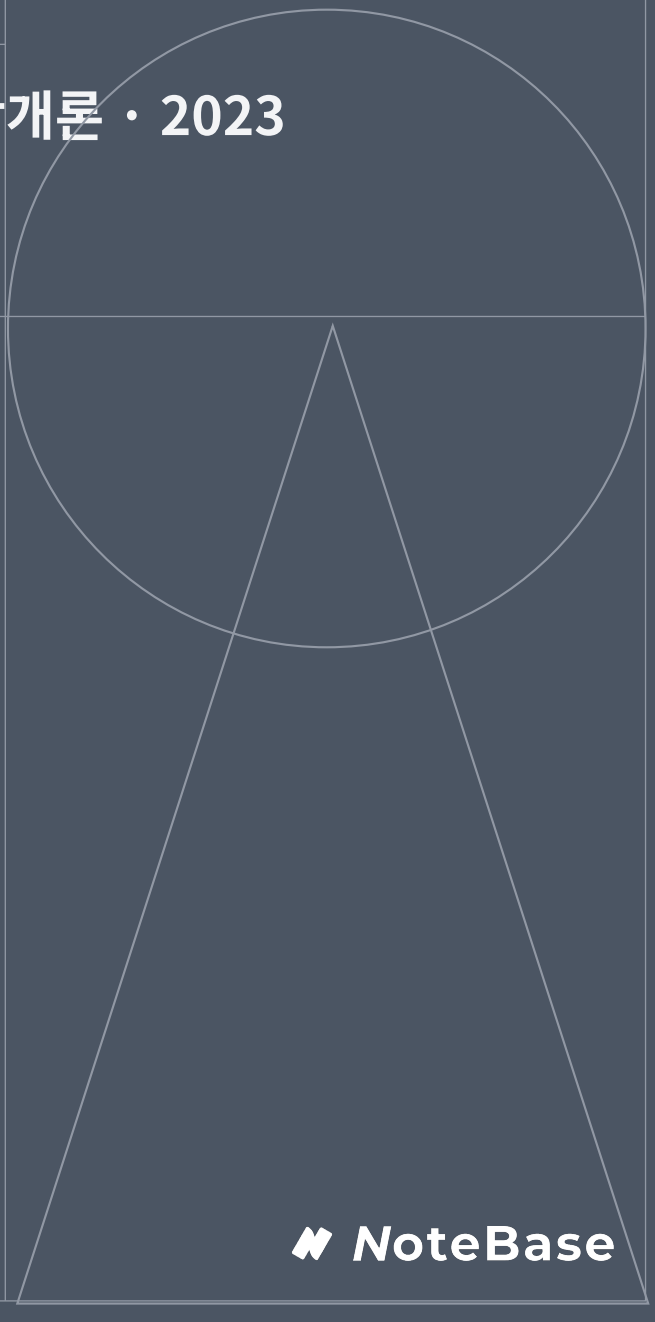
기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2023



NoteBase

기출OX 해설집

공무원/국회직/9급 · 정보학개론 · 2023

공무원/국회직/9급 · 정보학개론 · 2023

1. 다음 정보검색 분야에 업적을 남긴 학자는?

> SMART(System for the Manipulation And Retrieval of Texts), 벡터공간 모델, 용어가중치, 적합성 피드백, 클러스터링, 확장불리언 검색, 텍스트 이해와 구조화, 자동텍스트 처리 등

① 유진 가필드(Eugene Garfield)
② 한스 피터 룬(Hans Peter Luhn)
③ 캘빈 무어스(Calvin Mooers)
④ 제라드 설튼(Gerard Salton)
⑤ 모티머 토브(Mortimer Taube)

정답 4번

출제 포인트

이 문제는 정보검색 발전에 기여한 학자와 대표 업적을 연결하는 문제이다. 제시문에서 결정적인 단서는 SMART 정보검색 시스템과 벡터공간 모델이다.

핵심 정리

제라드 설톤(Gerard Salton): SMART 정보검색 시스템을 개발하고 벡터공간 모델, 용어가중치, 적합성 피드백 등 현대 정보검색의 핵심 기법을 발전시켰다.
유진 가필드(Eugene Garfield): 인용색인과 과학인용색인의 발전에 기여하였다.
한스 피터 룬(Hans Peter Luhn): KWIC 색인, 자동색인, 자동초록 등 문헌정보 처리의 자동화에 기여하였다.
캘빈 무어스(Calvin Mooers): 'information retrieval'이라는 용어를 제시하고 정보검색 개념의 정립에 기여하였다.
모티머 토브(Mortimer Taube): 유니텀(Uniterm)과 좌표색인법을 발전시켰다.

NoteBase

선지별 해설

① X — 유진 가필드의 대표 업적은 인용색인과 과학인용색인이다. SMART와 벡터공간 모델의 개발자로 연결한 것이 틀리며, 올바르게는 제라드 설톤이다.
② X — 한스 피터 룬은 KWIC 색인, 자동색인, 자동초록 등 정보처리 자동화와 관련된다. 제시된 업적의 학자로 연결한 것이 틀리며, 올바르게는 제라드 설톤이다.
③ X — 캘빈 무어스는 'information retrieval'이라는 용어와 정보검색 개념의 정립으로 알려져 있다. SMART와 벡터공간 모델의 학자로 연결한 것이 틀리며, 올바르게는 제라드 설톤이다.
④ O — 제라드 설톤은 SMART 시스템을 개발하였고 벡터공간 모델, 용어가중치, 적합성 피드백, 클러스터링 등 정보검색 기법의 발전에 큰 업적을 남겼다.
⑤ X — 모티머 토브는 유니텀과 좌표색인법을 발전시킨 인물이다. 제시된 업적의 학자로 연결한 것이 틀리며, 올바르게는 제라드 설톤이다.

출제자가 노린 함정

선지의 다섯 인물이 모두 정보검색과 색인의 역사에 관련되어 있다는 점이 함정이다. 특히 정보검색이라는 용어를 제시한 무어스와 실제 검색 시스템 및 수학적 검색 모델을 발전시킨 설톤을 혼동하기 쉽다.

다시 나오면 이렇게 판단

SMART 또는 벡터공간 모델이 보이면 제라드 설톤을 선택한다. 인용색인은 가필드, KWIC는 룬, 정보검색이라는 용어는 무어스, 유니텀·좌표색인은 토브로 구분한다.

암기 포인트

설톤 = SMART·벡터공간, 가필드 = 인용색인, 룬 = KWIC, 무어스 = 정보검색 용어, 토브 = 유니텀·좌표색인이다.

NoteBase

2. 다음 테일러(Taylor)가 제시한 이용자의 정보요구 구체화 과정을 순서대로 바르게 나열한 것은?

