱ과 ㄴ도 옳으므로 올바른 조합은 ㄱ, ㄴ, ㄷ이다.
③ X — ㄱ과 ㄷ만 고른 것이 잘못되었다. '정보탐색' 항목에서 '정보검색'은 RT로 제시되므로 ㄴ도 옳다. 올바른 조합은 ㄱ, ㄴ, ㄷ이다.
④ X — ㄴ과 ㄷ만 고른 것이 잘못되었다. '온라인 탐색'과 '정보응용'은 각각 UF에 놓인 비디스크립터이므로 ㄱ도 옳다. 올바른 조합은 ㄱ, ㄴ, ㄷ이다.
⑤ O — ㄱ은 두 용어가 모두 UF에 제시되므로 옳다. ㄴ은 '정보검색'이 '정보탐색'의 RT이면서 '정보활용'의 NT이므로 옳다. ㄷ은 '학습활동'이 '정보탐색'의 BT이고 '정보배포'가 '정보활용'의 NT이므로 옳다. 따라서 ㄱ, ㄴ, ㄷ을 모두 고른 ⑤가 정답이다.

오답 포인트

RT와 NT를 한 용어에 동시에 붙일 수 없다고 생각하기 쉽다. 그러나 관계의 상대가 다르면 가능하다.
'정보검색'은 '정보탐색'과는 관련어(RT) 관계이고, '정보활용'에 대해서는 하위어(NT) 관계이다. 또한 UF에 제시된 말은 채택된 디스크립터가 아니라 비디스크립터라는 점이 핵심 함정이다.

다시 나오면 이렇게 판단

각 기호를 먼저 UF=비디스크립터, BT=상위어, NT=하위어, RT=관련어로 바꾼 뒤, 기호 바로 아래에 이어지는 모든 용어를 같은 관계로 읽는다. 그다음 ㄱ·ㄴ·ㄷ을 각각 독립적으로 판정하고 마지막에 옳은 문장의 조합과 선지를 대조한다.

암기 포인트

UF는 비우선어, BT는 위, NT는 아래, RT는 옆의 관련어이다.

NoteBase

10. <보기>는 정보환경을 구성하는 주요한 요소기술과 표준화를 연결한 것이다. 옳게 연결된 것을 모두 고르면?

<보 기>

ㄱ. 분산검색 - Z30.50
ㄴ. 메타데이터 수확 - OAIS
ㄷ. 뉴스와 공지 - RSS
ㄹ. 문맥기반의 연계 - OpenURL
ㅁ. 처리형 서비스 - SOAP

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

정답 4번

출제 포인트

정보환경의 주요 요소기술과 이를 구현하는 표준의 정확한 연결을 묻는 문제이다. 표준명의 숫자와 약어가 비슷한 기술의 용도를 구별하는 것이 핵심이다.

핵심 개념

분산검색은 여러 원격 정보검색 시스템을 연계해 검색하는 기술이며 대표 표준은 Z39.50이다. 따라서 ㄱ의 Z30.50은 잘못된 표기이다.
메타데이터 수확은 여러 저장소의 메타데이터를 모으는 기술이며 OAI-PMH를 사용한다. OAIS는 디지털 정보의 장기보존을 위한 참조모형이므로 ㄴ과 연결되지 않는다.
RSS는 뉴스와 공지 등의 갱신 정보를 배포하는 형식이다. OpenURL은 문헌의 메타데이터를 링크 해석기로 전달하여 이용 가능한 원문·관련 서비스로 연결하는 표준이다. SOAP은 구조화된 메시지를 교환하여 웹 기반 처리형 서비스를 연계하는 프로토콜이다. 따라서 ㄷ, ㄹ, ㅁ이 옳게 연결된다.

NoteBase

선지별 해설

① X — ㄷ(RSS)는 옳지만 ㄱ의 Z30.50은 잘못된 표기이고, 옳은 ㄹ과 ㅁ도 빠져 있다. 틀린 표현인 '분산검색-Z30.50'은 '분산검색-Z39.50'으로 고치며, 옳게 연결된 항목 전체는 ㄷ, ㄹ, ㅁ이다.
② X — ㄱ의 Z30.50과 ㄴ의 OAIS가 모두 잘못 연결되었다. 올바르게는 '분산검색-Z39.50', '메타데이터 수확-OAI-PMH'이며, 옳게 연결된 항목 전체는 ㄷ, ㄹ, ㅁ이다.
③ X — ㄷ과 ㄹ은 옳지만 ㄴ의 OAIS는 메타데이터 수확 표준이 아니다. 틀린 표현인 '메타데이터 수확-OAIS'는 '메타데이터 수확-OAI-PMH'로 고치며, 옳은 조합에는 ㅁ도 포함되어 ㄷ, ㄹ, ㅁ이 된다.
④ O — ㄷ의 뉴스와 공지-RSS, ㄹ의 문맥기반 연계-OpenURL, ㅁ의 처리형 서비스-SOAP은 모두 각 기술의 기능과 표준이 정확히 연결되어 있다.
⑤ X — ㄷ, ㄹ, ㅁ은 옳지만 ㄱ의 Z30.50이 잘못되었다. 틀린 표현인 '분산검색-Z30.50'은 '분산검색-Z39.50'으로 고쳐야 하므로 옳은 조합은 ㄱ을 제외한 ④의 ㄷ, ㄹ, ㅁ이다.

오답 포인트

ㄱ의 Z30.50을 익숙한 정보검색 표준인 Z39.50으로 잘못 읽기 쉽고, ㄴ의 OAIS를 OAI-PMH와 혼동하기 쉽다. 숫자 하나의 차이와 약어 뒤의 PMH가 각각 정오를 가른다.