ㄱ. 명시적 정보요구(Formalized Need)
ㄴ. 의식적 정보요구(Conscious Need)
ㄷ. 잠재적 정보요구(Visceral Need)
ㄹ. 협의적 정보요구(Compromised Need)

① ㄴ → ㄷ → ㄱ → ㄹ

② ㄴ → ㄹ → ㄱ → ㄷ

③ ㄷ → ㄴ → ㄱ → ㄹ

④ ㄷ → ㄴ → ㄹ → ㄱ

⑤ ㄷ → ㄹ → ㄴ → ㄱ

정답 3번

출제 포인트

이 문제는 테일러(Taylor)가 제시한 이용자의 정보요구가 내면의 막연한 필요에서 실제 검색시스템에 제시할 수 있는 요구로 구체화되는 순서를 묻는 문제이다.

올바른 순서

1. ㄷ 잠재적 정보요구(Visceral Need) — 이용자가 부족함을 느끼지만 그 요구를 아직 명확한 말로 표현하지 못한 단계이다.
2. ㄴ 의식적 정보요구(Conscious Need) — 이용자가 자신의 정보 부족을 의식하고 머릿속에서 요구를 어느 정도 구체화한 단계이다.
3. ㄱ 명시적 정보요구(Formalized Need) — 이용자가 정보요구를 언어로 명확하게 표현한 단계이다.
4. ㄹ 협의적 정보요구(Compromised Need) — 이용자가 정보시스템이나 정보원의 범위와 제약을 고려하여 실제 검색에 사용할 질문으로 조정한 단계이다.

왜 이 순서인가: 먼저 막연한 필요를 인식한 뒤 이를 언어로 공식화해야 하며, 그렇게 표현된 요구를 실제 검색 환경에 맞게 조정한 결과가 협의적 정보요구이기 때문이다.

NoteBase

선지별 해설

① X — ㄴ 의식적 정보요구보다 ㄷ 잠재적 정보요구가 먼저 발생하며, ㄹ 협의적 정보요구는 ㄱ 명시적 정보요구 뒤에 와야 한다. 틀린 순서 'ㄴ → ㄷ → ㄱ → ㄹ'은 'ㄷ → ㄴ → ㄱ → ㄹ'로 고쳐야 한다.

② X — ㄹ 협의적 정보요구를 두 번째에 둔 것이 잘못이다. 협의적 정보요구는 명시적으로 표현된 요구를 실제 정보시스템에 맞게 조정한 마지막 단계이므로, 'ㄴ → ㄹ → ㄱ → ㄷ'은 'ㄷ → ㄴ → ㄱ → ㄹ'로 고쳐야 한다.

③ O — ㄷ 잠재적 정보요구에서 출발하여 ㄴ 의식적 정보요구, ㄱ 명시적 정보요구를 거쳐 ㄹ 협의적 정보요구로 구체화되므로 올바른 순서이다.

④ X — ㄷ 다음에 ㄴ이 오는 앞부분은 맞지만, ㄹ과 ㄱ의 선후관계가 뒤바뀌었다. 먼저 요구를 ㄱ 명시적으로 표현한 뒤 ㄹ 검색 환경에 맞게 조정하므로, 'ㄷ → ㄴ → ㄹ → ㄱ'은 'ㄷ → ㄴ → ㄱ → ㄹ'로 고쳐야 한다.

⑤ X — ㄹ 협의적 정보요구를 의식적·명시적 정보요구보다 앞에 둔 것이 잘못이다. 검색 환경과 타협하려면 먼저 요구를 의식하고 명시적으로 표현해야 하므로, 'ㄷ → ㄹ → ㄴ → ㄱ'은 'ㄷ → ㄴ → ㄱ → ㄹ'로 고쳐야 한다.

오답 포인트

'명시적'과 '협의적'을 혼동하기 쉽다. 명시적 정보요구는 이용자가 자신의 요구를 말로 표현한 상태이고, 협의적 정보요구는 그 표현을 실제 시스템과 정보원의 제약에 맞추어 검색 가능한 질문으로 바꾼 상태이므로 협의적 정보요구가 마지막이다.

다시 나오면 이렇게 판단

정보요구가 구체화되는 정도를 기준으로 판단한다. 말로 표현하기 전의 막연한 필요에서 시작하여, 필요를 의식하고, 언어로 명시한 다음, 실제 검색 환경에 맞게 조정하는 흐름이다.

암기 포인트

잠재 → 의식 → 명시 → 협의, 즉 '느낌 → 인식 → 표현 → 검색에 맞춘 조정'의 순서이다.

NoteBase

3. 문헌 내에서 문장을 분석하는 자동색인의 비통계적기법이 아닌 것은?

① 구문분석기법
② 단순기법
③ 문헌구조적기법
④ 문헌분리가중기법
⑤ 의미분석기법

정답 4번

출제 포인트

자동색인 기법을 문장의 언어적 구조와 의미를 이용하는 비통계적 기법과, 문헌 집합에서 산출한 수치·가중치를 이용하는 통계적 기법으로 구분하는 문제이다.

핵심 개념

비통계적 자동색인 기법은 문장의 구문, 문헌의 구조, 단어의 의미와 같은 언어적·구조적 규칙을 이용해 색인어를 선정하는 방식이다.
통계적 자동색인 기법은 단어의 출현 빈도나 문헌을 구별하는 정도처럼 문헌 집합에서 계산한 수치와 가중치를 이용한다.
문헌분리가중기법은 용어가 문헌들을 얼마나 잘 구별하는지를 수치화하여 가중치를 부여하므로 통계적 기법에 속한다.

선지별 해설

① O — 구문분석기법은 문장의 문법적 구조와 단어의 구문적 기능을 분석해 색인어 후보를 정하므로 비통계적 기법이다.
② O — 단순기법은 불용어를 제외하고 남은 내용어를 색인어 후보로 삼는 규칙 중심 방식이므로 비통계적 기법이다.
③ O — 문헌구조적기법은 표제, 초록, 문단 등 문헌 내부의 구조와 단어가 놓인 위치를 이용하므로 비통계적 기법이다.
④ X — 비통계적 기법이라는 분류가 틀렸다. 문헌분리가중기법은 용어의 문헌 구별 능력을 수치화해 가중치를 주는 통계적 기법이다. 올바르게는 '문헌분리가중기법은 통계적 자동색인 기법이다'라고 해야 한다.
⑤ O — 의미분석기법은 단어의 의미와 문맥, 개념 사이의 관계를 분석해 색인어를 선정하므로 비통계적 기법이다.

NoteBase

오답 포인트

‘문헌분리’라는 이름만 보면 문헌의 구조를 분석하는 방법처럼 보이지만, 핵심은 용어의 문헌 구별 능력에 수치적 가중치를 부여한다는 점이다. 따라서 ④를 문헌구조적기법과 같은 비통계적 기법으로 혼동하면 안 된다.