다시 나오면 이렇게 판단

기능을 먼저 기준으로 삼아 '검색-Z39.50, 수확-OAI-PMH, 배포-RSS, 문맥 링크-OpenURL, 처리 메시지-SOAP'의 순서로 대응시킨다. OAIS가 나오면 메타데이터 수확이 아니라 장기보존 참조모형인지 확인한다.

암기 포인트

검색은 Z39.50, 수확은 OAI-PMH, 소식은 RSS, 문맥 링크는 OpenURL, 처리는 SOAP이다. OAIS는 장기보존이다.

NoteBase

11. 문헌정보학 분야의 학자와 업적에 대한 설명으로 옳지 않은 것은?

① Gorman : 미래도서관 및 사서직에 대한 역할 정의, 신도서관학 5법칙 주장

② Dewey : 십진분류법 개발, 도서관학 교육기관 창설에 기여

③ Lancaster : 도서관·정보센터 평가에 중요한 역할, 정보검색시스템의 개념적 구조화에 일조

④ Butler : 도서관의 사회적 기능과 사서의 역사의식 및 전문성 강조, 도서관과 사회인식론 주장

⑤ Panizzi : 대영박물관 사서, 근대적 목록규칙의 기초 수립

정답 4번

출제 포인트

문헌정보학의 주요 인물과 대표 업적·이론을 정확히 연결하는 문제이다. 특히 Butler의 사회학적 도서관학과 Shera·Egan의 사회인식론을 구별해야 한다.

핵심 정리

Gorman: 미래 도서관과 사서의 역할을 논하고 신도서관학 5법칙을 제시하였다.
Dewey: 십진분류법을 개발하고 근대적인 도서관학 교육기관의 설립에 기여하였다.
Lancaster: 도서관·정보센터와 정보검색시스템의 평가 및 설계를 체계화하였다.
Butler: 도서관을 사회적 기관으로 보고 도서관학에 사회학적 관점을 도입하였다.
Panizzi: 대영박물관 도서관의 사서로서 91개 목록규칙을 마련하여 근대 목록법의 기초를 세웠다.
사회인식론: 지식의 생산·유통·이용을 사회적 과정으로 파악하는 관점으로, Shera와 Egan의 업적과 연결된다.

선지별 해설

① O — Gorman은 미래 도서관과 사서직의 역할을 정리하고 Ranganathan의 법칙을 현대적으로 재구성한 신도서관학 5법칙을 제시하였다.
② O — Dewey는 십진분류법을 개발했으며, 도서관학 전문교육의 제도화와 도서관학교 설립에 기여하였다.
③ O — Lancaster는 도서관·정보센터의 서비스와 정보검색시스템을 평가하는 방법을 발전시키고 정보검색시스템의 개념적 구조화에 기여하였다.
④ X — 틀린 부분은 Butler가 '사회인식론'을 주장했다는 연결이다. Butler는 도서관의 사회적 기능을 중시하고 도서관학에 사회학적 관점을 도입한 인물이며, 사회인식론은 올바르게는 Shera와 Egan의 업적과 연결된다.
⑤ O — Panizzi는 대영박물관 도서관의 사서로 활동하면서 91개 목록규칙을 작성하여 근대 목록규칙의 기초를 마련하였다.

NoteBase

출제자가 노린 함정

④는 Butler가 도서관을 사회적 기관으로 보았다는 사실에 '사회인식론'이라는 유사한 표현을 덧붙여 그럴듯하게 만든 선지이다. 사회학적 접근의 선구자인 Butler와 사회인식론을 전개한 Shera·Egan을 혼동하면 오답을 고르기 쉽다.

다시 나오면 이렇게 판단

인물명에서 대표 키워드를 먼저 대응시킨다. Butler는 도서관학의 사회학적 접근, Shera·Egan은 사회인식론, Dewey는 십진분류법, Panizzi는 91개 목록규칙으로 구별한다. 한 선지에 여러 업적이 함께 제시되면 마지막 개념까지 같은 인물의 업적인지 확인한다.

암기 포인트

Butler는 사회학적 접근, Shera·Egan은 사회인식론이다. Dewey는 십진분류, Panizzi는 91개 목록규칙, Gorman은 신도서관학 5법칙이다.

NoteBase

12. 더블린 코어의 기본 데이터 요소 중 '포맷(Format)'에 해당되는 세분화 요소로 옳게 나열한 것은?

① Issued, Modified
② Image, Dataset
③ Extent, Medium
④ Spatial, Temporal
⑤ Access Rights, Licence

정답 3번

출제 포인트

더블린 코어의 기본 데이터 요소와 각 요소를 구체화하는 세분화 요소의 대응 관계를 묻는 문제이다.

핵심 개념

Format은 자원의 파일 형식, 물리적 매체, 규모 등을 나타내는 기본 요소이며, Extent와 Medium이 이를 세분화한다. Extent는 자원의 크기나 분량을, Medium은 자원을 담거나 구현한 물리적 매체를 나타낸다. 주요 대응은 Date－Issued·Modified, Type－Image·Dataset, Format－Extent·Medium, Coverage－Spatial·Temporal, Rights－Access Rights·Licence이다.