다시 나오면 이렇게 판단

기법이 구문·구조·의미 같은 규칙을 직접 분석하면 비통계적 기법으로, 빈도·구별력·가중치처럼 계산된 수치를 사용하면 통계적 기법으로 판단한다.

암기 포인트

구문·단순·문헌구조·의미는 비통계, 문헌분리‘가중’은 통계이다.

NoteBase

4. 다음은 2017년부터 2022년까지 K저널에 게재된 논문의 수와 K저널을 인용한 논문의 수를 나타낸 것이다. K저널의 2022년 인용계수(Impact Factor)는?

년도	K저널에 게재된 논문의 수	2022년 출판 논문의 K저널 인용건수
2017	60	40
2018	70	50
2019	100	70
2020	90	40
2021	110	30
2022	70	20
계	500	250

① 0.35
② 0.37
③ 0.40
④ 0.50
⑤ 0.70

정답 1번

출제 포인트

이 문제는 특정 연도의 영향력지수(Impact Factor)를 구할 때 직전 2개년에 게재된 논문 수와 해당 논문들이 그 특정 연도에 받은 인용 수만 사용하는지를 묻는다.

조건 정리

주어진 값: 2020년 논문 수 90편, 2021년 논문 수 110편, 2020년 논문의 2022년 인용 수 40회, 2021년 논문의 2022년 인용 수 30회이다.
구해야 할 값: K저널의 2022년 영향력지수이다.
주의할 조건: 2022년을 기준으로 직전 2개년인 2020년과 2021년의 값만 사용한다. 2017~2019년과 2022년의 논문 수·인용 수 및 표의 합계는 계산에 넣지 않는다.

사용할 공식

2022년 영향력지수=2020·2021년 논문이 2022년에 받은 인용 수2020·2021년에 게재된 논문 수
이 공식은 특정 연도의 영향력지수가 그 연도의 직전 2개년에 게재된 논문이 해당 연도에 얼마나 인용되었는지를 나타내는 비율이기 때문에 사용한다.

NoteBase

계산 과정

1. 2020년과 2021년에 게재된 논문 수를 더하면 90+110=200편이다.
2. 두 연도 논문이 2022년에 받은 인용 수를 더하면 40+30=70회이다.
3. 공식에 대입하면 다음과 같다.

$\frac{40+30}{90+110} = \frac{70}{200} = 0.35$

따라서 K저널의 2022년 영향력지수는 0.35이다.

선지 판단

① O — 0.35는 계산 결과 70/200 = 0.35와 일치한다.
② X — 0.37은 계산 결과와 일치하지 않는다. 틀린 값 0.37 → 올바르게는 0.35이다.
③ X — 0.40은 계산 결과와 일치하지 않는다. 틀린 값 0.40 → 올바르게는 0.35이다.
④ X — 0.50은 표의 전체 인용 수 250회를 전체 논문 수 500편으로 나눈 값이지만, 2022년 영향력지수에는 전체 기간 합계를 사용하지 않는다. 틀린 값 0.50 → 올바르게는 직전 2개년의 값만 사용한 0.35이다.
⑤ X — 0.70은 계산 결과와 일치하지 않는다. 틀린 값 0.70 → 올바르게는 0.35이다.
계산 결과가 0.35이므로 정답은 1번이다.

오답 포인트

가장 큰 함정은 표의 마지막 행에 제시된 전체 합계 250과 500을 그대로 나누어 0.50을 고르는 것이다.
영향력지수의 분모와 분자는 같은 대상 기간인 직전 2개년으로 맞춰야 하므로, 2022년 계산에서는 2020년과 2021년 행만 선택한다.

다시 나오면 이렇게 판단

영향력지수 계산에서는 먼저 기준 연도를 확인하고, 그 직전 2개년의 논문 수를 분모로 합산한 뒤, 그 논문들이 기준 연도에 받은 인용 수를 분자로 합산한다. 마지막으로 분자와 분모가 동일한 두 출판연도를 대상으로 하는지 검산한다.

암기 포인트

2022년 영향력지수는 2020·2021년 논문의 2022년 인용 수 ÷ 2020·2021년 논문 수이다.

NoteBase

5. 더블린 코어(Dublin Core)에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. 더블린 코어는 15개의 기본 요소로 구성되어 있다.

ㄴ. 물리적 기술요소는 Date, Format, Publisher, Language가 있다.

ㄷ. 더블린 코어의 각 요소를 표기하기 위하여 인코딩 스킴을 필수적으로 사용한다.

ㄹ. 더블린 코어의 Description과 대응되는 MODS의 요소는 abstract, tableOfContents, subject이다.

ㅁ. 더블린 코어 형식으로 작성된 레코드의 효율성과 정확성을 확보하기 위해 기본 데이터 요소의 의미를 보완할 수 있는 다양한 한정어를 마련하고 있다.

① ㄱ, ㄴ

② ㄱ, ㅁ

③ ㄱ, ㄴ, ㄷ

④ ㄱ, ㄷ, ㄹ

⑤ ㄱ, ㄴ, ㄹ, ㅁ

정답 2번

출제 포인트

더블린 코어의 15개 기본 요소, 요소의 기능별 분류, 한정어(qualifier)의 역할, MODS 요소와의 대응 관계를 구별하는 문제이다.

핵심 개념

더블린 코어는 다양한 정보자원을 단순하고 일관되게 기술하기 위한 15개의 기본 메타데이터 요소 집합이다.
기본 요소는 내용 기술, 지적재산권, 자원 인스턴스 기술의 성격으로 구분할 수 있다.
Date·Format·Identifier·Language는 자원의 물리적·형식적 구현과 관련된 요소이고, Publisher는 지적재산권 관련 요소이다.
한정어는 기본 요소의 의미를 세분하거나 값의 표현 방식을 지정하여 레코드의 정확성과 효율성을 높인다. 다만 한정어의 사용은 필수가 아니다.
MODS의 abstract와 tableOfContents는 더블린 코어의 Description에 대응하지만, MODS의 subject는 더블린 코어의 Subject에 대응한다.

NoteBase

선지별 해설

① X
ㄱ은 옳지만 ㄴ이 틀렸으므로 옳은 보기의 조합이 아니다. ㄴ의 틀린 표현은 물리적 기술요소에 'Publisher'를 포함한 부분이다. 올바르게는 Date, Format, Identifier, Language가 해당하고, Publisher는 지적재산권 관련 요소이다.

② O
ㄱ과 ㅁ은 모두 옳다. 더블린 코어는 15개의 기본 요소로 구성되며, 기본 요소의 의미를 보완하는 다양한 한정어를 두어 기술의 정확성과 활용 효율을 높일 수 있다.

③ X
ㄱ은 옳지만 ㄴ과 ㄷ이 틀렸다. ㄴ은 Publisher를 Identifier로 고쳐야 하며, ㄷ의 '인코딩 스킴을 필수적으로 사용한다'는 표현은 틀리다. 올바르게는 인코딩 스킴을 선택적으로 사용하여 값의 형식이나 통제어휘를 명시할 수 있다.

④ X
ㄱ은 옳지만 ㄷ과 ㄹ이 틀렸다. ㄷ은 인코딩 스킴이 필수가 아니라 선택 사항이라는 내용으로 고쳐야 한다. ㄹ의 틀린 부분은 MODS의 'subject'를 더블린 코어 Description에 대응시킨 것이다. 올바르게는 abstract와 tableOfContents가 Description에 대응하고, subject는 Subject에 대응한다.

⑤ X
ㄱ과 ㅁ은 옳지만 ㄴ과 ㄹ이 틀렸으므로 옳은 보기만의 조합이 아니다. ㄴ은 Publisher 대신 Identifier가 들어가야 하고, ㄹ은 subject를 Description 대응 요소에서 제외하여 Subject에 대응시켜야 한다. 따라서 옳은 것은 ㄱ과 ㅁ뿐이다.

오답 포인트

ㄴ에서는 Publisher와 Identifier의 분류를 바꾸어 놓았고, ㄷ에서는 선택적으로 쓰는 인코딩 스킴을 '필수'라고 표현하였다. ㄹ에서는 이름이 비슷한 Description과 subject의 대응 관계를 섞어 놓았다. 이 세 표현을 각각 바로잡아야 조합형 선지를 정확히 고를 수 있다.

다시 나오면 이렇게 판단

먼저 15개 기본 요소라는 사실을 확인한 뒤, Date·Format·Identifier·Language를 자원 인스턴스 기술 요소로 묶는다. 다음으로 한정어와 인코딩 스킴은 기본 요소를 정교하게 만들지만 필수는 아니라는 점을 적용한다. MODS 대응에서는 abstract·tableOfContents는 Description, subject는 Subject로 구별한다.

NoteBase

암기 포인트

더블린 코어는 기본 15요소이고, D-F-I-L(Date·Format·Identifier·Language)은 자원 인스턴스 기술 요소이다. 한정어는 선택적으로 보완하며, MODS subject는 DC Subject에 대응한다.

NoteBase

6. 클러스터링에 대한 설명으로 옳지 않은 것은?

① 싱글패스 기법은 비계층적 기법이다.

② 계층적 클러스터링 기법은 비계층적 기법에 비해 클러스터링 품질이 좋다.

③ 비계층적 클러스터링 기법은 계층적 기법에 비해 클러스터링 처리시간이 짧다.

④ 계층적 클러스터링은 비계층적 기법에 비해 상대적으로 많은 양의 문헌에 적용하기 어렵다.

⑤ 비교하고자 하는 두 대상을 표현하고 있는 속성 간의 일치하는 정도를 측정하는 유사도 계수는 상관계수이다.

정답 5번

출제 포인트

이 문제는 계층적 클러스터링과 비계층적 클러스터링의 처리 방식·품질·속도·적용 규모를 비교하고, 유사도 계수와 상관계수를 구별하는 문제이다.

핵심 개념

계층적 클러스터링: 대상을 단계적으로 병합하거나 분할하여 군집의 계층 구조를 만드는 기법이다. 군집 관계를 세밀하게 파악할 수 있지만 계산량이 커질 수 있다.
비계층적 클러스터링: 계층 구조를 만들지 않고 대상을 정해진 군집에 직접 배정하거나 재배치하는 기법이다. 싱글패스는 문헌을 순차적으로 한 번 처리하며 군집에 배정하는 대표적인 비계층적 기법이다.
유사도 계수: 두 대상이 얼마나 비슷한지를 수치화한 값이다. 속성값의 일치 정도를 재는 경우에는 일치계수 등을 사용하며, 상관계수는 수치적 속성의 변화 양상이나 프로파일 간 관련성을 재는 하나의 방법이지 유사도 계수 자체와 같은 개념은 아니다.

NoteBase

선지별 해설

① O — 싱글패스 기법은 문헌을 순서대로 한 번씩 살펴 기존 군집에 넣거나 새 군집을 만드는 비계층적 클러스터링 기법이다.
② O — 계층적 기법은 대상 간 관계를 단계적으로 비교하여 군집 구조를 형성하므로, 이 문항의 일반적 비교 기준에서는 비계층적 기법보다 클러스터링 품질이 좋다. 다만 그만큼 계산 부담이 크다.
③ O — 비계층적 기법은 모든 단계의 계층 구조를 만들지 않으므로 계층적 기법보다 처리시간이 짧다. 특히 싱글패스는 각 대상을 순차적으로 처리하여 속도가 빠르다.
④ O — 계층적 클러스터링은 대상 간 유사도를 반복적으로 계산하면서 병합 또는 분할해야 하므로 문헌 수가 많아질수록 계산량과 저장 부담이 커진다. 따라서 비계층적 기법보다 대량 문헌에 적용하기 어렵다.
⑤ X — 틀린 표현은 '유사도 계수는 상관계수이다'이다. 유사도 계수는 두 대상 사이의 유사성 정도를 나타내는 포괄적 개념이며, 속성의 일치 정도는 일치계수 등으로 측정한다. 상관계수는 수치적 속성의 변화 양상이나 프로파일 간 관련성을 측정하여 유사도 판단에 활용할 수 있는 하나의 방법일 뿐이다. 올바르게는 '두 대상의 속성값이 일치하는 정도를 측정하는 것은 일치계수이며, 상관계수는 유사도 측정에 활용되는 별도의 계수이다'라고 해야 한다.

오답 포인트

⑤는 '유사도 계수'라는 상위 개념과 그 측정에 활용될 수 있는 '상관계수'를 같은 개념처럼 단정한 것이 함정이다. 또한 속성의 일치 정도를 재는 계수는 상관계수가 아니라 일치계수로 구분해야 한다.

다시 나오면 이렇게 판단

클러스터링 비교 문제에서는 먼저 계층 구조를 만드는지를 확인한다. 계층적 기법은 품질과 구조 파악에 유리하지만 계산 부담이 크고, 비계층적 기법은 처리 속도와 대규모 적용에 유리하다. 계수 문제에서는 '값의 일치'는 일치계수, '수치적 변화 양상의 관련성'은 상관계수로 구별한다.

암기 포인트

계층적은 정교하지만 무겁고, 비계층적은 빠르고 대량 처리에 유리하다. 값의 일치는 일치계수, 변화 양상의 관련성은 상관계수이다.

NoteBase

7. 학술잡지의 법칙에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. Bradford 법칙은 제한된 예산에서 이용자를 만족시키는 중요한 학술잡지를 선택하는 데 응용할 수 있다.

ㄴ. Goffman은 고출현빈도의 단어에서 저출현빈도의 단어로 넘어가는 전환점을 밝혔다.