선지별 해설

① X — Issued와 Modified는 발행일과 수정일을 나타내므로 Date의 세분화 요소이다. 틀린 부분은 Format에 대응시킨 것이며, 올바르게는 'Date－Issued, Modified'이다.
② X — Image와 Dataset은 자원의 성격이나 장르를 나타내는 Type의 유형이다. 틀린 부분은 Format에 대응시킨 것이며, 올바르게는 'Type－Image, Dataset'이다.
③ O — Extent는 자원의 크기·분량을, Medium은 물리적 매체를 나타내므로 둘 다 Format을 세분화하는 요소이다.
④ X — Spatial과 Temporal은 자원의 공간적·시간적 범위를 나타내므로 Coverage의 세분화 요소이다. 틀린 부분은 Format에 대응시킨 것이며, 올바르게는 'Coverage－Spatial, Temporal'이다.
⑤ X — Access Rights와 Licence는 자원의 접근 및 이용 권한을 나타내므로 Rights의 세분화 요소이다. 틀린 부분은 Format에 대응시킨 것이며, 올바르게는 'Rights－Access Rights, Licence'이다.

오답 포인트

선지의 용어가 모두 실제 더블린 코어 관련 용어이므로, 용어의 존재 여부만으로 판단하면 혼동하기 쉽다. 각 용어가 무엇을 설명하는지를 기준으로 상위 기본 요소를 연결해야 한다.

NoteBase

다시 나오면 이렇게 판단

크기·분량 또는 물리적 매체를 설명하면 Format, 날짜 상태를 설명하면 Date, 자원 유형을 설명하면 Type, 공간·시간 범위를 설명하면 Coverage, 접근·이용 권한을 설명하면 Rights로 판단한다.

암기 포인트

Format은 '얼마나 큰가(Extent), 무엇에 담겼는가(Medium)'로 기억한다.

NoteBase

13. <보기>에서 설명하는 것은?

<보 기>

ㄱ. 시스템 간의 복합디지털 객체의 전송 및 아카이빙을 위한 디지털 자원에 대한 허브 문서의 인코딩 규칙을 규정하는 XML 스키마 기반의 명세

ㄴ. 미국 의회도서관에서 개발

ㄷ. 오디오 또는 비디오 형태와 같은 시간에 기반을 둔 디지털 자원의 구조에 대한 기술과 외부자원으로서 연결 기능 등 기술적, 관리적 메타데이터에 더 많은 확장성을 제공하기 위하여 개발

① MODS
② METS
③ OAIS
④ Topic Maps
⑤ FRBR

정답 2번

출제 포인트

이 문제는 XML 기반 메타데이터 표준들의 목적과 기능을 구별하여, 디지털 객체의 전송·보존을 위해 구조·기술·관리 메타데이터를 묶는 METS를 찾는 문제이다.

핵심 개념

METS(Metadata Encoding and Transmission Standard)는 디지털 객체의 전송과 아카이빙을 위해 기술·관리·구조 메타데이터 등을 XML 문서로 표현하고 서로 연결하는 표준이다. 미국 의회도서관이 제공·관리하며, 외부 메타데이터를 가리키거나 문서 안에 포함할 수 있어 오디오·비디오 같은 복합 디지털 자원의 구조도 기술할 수 있다.

MODS는 서지정보를 기술하는 XML 기반 서지 메타데이터 스키마이고, OAIS는 디지털 장기보존 시스템을 설명하는 참조모형이다. Topic Maps는 주제와 주제 사이의 관계를 표현하는 지식조직 표준이며, FRBR은 서지적 개체와 그 관계를 설명하는 개념모형이다.

NoteBase

선지별 해설

① X — MODS는 MARC 계열의 서지요소를 XML로 표현하는 서지 메타데이터 스키마이다. 보기처럼 디지털 객체의 구조·기술·관리 메타데이터를 포괄하여 전송·보존하는 표준이라는 표현은 METS로 고쳐야 한다.
② O — METS는 XML 스키마를 바탕으로 디지털 객체의 구조를 기술하고 관련 메타데이터를 내부에 포함하거나 외부 자원으로 연결할 수 있으므로 ㄱ, ㄴ, ㄷ의 설명에 부합한다.
③ X — OAIS는 보존정보패키지, 기능 구성 등 디지털 아카이브의 장기보존 체계를 설명하는 참조모형이다. 보기의 XML 기반 메타데이터 인코딩·전송 표준은 METS로 고쳐야 한다.
④ X — Topic Maps는 주제, 주제 간 연관, 주제와 정보자원의 연결을 표현하는 지식조직 표준이다. 보기에서 설명하는 복합 디지털 객체의 구조·메타데이터 포장 표준은 METS로 고쳐야 한다.
⑤ X — FRBR은 저작·표현형·구현형·개별자료 등의 서지적 개체와 관계를 구조화한 개념모형이다. 보기의 디지털 객체 전송·아카이빙용 XML 명세는 METS로 고쳐야 한다.

오답 포인트

METS와 MODS가 모두 XML 기반이고 미국 의회도서관과 관련되어 있어 혼동하기 쉽다. 그러나 보기의 핵심 표현은 '복합 디지털 객체의 전송 및 아카이빙', '구조', '기술·관리 메타데이터', '외부자원 연결'이므로 서지기술 중심의 MODS가 아니라 METS이다.

다시 나오면 이렇게 판단

복합 디지털 객체를 하나의 XML 문서로 포장하면서 구조를 기술하고 여러 종류의 메타데이터를 연결하면 METS로 판단한다. 서지정보 자체의 기술이면 MODS, 장기보존 시스템의 기능모형이면 OAIS, 주제 관계 표현이면 Topic Maps, 서지적 개체 관계모형이면 FRBR로 구별한다.

암기 포인트

METS는 메타데이터의 포장·전송·구조, MODS는 서지기술, OAIS는 장기보존 참조모형이다.

NoteBase

14. 벡터공간모델에 대한 설명으로 옳지 않은 것은?