ㄷ. Lotka법칙을 활용하여 특정분야의 저자별 생산성에 대한 분포 비율을 계산할 수 있다.

ㄹ. Zipf는 학술지 구입결정이나 핵심 잡지군의 선택에 도움을 주는 서술법칙을 제시하였다.

① ㄱ, ㄴ

② ㄱ, ㄷ

③ ㄴ, ㄷ

④ ㄱ, ㄴ, ㄷ

⑤ ㄱ, ㄷ, ㄹ

정답 4번

출제 포인트

이 문제는 서지계량학의 대표 법칙과 학자, 그리고 각 법칙의 활용 대상을 정확히 연결하는 문제이다.

핵심 정리

Bradford 법칙: 특정 주제의 논문이 핵심 학술지와 그 밖의 학술지에 분산되는 양상을 설명하며, 제한된 예산에서 핵심 학술지를 선정하는 데 활용할 수 있다.
Goffman의 전환점: 단어의 출현 빈도 분포에서 고빈도어와 저빈도어 사이의 전환점을 이용하여 문헌의 내용어를 식별하는 개념이다.
Lotka 법칙: 저자의 논문 생산량과 그 생산량을 보이는 저자 수 사이의 분포를 설명하므로, 특정 분야의 저자별 생산성 분포를 계산하는 데 활용할 수 있다.
Zipf 법칙: 단어의 빈도와 빈도 순위가 반비례하는 관계를 설명하는 단어 빈도 법칙이다. 학술지 구입이나 핵심 잡지군 선정에 직접 적용되는 법칙은 Bradford 법칙이다.
따라서 옳은 것은 ㄱ, ㄴ, ㄷ이다.

NoteBase

선지별 해설

① X — ㄱ과 ㄴ은 옳지만, 옳은 ㄷ이 빠져 있다. 올바르게는 ㄱ, ㄴ, ㄷ을 모두 골라야 한다.
② X — ㄱ과 ㄷ은 옳지만, Goffman의 전환점에 관한 옳은 설명인 ㄴ이 빠져 있다. 올바르게는 ㄱ, ㄴ, ㄷ이다.
③ X — ㄴ과 ㄷ은 옳지만, Bradford 법칙의 활용을 바르게 설명한 ㄱ이 빠져 있다. 올바르게는 ㄱ, ㄴ, ㄷ이다.
④ O — ㄱ의 Bradford 법칙, ㄴ의 Goffman 전환점, ㄷ의 Lotka 법칙 설명이 모두 옳고, 틀린 ㄹ은 포함하지 않았다.
⑤ X — 옳은 ㄴ을 빠뜨리고 틀린 ㄹ을 포함했다. ㄹ의 'Zipf가 학술지 구입결정이나 핵심 잡지군 선택에 도움을 주는 서술법칙을 제시하였다'는 연결이 잘못되었으며, 올바르게는 Zipf 법칙은 단어의 빈도와 순위의 관계를 설명하고 학술지의 분산과 핵심 잡지군 선정은 Bradford 법칙과 연결된다. 따라서 올바른 조합은 ㄱ, ㄴ, ㄷ이다.

출제자가 노린 함정

Zipf 법칙도 단어의 출현 빈도를 다루므로 문헌 선택과 관련된 법칙처럼 보이게 만든 것이 함정이다. '단어의 빈도와 순위'는 Zipf, '저자 생산성'은 Lotka, '학술지 분산과 핵심 잡지군'은 Bradford로 구분해야 한다.

다시 나오면 이렇게 판단

학자 이름을 보면 적용 대상을 먼저 대응시킨다. Bradford는 학술지, Lotka는 저자, Zipf는 단어의 빈도·순위, Goffman은 단어 빈도의 전환점으로 판단한다. 그런 다음 옳은 보기의 조합이 빠짐없이 들어 있는 선택지를 고른다.

암기 포인트

Bradford=학술지 분산, Lotka=저자 생산성, Zipf=단어 빈도 순위, Goffman=전환점이다.

NoteBase

8. 다음 설명하고 있는 오픈액세스 학술지의 유형은?

> 출판사는 논문을 자유롭게 이용 가능하도록 하지만, 저자나 저자의 소속기관에서 저널처리비용을 부담하도록 함으로써 출판비용을 충당한다.

① Black OA
② Bronze OA
③ Hybrid OA
④ Gold OA
⑤ Green OA

정답 4번

출제 포인트

이 문제는 논문 이용 방식과 출판비용 부담 구조를 기준으로 오픈액세스(OA)의 유형을 구별하는 문제이다.

핵심 개념

Gold OA: 저자 또는 소속기관 등이 논문처리비용(APC)을 부담하고 출판사가 논문을 공개하여 누구나 자유롭게 이용하게 하는 방식이다.
Black OA: 저작권자의 허락 없이 논문을 불법적으로 공유하는 방식이다.
Bronze OA: 출판사 웹사이트에서 무료로 읽을 수 있지만 명시적인 공개 라이선스가 없어 지속적인 개방이 보장되지 않는 방식이다.
Hybrid OA: 구독형 학술지에서 저자가 비용을 내고 특정 논문만 오픈액세스로 공개하는 방식이다.
Green OA: 저자가 논문을 기관·주제 리포지터리 등에 셀프 아카이빙하여 공개하는 방식이다.

NoteBase

선지별 해설

① X — Black OA는 출판사가 저자 측의 논문처리비용을 받아 합법적으로 공개하는 유형이 아니다. 틀린 표현은 'Black OA'이며, 올바르게는 저작권자의 허락 없이 논문이 불법 공유되는 유형이다.
② X — Bronze OA는 저자 측의 비용 부담을 전제로 출판사가 공개하는 유형이 아니다. 틀린 표현은 'Bronze OA'이며, 올바르게는 출판사 사이트에서 무료 열람은 가능하지만 명시적 공개 라이선스가 없는 유형이다.
③ X — Hybrid OA라면 구독형 학술지라는 조건과 그 안의 특정 논문만 비용을 받고 공개한다는 특징이 제시되어야 한다. 틀린 표현은 'Hybrid OA'이며, 올바르게는 이러한 구독형 학술지 내 선택적 공개 방식이다.
④ O — 저자나 소속기관이 논문처리비용을 부담하고 출판사가 논문을 자유롭게 이용할 수 있도록 공개하는 방식은 Gold OA이다.
⑤ X — Green OA는 출판사의 논문처리비용으로 공개하는 방식이 아니라 저자가 리포지터리에 논문을 보존·공개하는 방식이다. 틀린 표현은 'Green OA'이며, 올바르게는 셀프 아카이빙 방식이다.

오답 포인트

③ Hybrid OA도 논문처리비용을 낼 수 있어 ④ Gold OA와 혼동하기 쉽다. 그러나 Hybrid OA는 '구독형 학술지에서 특정 논문만 공개'라는 조건이 핵심이며, 제시문처럼 저자 측의 비용 부담과 출판사의 공개라는 일반적 구조를 설명하면 Gold OA로 판단한다.