① 이용자는 벡터공간 정보검색시스템에서 탐색을 수행할 때 불리언 논리를 이해하거나 사용할 필요가 없다.

② 벡터공간모델에서는 질의 또는 문헌을 표현하기 위해 선택된 용어 또는 개념에 대해 상대적 중요성을 나타낼 수 있는 가중치가 부여될 수 있다.

③ 벡터공간모델에서 선택된 용어들 사이에는 서로 밀접한 관련이 있다고 전제한다.

④ 적합성 피드백은 검색성능을 개선하기 위해 벡터공간모델에서 구현된 메커니즘이다.

⑤ 벡터공간모델을 이용하여 정보검색의 결과를 순위화할 수 있다.

정답 3번

출제 포인트

벡터공간모델의 핵심 가정과 검색 방식, 가중치·순위화·적합성 피드백의 기능을 구별하는 문제이다.

핵심 개념

벡터공간모델은 문헌과 질의를 용어별 가중치로 이루어진 벡터로 표현하고, 두 벡터의 유사도를 계산하여 검색 결과의 관련성 순위를 정하는 정보검색 모델이다. 기본적으로 각 용어는 서로 독립적인 차원으로 취급되므로, 선택된 용어들이 서로 밀접하게 관련되어 있다고 전제하지 않는다.

선지별 해설

① O — 이용자는 자연어 질의를 용어 벡터로 표현하여 검색할 수 있으므로 불리언 연산자를 직접 이해하거나 사용할 필요가 없다.
② O — 질의나 문헌을 구성하는 각 용어에 가중치를 부여하여 그 용어의 상대적 중요성을 나타낼 수 있다.
③ X — 틀린 표현은 '선택된 용어들 사이에는 서로 밀접한 관련이 있다고 전제한다'이다. 올바르게는 벡터공간모델은 각 용어를 서로 독립적인 차원으로 취급한다고 본다.
④ O — 적합성 피드백은 이용자가 적합하다고 판단한 문헌의 용어 정보를 반영해 질의 벡터를 조정하고 검색 성능을 높이는 메커니즘이다.
⑤ O — 질의 벡터와 문헌 벡터 사이의 유사도를 계산할 수 있으므로 검색 결과를 유사도가 높은 순서로 순위화할 수 있다.

오답 포인트

③은 실제 의미 관계와 수학적 모델의 가정을 혼동하게 한다. 현실의 용어들은 의미상 관련될 수 있지만, 기본 벡터공간모델에서는 각 용어를 독립적인 축으로 표현한다.

NoteBase

다시 나오면 이렇게 판단

벡터공간모델은 '용어별 가중치로 벡터화 → 질의와 문헌의 유사도 계산 → 결과 순위화'의 흐름으로 판단한다. 용어 간 밀접한 관련성을 기본 전제로 삼는다는 설명은 벡터공간모델의 독립 차원 가정과 어긋난다.

암기 포인트

벡터공간모델은 용어 독립, 가중치 부여, 유사도 계산, 순위화가 핵심이다.

NoteBase

15. 확장불리언 검색모형으로 옳은 것은?

① 추론망 모델, 퍼지탐색
② MMM 모델, P-norm 모델
③ 베이지언 네트워크, 추론망 모델
④ 신경망 모델, 신념망 모델
⑤ 퍼지집합 모델, 잠재의미색인 모델

정답 2번

출제 포인트

확장 불리언 검색모형의 대표 유형인 MMM 모델과 P-norm 모델을 다른 정보검색 모형과 구별하는 문제이다.

핵심 개념

확장 불리언 검색모형은 불리언 모델의 AND·OR 질의를 유지하면서 문서와 질의의 일치 정도를 연속적인 값으로 산출하여 부분 일치와 검색 결과의 순위화를 가능하게 한 모형이다.
MMM 모델은 Mixed Min and Max 모델로, min과 max 연산을 혼합하여 불리언 연산의 엄격성을 완화한다.
P-norm 모델은 p-노름을 이용해 질의와 문서 사이의 일치 정도를 계산하며, p값에 따라 불리언 연산의 엄격한 정도를 조절한다.

선지별 해설

① X — 추론망 모델은 확률적 추론망에 기반한 검색모형이고, 퍼지탐색은 퍼지집합 이론을 이용한 검색 방식이므로 확장 불리언의 대표 모형 두 가지가 아니다. 틀린 표현은 '추론망 모델, 퍼지탐색'이며, 올바르게는 'MMM 모델, P-norm 모델'이다.
② O — MMM 모델과 P-norm 모델은 모두 불리언 질의의 엄격한 일치 조건을 완화하여 부분 일치와 순위화를 지원하는 확장 불리언 검색모형이다.
③ X — 베이지언 네트워크와 추론망 모델은 확률과 추론망을 이용하는 확률적 검색 계열이지 확장 불리언 검색모형이 아니다. 틀린 표현은 '베이지언 네트워크, 추론망 모델'이며, 올바르게는 'MMM 모델, P-norm 모델'이다.
④ X — 신경망 모델은 연결 가중치 학습을 이용하는 모형이고, 신념망 모델은 확률적 의존 관계를 표현하는 네트워크 모형이므로 확장 불리언의 대표 모형 조합이 아니다. 틀린 표현은 '신경망 모델, 신념망 모델'이며, 올바르게는 'MMM 모델, P-norm 모델'이다.
⑤ X — 퍼지집합 모델은 퍼지 소속도를 이용하고, 잠재의미색인 모델은 용어-문서 행렬을 저차원 잠재 의미 공간으로 축소하는 대수적 모형이므로 둘 다 확장 불리언 검색모형이 아니다. 틀린 표현은 '퍼지집합 모델, 잠재의미색인 모델'이며, 올바르게는 'MMM 모델, P-norm 모델'이다.