다시 나오면 이렇게 판단

비용 부담과 출판사 공개가 핵심이면 Gold OA, 구독형 학술지 속 일부 논문만 공개하면 Hybrid OA, 리포지터리 셀프 아카이빙이면 Green OA로 구별한다. 무료 열람만 가능하고 공개 라이선스가 없으면 Bronze OA, 무단 공유이면 Black OA이다.

암기 포인트

Gold는 출판사 공개, Green은 저장소 공개, Hybrid는 구독지 속 선택 공개이다. Bronze는 라이선스 없는 무료 열람, Black은 무단 공유이다.

NoteBase

9. 정보검색의 효율성에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. 부적합률과 배제율의 합은 1이다.

ㄴ. 잡음률과 누락률은 낮을수록 좋다.

ㄷ. 재현율과 정확률을 각각 50%로 유지하는 것이 이상적이다.

ㄹ. 배제율은 전체 부적합문헌 가운데 검색되지 않은 부적합문헌의 비율이다.

ㅁ. 시스템 내 포함되어 있는 적합문헌의 양은 재현율과 정확률에 의한 효율 측정에 영향을 준다.

① ㄱ, ㄴ, ㄹ

② ㄱ, ㄴ, ㅁ

③ ㄴ, ㄹ, ㅁ

④ ㄱ, ㄴ, ㄷ, ㄹ

⑤ ㄱ, ㄴ, ㄹ, ㅁ

정답 5번

출제 포인트

정보검색 효율성 평가 지표인 재현율·정확률·배제율·부적합률과 잡음률·누락률의 의미 및 바람직한 방향을 구별하는 문제이다.

핵심 개념

재현율은 전체 적합문헌 중 검색된 적합문헌의 비율이고, 정확률은 검색된 문헌 중 적합문헌의 비율이다. 두 비율은 높을수록 좋다.

부적합률은 전체 부적합문헌 중 검색된 부적합문헌의 비율이고, 배제율은 전체 부적합문헌 중 검색되지 않은 부적합문헌의 비율이므로 두 비율의 합은 1이다. 배제율은 높을수록 좋다.

잡음률은 검색 결과에 부적합문헌이 섞인 정도이고, 누락률은 적합문헌을 검색하지 못한 정도이므로 두 비율은 낮을수록 좋다.

NoteBase

선지별 해설

① X — ㄱ·ㄴ·ㄹ은 옳지만, 적합문헌의 양이 재현율과 정확률의 측정에 영향을 준다는 ㅁ도 옳으므로 옳은 내용을 모두 고른 조합이 아니다. 올바른 조합은 ㄱ·ㄴ·ㄹ·ㅁ이다.
② X — ㄱ·ㄴ·ㅁ은 옳지만, 배제율의 정의를 설명한 ㄹ도 옳으므로 불완전한 조합이다. 올바른 조합은 ㄱ·ㄴ·ㄹ·ㅁ이다.
③ X — ㄴ·ㄹ·ㅁ은 옳지만, 부적합률과 배제율의 합이 1이라는 ㄱ도 옳으므로 불완전한 조합이다. 올바른 조합은 ㄱ·ㄴ·ㄹ·ㅁ이다.
④ X — ㄱ·ㄴ·ㄹ은 옳지만 ㄷ이 틀렸다. 재현율과 정확률은 각각 50%로 유지하는 것이 이상적인 것이 아니라 모두 높을수록 이상적이다. 또한 옳은 ㅁ이 빠져 있다. 올바른 조합은 ㄱ·ㄴ·ㄹ·ㅁ이다.
⑤ O — ㄱ은 부적합률과 배제율이 상호 보완 관계라는 설명, ㄴ은 잡음률과 누락률이 낮을수록 좋다는 설명, ㄹ은 배제율의 정의, ㅁ은 적합문헌의 양이 재현율과 정확률에 영향을 준다는 설명으로 모두 옳다.

오답 포인트

ㄷ의 '각각 50%'를 재현율과 정확률 사이의 일반적인 절충점으로 오인하기 쉽다. 두 지표 사이에 상충 관계가 나타날 수는 있지만 50%가 이상적 기준인 것은 아니며, 평가 방향은 둘 다 높을수록 좋다는 것이다. 또한 ㄱ·ㄹ은 같은 분모인 전체 부적합문헌을 기준으로 검색 여부를 나눈 상호 보완 지표이다.

다시 나오면 이렇게 판단

먼저 분모가 전체 적합문헌인지, 검색된 문헌인지, 전체 부적합문헌인지 확인한다. 그다음 검색 성공을 나타내는 재현율·정확률·배제율은 높을수록 좋고, 실패를 나타내는 잡음률·누락률·부적합률은 낮을수록 좋다고 판단한다.

암기 포인트

재현·정확·배제는 높게, 잡음·누락·부적합은 낮게 기억한다. 부적합률과 배제율의 합은 1이다.

NoteBase

10. 온톨로지 구축과정을 순서대로 바르게 나열한 것은?

> ㄱ. 슬롯의 패싯 정의
> ㄴ. 개별 인스턴스 생성
> ㄷ. 클래스와 클래스 계층 정의
> ㄹ. 클래스의 속성인 슬롯 정의
> ㅁ. 기존 온톨로지의 재활용 검토
> ㅂ. 온톨로지 대상 도메인 및 범위 결정
> ㅅ. 온톨로지 주요용어 수집 및 나열

① ㅂ → ㅁ → ㄷ → ㄹ → ㅅ → ㄴ → ㄱ
② ㅂ → ㅁ → ㅅ → ㄷ → ㄹ → ㄱ → ㄴ
③ ㅂ → ㅁ → ㅅ → ㄹ → ㄷ → ㄴ → ㄱ
④ ㅂ → ㅅ → ㅁ → ㄷ → ㄹ → ㄴ → ㄱ
⑤ ㅂ → ㅅ → ㅁ → ㄹ → ㄷ → ㄱ → ㄴ

정답 2번

출제 포인트

이 문제는 온톨로지 구축에서 범위 설정부터 개별 인스턴스 생성까지의 표준적인 작업 순서를 묻는 문제이다.

올바른 순서

1. ㅂ 온톨로지 대상 도메인 및 범위 결정 — 무엇을 표현할지 경계를 정한다.
2. ㅁ 기존 온톨로지의 재활용 검토 — 이미 정의된 개념을 가져와 활용할 수 있는지 확인한다.
3. ㅅ 온톨로지 주요용어 수집 및 나열 — 도메인에 필요한 핵심 개념과 관계 후보를 모은다.
4. ㄷ 클래스와 클래스 계층 정의 — 수집한 용어를 개념 클래스로 조직하고 상하관계를 정한다.
5. ㄹ 클래스의 속성인 슬롯 정의 — 각 클래스가 가져야 할 속성과 관계를 정한다.
6. ㄱ 슬롯의 패싯 정의 — 슬롯의 값 유형, 허용 범위, 개수와 같은 제약을 정한다.
7. ㄴ 개별 인스턴스 생성 — 정의된 클래스와 슬롯 구조에 실제 개체를 입력한다.