NoteBase

오답 포인트

퍼지집합 모델도 0과 1 사이의 값을 사용하고 부분 일치를 허용하므로 확장 불리언과 혼동하기 쉽다. 그러나 이 문항은 '부분 일치를 지원하는가'가 아니라 '확장 불리언 검색모형의 대표 유형인가'를 묻는다. 대표 유형은 MMM 모델과 P-norm 모델이다.

다시 나오면 이렇게 판단

검색모형의 이름을 계열별로 묶어 판단한다. MMM·P-norm은 확장 불리언, 베이지언 네트워크·추론망·신념망은 확률적 네트워크, 잠재의미색인은 대수적 모형, 퍼지집합은 퍼지집합 기반 모형으로 구분한다.

암기 포인트

확장 불리언은 'MMM + P-norm'이다.

NoteBase

16. 인용색인에 대한 설명으로 옳지 않은 것은?

① 문헌 A의 참고문헌과 문헌 B의 참고문헌 중 동일한 문헌이 10개 있다면 동시인용도는 10이다.
② 서지결합법은 미국 MIT의 Kessler가 제시하였다.
③ 서지결합법에서 여러 개의 문헌이 공통되는 인용문헌을 하나 이상 가질 때 이 문헌들은 서로 주제적으로 연관되어 있다.
④ SCI, SSCI, A&HCI는 대표적인 인용색인 데이터베이스이다.
⑤ 동시인용법은 문헌들의 외부적 관계에 근거한 동적 연결이라 할 수 있다.

정답 1번

출제 포인트

서지결합법, 동시인용법, 동시이용법이 문헌 사이의 관계를 무엇으로 측정하는지 구별하는 문제이다.

핵심 개념

서지결합법은 두 문헌이 함께 인용한 공통 참고문헌의 수로 두 문헌의 관련성을 측정하는 방법이다.
동시인용법은 후속 문헌들이 두 문헌을 함께 인용한 횟수로 두 문헌의 관련성을 측정하는 방법이다.
동시이용법은 이용자가 두 문헌을 함께 이용한 기록을 바탕으로 문헌 간 관계를 파악하는 방법이다.
따라서 서지결합은 공통 참고문헌, 동시인용은 제3의 문헌에 의한 공동 인용, 동시이용은 이용 기록이 판단 기준이다.

선지별 해설

① X — 문헌 A와 문헌 B의 참고문헌 가운데 동일한 문헌이 10개라는 것은 서지결합도가 10이라는 뜻이다. 틀린 표현은 '동시인용도는 10'이며, 올바르게는 '서지결합도는 10'이다.
② O — 서지결합법은 미국 MIT의 Kessler가 제시한 문헌 관계 분석 방법이다.
③ O — 여러 후속 문헌이 두 문헌을 함께 인용한다면 두 문헌이 공통된 주제나 지식 기반과 관련되어 있다고 해석할 수 있다. 이는 동시인용법의 기본 원리이다.
④ O — SCI, SSCI, A&HCI는 각각 과학, 사회과학, 예술·인문 분야의 대표적인 인용색인 데이터베이스이다.
⑤ O — 동시이용법은 두 문헌이 함께 이용되는 외부의 이용 관계를 활용하므로 새로운 이용 기록에 따라 관계가 변할 수 있는 동적 연결로 볼 수 있다.

오답 포인트

①의 '동일한 참고문헌 10개'를 보고 숫자 10만 확인하면 동시인용도로 오인하기 쉽다. 두 문헌이 과거의 동일 문헌을 인용하면 서지결합이고, 이후의 제3 문헌이 두 문헌을 함께 인용하면 동시인용이다.

NoteBase

다시 나오면 이렇게 판단

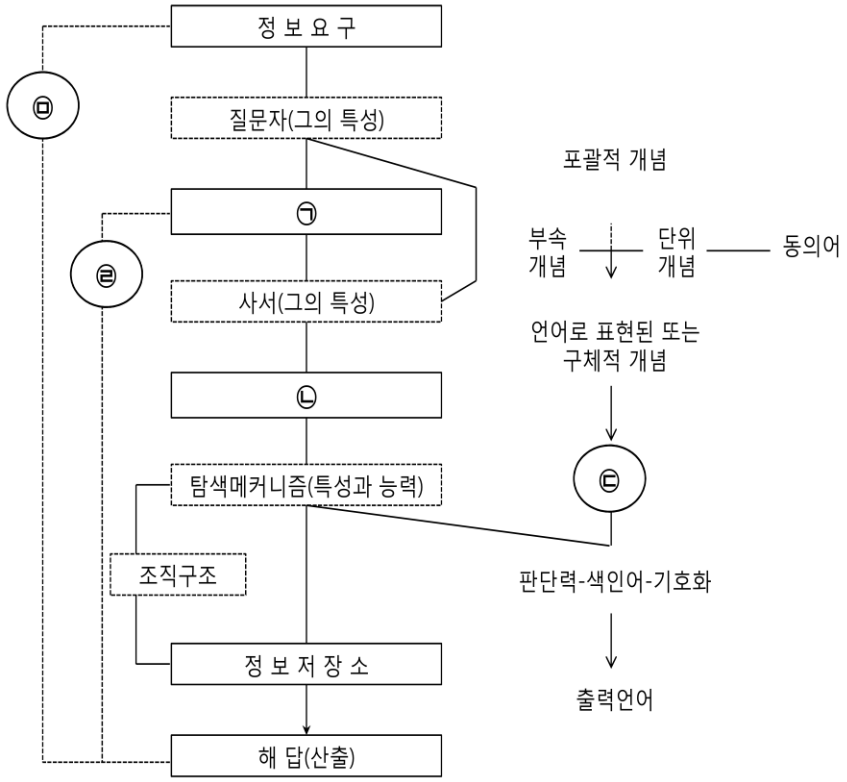
관계의 근거가 어디에 있는지를 확인한다. 두 문헌의 참고문헌 목록에 있으면 서지결합, 후속 문헌의 인용 목록에 있으면 동시인용, 이용자의 사용 기록에 있으면 동시이용으로 판단한다.