왜 이 순서인가: 먼저 도메인과 재활용 가능성을 정한 뒤 용어를 수집해야 클래스 구조를 만들 수 있고, 슬롯을 정의해야 그 제약인 패싯을 정할 수 있으며, 완성된 스키마를 바탕으로 마지막에 인스턴스를 생성하기 때문이다.

NoteBase

선지별 해설

① X — 'ㄷ 클래스와 클래스 계층 정의' 뒤에 'ㅅ 주요용어 수집 및 나열'을 둔 점과 'ㄴ 인스턴스 생성' 뒤에 'ㄱ 슬롯의 패싯 정의'를 둔 점이 틀리다. 올바르게는 주요용어를 먼저 수집한 뒤 클래스 계층을 정의하고, 슬롯의 패싯을 정한 뒤 마지막에 인스턴스를 생성한다.

② O — ㅂ → ㅁ → ㅅ → ㄷ → ㄹ → ㄱ → ㄴ의 순서로, 범위 결정과 재활용 검토를 거쳐 용어·클래스·슬롯·패싯을 정의한 다음 인스턴스를 생성한다.

③ X — 'ㄹ 슬롯 정의'를 'ㄷ 클래스와 클래스 계층 정의'보다 앞에 둔 점과 'ㄴ 인스턴스 생성'을 'ㄱ 패싯 정의'보다 앞에 둔 점이 틀리다. 올바르게는 ㄷ → ㄹ → ㄱ → ㄴ이다.

④ X — 'ㅅ 주요용어 수집'을 'ㅁ 기존 온톨로지 재활용 검토'보다 앞에 둔 점과 'ㄴ 인스턴스 생성'을 'ㄱ 패싯 정의'보다 앞에 둔 점이 틀리다. 올바르게는 ㅂ → ㅁ → ㅅ으로 시작하고 ㄱ → ㄴ으로 끝난다.

⑤ X — 'ㅅ 주요용어 수집'을 'ㅁ 재활용 검토'보다 앞에 두고, 'ㄹ 슬롯 정의'를 'ㄷ 클래스 계층 정의'보다 앞에 둔 점이 틀리다. 올바르게는 ㅂ → ㅁ → ㅅ → ㄷ → ㄹ → ㄱ → ㄴ이다.

오답 포인트

출제자가 노린 함정은 서로 가까워 보이는 클래스·슬롯·패싯·인스턴스의 선후관계를 뒤섞는 것이다. 특히 슬롯을 먼저 정의한 뒤 패싯으로 그 슬롯의 제약을 정하며, 인스턴스는 이 구조가 모두 완성된 후 생성된다는 점이 핵심이다.

다시 나오면 이렇게 판단

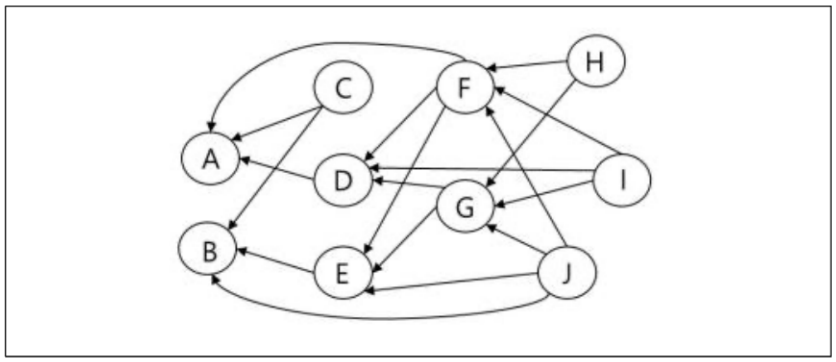
절차형 문제에서는 앞 단계의 산출물이 뒤 단계의 입력이 되는지를 확인한다. 온톨로지는 '범위 결정 → 재활용 검토 → 용어 수집 → 클래스 계층 → 슬롯 → 슬롯 제약 → 실제 인스턴스'의 흐름으로 판단한다.

암기 포인트

범위·재활용·용어를 먼저 정하고, 클래스·슬롯·패싯으로 틀을 만든 뒤 인스턴스를 넣는다.

NoteBase

11. 다음 인용문헌망에 의한 주제관계에 대한 설명이 옳지 않은 것은?



① 문헌 C와 E의 서지결합도는 1이다.

② 문헌 F와 G의 동시인용도는 문헌 C와 E의 동시인용도보다 크다.

③ 문헌 A와 B의 동시인용도는 문헌 D와 E의 동시인용도와 같다.

④ 문헌 F와 G의 서지결합도는 문헌 C와 D의 서지결합도보다 크다.

⑤ 문헌 A와 B의 동시인용도는 서지결합도와는 달리 고정되어 있지 않고 동적이다.

정답 3번

출제 포인트

이 문제는 인용문헌망의 화살표 방향을 읽고 서지결합도와 동시인용도를 구분하는 문제이다. 서지결합도는 두 문헌이 함께 인용한 문헌의 수이고, 동시인용도는 다른 문헌들이 두 문헌을 함께 인용한 횟수이다.

NoteBase

선지별 해설

① O
C는 A와 B를 인용하고 E는 B를 인용하므로 두 문헌이 공통으로 인용한 문헌은 B 하나이다. 따라서 C와 E의 서지결합도는 1이다.

② O
F와 G를 함께 인용한 문헌은 H, I, J의 3개인 반면, C와 E를 함께 인용한 문헌은 없다. 따라서 F와 G의 동시인용도가 C와 E의 동시인용도보다 크다.

③ X
A와 B를 함께 인용한 문헌은 C 하나이므로 동시인용도는 1이다. D와 E를 함께 인용한 문헌은 F와 G이므로 동시인용도는 2이다. '두 동시인용도가 같다'는 표현이 틀렸으며, 올바르게는 A와 B의 동시인용도가 D와 E의 동시인용도보다 작다.

④ O
F와 G가 공통으로 인용한 문헌은 D와 E의 2개이고, C와 D가 공통으로 인용한 문헌은 A 하나이다. 따라서 F와 G의 서지결합도가 C와 D의 서지결합도보다 크다.

⑤ O
서지결합도는 이미 발표된 두 문헌의 참고문헌 목록을 비교하므로 고정된다. 반면 동시인용도는 이후 발표되는 문헌이 A와 B를 함께 인용할 때마다 증가할 수 있으므로 동적이다.