암기 포인트

서지결합은 뒤를 함께 보고, 동시인용은 앞에서 함께 인용되며, 동시이용은 함께 사용된다.

NoteBase

17. 다음은 세라의 참고과정 모델이다. ㉠~㉤에 들어갈 용어로 옳은 것은?



	㉠	㉡	㉢	㉣	㉤
①	개념정리	탐색전략	질문언어	정확성평가	적합성평가
②	개념정리	탐색실행	반응	정확성평가	적합성평가
③	질문 (언어적 표현)	탐색실행	반응	연관성평가	유용성평가
④	질문 (언어적 표현)	탐색전략	질문언어	연관성평가	적합성평가
⑤	질문 (언어적 표현)	탐색전략	질문언어	연관성평가	유용성평가

정답 4번

출제 포인트

세라의 참고과정 모델에서 정보요구가 질문의 언어적 표현으로 바뀌고, 사서가 탐색전략을 세워 질문언어로 변환하는 흐름과 두 평가 단계의 대상을 구분하는 문제이다.

NoteBase

선지별 해설

① X
ㄱ을 '개념정리', ㄹ을 '정확성평가'로 제시한 부분이 틀렸다. ㄱ은 질문(언어적 표현)이고 ㄹ은 연관성평가이다. ㄴ은 탐색전략, ㄷ은 질문언어, ㅁ은 적합성평가가 맞다.

② X
ㄱ '개념정리', ㄴ '탐색실행', ㄷ '반응', ㄹ '정확성평가'가 틀렸다. 올바르게는 ㄱ 질문(언어적 표현), ㄴ 탐색전략, ㄷ 질문언어, ㄹ 연관성평가이며 ㅁ은 적합성평가이다.

③ X
ㄴ을 '탐색실행', ㄷ을 '반응', ㅁ을 '유용성평가'로 제시한 부분이 틀렸다. 올바르게는 ㄴ 탐색전략, ㄷ 질문언어, ㅁ 적합성평가이다. ㄱ 질문(언어적 표현)과 ㄹ 연관성평가는 맞다.

④ O
도식의 흐름에 따라 ㄱ은 질문(언어적 표현), ㄴ은 탐색전략, ㄷ은 질문언어이다. 검색 결과와 질문의 관계를 확인하는 ㄹ은 연관성평가이고, 산출된 해답이 이용자의 정보요구에 맞는지를 확인하는 ㅁ은 적합성평가이다.

⑤ X
ㅁ을 '유용성평가'로 제시한 부분이 틀렸다. ㅁ은 해답이 처음의 정보요구에 부합하는지를 판단하는 적합성평가이다. ㄱ 질문(언어적 표현), ㄴ 탐색전략, ㄷ 질문언어, ㄹ 연관성평가는 맞다.

오답 포인트

'탐색전략'과 '탐색실행', '질문언어'와 '반응'을 바꾸어 놓은 선지가 함정이다. 또한 연관성평가는 검색 결과가 질문과 관련되는지를 보는 단계이고, 적합성평가는 최종 해답이 이용자의 정보요구에 맞는지를 보는 단계이므로 평가 대상이 다르다.

다시 나오면 이렇게 판단

중앙의 진행 순서를 '정보요구 → 질문의 언어적 표현 → 탐색전략 → 질문언어 → 탐색메커니즘 → 정보저장소 → 해답'으로 먼저 복원한다. 그다음 검색 결과와 질문의 관계는 연관성평가, 해답과 정보요구의 관계는 적합성평가로 대응한다.

암기 포인트

질문을 표현하고 전략을 세워 질문언어로 검색한다. 결과는 질문과의 연관성, 해답은 요구와의 적합성을 평가한다.

NoteBase

18. <보기> 중 검색된 웹페이지를 순위화하는 알고리즘에 대한 설명으로 옳은 것을 모두 고르면?

<보 기>

ㄱ. PageRank와 HITS는 웹페이지들의 링크정보를 분석하여 검색결과를 순위화하는 알고리즘이다.

ㄴ. PageRank 알고리즘은 특정한 페이지가 후진링크를 많이 가질수록 중요성이 높다고 할 수 있다.

ㄷ. PageRank 알고리즘은 전진링크와 후진링크를 활용하여 Authority와 Hub 페이지를 찾는다.

ㄹ. PageRank 알고리즘은 인용색인의 원리와 유사하다.

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

정답 3번

출제 포인트

PageRank와 HITS가 링크 구조를 이용하는 방식의 공통점과 차이점을 구별하는 문제이다.

핵심 개념

PageRank: 다른 웹페이지로부터 링크를 많이 받을수록, 특히 중요도가 높은 페이지로부터 링크를 받을수록 해당 페이지를 중요하게 평가하는 링크 분석 알고리즘이다. 무작위 이용자가 링크를 따라 이동한다는 관점으로 설명할 수 있으며 인용색인과 유사한 발상을 가진다.
HITS: 링크를 많이 받는 Authority 페이지와 좋은 Authority 페이지들을 많이 가리키는 Hub 페이지를 함께 찾는 알고리즘이다.
핵심 대비: PageRank는 페이지별 중요도 점수를 계산하고, HITS는 Authority와 Hub라는 두 역할의 점수를 계산한다.

NoteBase

선지별 해설

① X — ㄱ과 ㄴ은 옳지만 옳은 ㄹ이 빠져 있다. 올바른 조합은 ㄱ, ㄴ, ㄹ이다.
② X — ㄱ과 ㄴ은 옳지만 ㄷ은 틀리고, 옳은 ㄹ이 빠져 있다. 올바른 조합은 ㄱ, ㄴ, ㄹ이다.
③ O — ㄱ, ㄴ, ㄹ이 모두 옳고 ㄷ은 틀리므로 정확한 조합이다.
④ X — ㄴ과 ㄹ은 옳지만 ㄷ은 틀리고, 옳은 ㄱ이 빠져 있다. 올바른 조합은 ㄱ, ㄴ, ㄹ이다.
⑤ X — ㄷ을 옳은 설명으로 포함한 부분이 틀렸다. ㄷ의 'PageRank'는 'HITS'로 고쳐야 하므로 올바른 조합은 ㄱ, ㄴ, ㄹ이다.

보기 판단
ㄱ O — 두 알고리즘 모두 웹페이지 사이의 링크 구조를 분석하여 검색 결과의 순위를 정하는 데 활용된다.
ㄴ O — PageRank에서는 다른 페이지로부터 받는 후진링크가 중요도 판단의 핵심이며, 링크를 주는 페이지의 중요도도 함께 반영된다.
ㄷ X — 전진링크와 후진링크를 이용하여 Authority와 Hub 페이지를 찾는 것은 PageRank가 아니라 HITS이다. 올바르게는 'HITS 알고리즘은 링크 관계를 이용하여 Authority와 Hub 페이지를 찾는다'이다.
ㄹ O — 다른 문헌에 많이 인용되는 문헌을 중요하게 보는 인용색인의 발상처럼, PageRank도 다른 페이지에서 받는 링크를 페이지 중요도의 신호로 사용한다.

오답 포인트

ㄷ의 'Authority와 Hub'라는 표현을 PageRank의 특징으로 착각하기 쉽다. Authority와 Hub를 구분하는 알고리즘은 HITS이고, PageRank는 링크를 통해 전달되는 페이지 중요도를 하나의 점수로 계산한다. 또한 조합형 문항이므로 옳은 보기 하나를 빠뜨린 ①도 오답이다.

다시 나오면 이렇게 판단

'Authority·Hub'가 나오면 HITS, '중요한 페이지가 건 링크·무작위 이용자·인용 관계'가 나오면 PageRank로 판단한다. 각 보기의 O/X를 먼저 확정한 뒤 그 결과와 일치하는 조합을 고른다.

암기 포인트

PageRank는 링크가 전달하는 중요도, HITS는 Authority와 Hub의 상호 강화이다.

NoteBase

19. OpenURL에 대한 설명 중 빈칸에 해당하는 개념은?

> OpenURL 프레임워크는 참조대상이 되는 연계자원에 대한 메타데이터와 식별자를 기반으로 하여 (　　)을/를 지원하는 아키텍처이다. 또한 링크 생성자가 주기적으로 링크의 연결여부를 관리해야만 하는 부담을 줄임과 더불어 안정적인 연결 서비스를 지원한다.

① 상황인식(context-sensitive)연계
② 적합자원(appropriate copy)연계
③ 연계검색(linking search)모델
④ 링크드데이터(linked data)
⑤ 브라우징(browsing)

정답 1번

출제 포인트

OpenURL 프레임워크가 서지 메타데이터와 식별자뿐 아니라 이용자의 소속·접근 권한 등 맥락을 해석하여 적절한 연결 대상을 제시하는 상황인식 연계를 지원한다는 점을 묻는 문제이다.

핵심 개념

OpenURL은 인용된 정보자원의 메타데이터와 식별자를 컨텍스트 객체로 전달하고, 링크 리졸버가 그 정보와 이용자의 상황을 해석하여 이용 가능한 원문이나 관련 서비스를 연결하는 프레임워크이다.
상황인식 연계는 동일한 인용정보라도 이용자의 소속기관, 접근 권한, 기관 구독 범위 등의 맥락에 따라 적합한 연결 서비스를 제공하는 방식이다. 링크 생성자가 개별 목적지 URL을 계속 갱신하는 대신 리졸버가 연결을 결정하므로 링크 유지 부담을 줄이고 안정적인 서비스를 제공할 수 있다.

NoteBase

선지별 해설

① O — 메타데이터와 식별자를 바탕으로 이용자의 맥락에 적합한 서비스를 선택하는 것은 상황인식 연계이다. 따라서 빈칸에 알맞다.
② X — '적합자원 연계'는 이용자가 접근할 수 있는 적절한 복제본을 선택하는 개별 서비스에 가까우며, 제시문이 설명하는 OpenURL의 포괄적 연결 개념명은 아니다. 틀린 표현 '적합자원 연계'는 '상황인식 연계'로 고쳐야 한다.
③ X — 연계검색 모델은 여러 정보원이나 검색 대상을 함께 탐색하는 검색 방식에 관한 개념이므로, 이용자 맥락에 따라 링크 목적지를 결정하는 OpenURL의 기능을 가리키지 않는다. 틀린 표현 '연계검색 모델'은 '상황인식 연계'로 고쳐야 한다.
④ X — 링크드데이터는 URI로 대상을 식별하고 RDF 형식의 관계를 통해 웹의 데이터를 연결하는 방식이다. 제시문의 OpenURL 기반 맥락별 연결 서비스와는 다른 개념이므로 '링크드데이터'는 '상황인식 연계'로 고쳐야 한다.
⑤ X — 브라우징은 목록이나 링크를 따라 정보자원을 탐색하는 행위이다. 메타데이터와 이용자 맥락을 해석하여 적절한 연결 서비스를 선택하는 아키텍처를 뜻하지 않으므로 '브라우징'은 '상황인식 연계'로 고쳐야 한다.

오답 포인트

②의 '적합자원'과 ④의 '링크드데이터'도 자원을 연결한다는 인상을 주지만, 제시문의 결정적 단서는 이용자의 상황에 따라 연결 결과가 달라진다는 점이다. 또한 링크 생성자가 목적지 변동을 직접 관리하지 않는다는 설명은 리졸버가 중간에서 적절한 서비스를 선택하는 구조를 뜻한다.

다시 나오면 이렇게 판단

OpenURL 문제에서는 '메타데이터·식별자 전달', '링크 리졸버', '이용자나 기관의 맥락에 맞는 서비스 선택'이 함께 제시되면 상황인식 연계로 판단한다. URI와 RDF의 관계망이면 링크드데이터, 여러 정보원의 동시 검색이면 연계검색으로 구분한다.

암기 포인트

OpenURL은 인용정보와 이용자 맥락을 리졸버에 전달하고, 리졸버는 그 상황에 맞는 링크를 선택한다.

NoteBase

20. 디지털 파일 형식과 그에 대한 설명으로 옳지 않은 것은?

① TIFF : 장기보존 또는 마스터 디지털화가 필요한 경우 사용되는 저장형식이다.

② JPEG 2000 : 무손실 압축을 지원하며 한 이미지에 여러 해상도의 이미지를 포함할 수 있는 정지 디지털 이미지의 압축을 위한 국제표준포맷이다.

③ Raw : 파일디지털 카메라나 스캐너의 이미지 센서로부터 최소한으로만 처리된 이미지파일 포맷이다.

④ PDF/A : 전자문서의 디지털 보존을 위해 특별히 고안된 포터블 문서 포맷이다.

⑤ PNG : 손실 압축을 사용하므로 이미지 디테일 손실이 있지만 결과적으로 JPEG 형식보다 파일 크기는 더 작아진다.

정답 5번

출제 포인트

디지털 파일 형식별 압축 방식과 보존·활용 목적을 구분하는 문제이다. 특히 PNG의 무손실 압축을 JPEG의 손실 압축과 혼동하지 않아야 한다.

핵심 개념

TIFF: 비압축 또는 무손실 압축으로 고품질 이미지를 저장할 수 있어 보존용 마스터 파일에 적합한 형식이다.
JPEG 2000: 정지영상 압축 표준으로, 무손실 압축과 해상도 단계별 접근을 지원한다.
RAW: 디지털카메라나 스캐너의 이미지 센서가 수집한 데이터를 최소한으로 처리하여 저장한 형식이다.
PDF/A: 전자문서의 장기 보존을 목적으로 표준화된 PDF 계열 형식이다.
PNG: 무손실 압축을 사용하여 저장과 복원 과정에서 이미지 정보가 손실되지 않는 형식이다.

NoteBase

선지별 해설

① O — TIFF는 비압축 또는 무손실 방식으로 높은 화질을 유지할 수 있으므로 장기 보존용이나 마스터 디지털화 파일에 널리 활용된다.
② O — JPEG 2000은 무손실 압축을 지원하며, 하나의 압축 데이터에서 여러 해상도 수준의 이미지를 이용할 수 있는 정지영상 압축 표준이다.
③ O — RAW는 카메라나 스캐너의 이미지 센서에서 얻은 데이터를 색상·선명도 등의 후처리를 최소화한 상태로 저장한다.
④ O — PDF/A는 글꼴·색상 정보 등 장기 재현에 필요한 요소를 문서 안에 보존하도록 제약을 둔 전자문서 장기 보존용 PDF 형식이다.
⑤ X — 틀린 표현은 '손실 압축을 사용하므로 이미지 디테일 손실이 있지만 결과적으로 JPEG 형식보다 파일 크기는 더 작아진다'이다. PNG는 무손실 압축을 사용하므로 이미지 정보가 손실되지 않으며, 사진과 같은 이미지에서는 일반적으로 손실 압축 방식인 JPEG보다 파일 크기가 더 커진다.

오답 포인트

⑤는 '압축'이라는 말만 보고 모든 압축에서 화질 손실이 생긴다고 생각하게 만드는 선지이다. PNG의 압축은 무손실이고 JPEG의 압축은 일반적으로 손실이라는 차이가 핵심이다. 또한 무손실 PNG가 사진에서도 JPEG보다 항상 더 작아지는 것은 아니다.

다시 나오면 이렇게 판단

파일 형식 문제에서는 먼저 압축이 손실인지 무손실인지 확인한 뒤 주된 용도를 연결한다. TIFF는 보존용 마스터, JPEG 2000은 무손실·다중 해상도, RAW는 센서 원자료, PDF/A는 전자문서 장기 보존, PNG는 무손실 이미지 형식으로 판단한다.

암기 포인트

PNG는 무손실, JPEG는 손실, RAW는 센서 원자료, TIFF는 보존용 마스터, PDF/A는 장기 보존 문서이다.

NoteBase