오답 포인트

③은 공통으로 들어오는 화살표를 세지 않고 두 쌍의 값을 같다고 판단하게 만드는 선지이다. A와 B에는 C에서 출발한 공통 화살표 1개가 들어오지만, D와 E에는 F와 G에서 출발한 공통 화살표 2개가 들어온다.

다시 나오면 이렇게 판단

서지결합도는 두 문헌에서 출발하는 화살표의 공통 도착점을 세고, 동시인용도는 두 문헌으로 들어오는 화살표의 공통 출발점을 센다. 각 문헌의 공통 대상을 먼저 목록으로 적으면 방향을 혼동하지 않는다.

암기 포인트

서지결합은 '같이 인용한 과거 문헌'이라 고정적이고, 동시인용은 '나중에 함께 인용됨'이라 동적이다.

NoteBase

12. 참고과정의 '적합성(Pertinence)평가'에 대한 설명으로 옳은 것은?

① 검색어에 대한 검색결과의 적합성 평가
② 이용자의 정보요구에 대한 검색어의 적합성 평가
③ 이용자의 정보요구에 대한 검색전략의 적합성 평가
④ 정보검색에 소요된 시간과 비용에 대한 적합성 평가
⑤ 이용자의 정보요구에 대한 검색결과의 적합성 평가

정답 5번

출제 포인트

이 문제는 정보검색 평가에서 적합성과 적절성(Pertinence)이 각각 무엇과 무엇의 관계를 평가하는지 구별하는 문제이다.

핵심 개념

적절성(Pertinence): 검색결과가 이용자의 실제 정보요구를 얼마나 충족하는지를 평가하는 개념이다. 판단의 두 축은 '이용자의 정보요구'와 '검색결과'이다.
적합성(Relevance): 검색결과가 입력된 검색어 또는 질의와 얼마나 부합하는지를 평가하는 개념이다. 판단의 두 축은 '검색어'와 '검색결과'이다.
따라서 적합성은 질의 중심, 적절성은 이용자의 정보요구 중심으로 구별한다.

선지별 해설

① X — '검색어에 대한 검색결과의 적합성'은 검색어와 검색결과의 부합 정도를 뜻하는 적합성(Relevance)의 설명이다. 올바르게는 적절성을 '이용자의 정보요구에 대한 검색결과의 적합성'으로 표현한다.
② X — '이용자의 정보요구에 대한 검색어의 적합성'은 이용자의 요구가 검색어로 제대로 표현되었는지를 보는 관계이며, 적절성의 정의가 아니다. 올바르게는 검색어가 아니라 검색결과가 이용자의 정보요구에 부합하는지를 평가한다.
③ X — '이용자의 정보요구에 대한 검색전략의 적합성'은 검색전략 선택의 타당성을 평가하는 설명이다.
올바르게는 검색전략이 아니라 검색결과와 이용자의 정보요구 사이의 부합 정도를 평가한다.
④ X — '시간과 비용에 대한 적합성'은 검색 과정의 효율성 또는 경제성과 관련된 설명이다. 올바르게는 적절성을 시간·비용이 아니라 이용자의 정보요구와 검색결과의 관계로 설명한다.
⑤ O — 이용자의 정보요구에 대해 검색결과가 얼마나 부합하는지를 평가하므로 적절성(Pertinence)의 정확한 설명이다.

NoteBase

오답 포인트

①과 ⑤는 모두 '검색결과의 적합성'을 말하지만 비교 기준이 다르다. ①은 검색어를 기준으로 하므로 적합성(Relevance)이고, ⑤는 이용자의 실제 정보요구를 기준으로 하므로 적절성(Pertinence)이다.

다시 나오면 이렇게 판단

선지에서 검색결과의 비교 대상을 확인한다. 비교 대상이 검색어이면 적합성(Relevance), 이용자의 정보요구이면 적절성(Pertinence)으로 판단한다.

암기 포인트

검색어-검색결과는 적합성, 정보요구-검색결과는 적절성이다.

NoteBase

13. 식별자에 대한 설명으로 옳지 않은 것은?

① URL은 위치 의존적이다.

② DOI 접두어는 디렉코리 코드와 항목ID로 구성된다.

③ DOI는 접두어와 접미어로 구성되며, 이 둘은 슬래시(/)로 나누어진다.

④ URL에서는 자원에 접근하고자 하는 이용자의 상황이 고려되지 않는다.

⑤ OpenURL에서 연결대상이 되는 정보자원은 특정 메타데이터 집합이나 하나 이상의 고유한 식별자로 식별 가능하다.

정답 2번

출제 포인트

URL, DOI, OpenURL의 식별 방식과 DOI 접두어·접미어의 구성을 구별하는 문제이다.

핵심 개념

URL: 인터넷 자원의 위치를 나타내는 위치 의존적 식별자이다.
DOI: 디지털 객체에 부여하는 지속적 식별자로, 접두어와 접미어를 슬래시(/)로 구분한다. 접두어는 디렉터리 지시자 10과 등록자 코드로 구성되고, 접미어는 등록자가 정한 객체 식별 문자열이다.
OpenURL: 이용자의 소속기관·접근 권한 등 상황 정보를 반영하여 적절한 서비스나 원문으로 연결하는 문맥 인식형 링크 기술이다. 연결 대상 자원은 메타데이터 집합이나 고유 식별자로 표현할 수 있다.

선지별 해설

① O — URL은 자원의 네트워크상 위치를 주소로 나타내므로 위치가 바뀌면 주소도 달라질 수 있는 위치 의존적 식별자이다.
② X — 틀린 부분은 DOI 접두어가 '디렉터리 코드와 항목 ID'로 구성된다는 표현이다. 올바르게는 DOI 접두어는 디렉터리 지시자 10과 등록자 코드로 구성되며, 개별 객체를 구별하는 문자열은 접미어에 놓인다.
③ O — DOI는 접두어와 접미어로 구성되며, 두 부분은 슬래시(/)로 구분된다.
④ O — URL 자체는 자원의 위치를 표시할 뿐 이용자의 소속기관, 접근 권한 같은 상황 정보를 반영하지 않는다. 이러한 문맥을 반영한 연결은 OpenURL의 특징이다.
⑤ O — OpenURL에서 연결 대상 정보자원은 제목·저자 등 메타데이터의 집합이나 DOI 같은 하나 이상의 고유 식별자로 식별할 수 있다.

오답 포인트

②는 DOI 접두어와 접미어의 역할을 뒤섞었다. 등록자를 나타내는 정보는 접두어에, 등록자가 부여한 개별 객체 식별 문자열은 접미어에 들어간다.

NoteBase

다시 나오면 이렇게 판단

URL은 '어디에 있는가', DOI는 '어떤 객체인가', OpenURL은 '이 이용자에게 어디로 연결할 것인가'를 중심으로 판단한다. DOI의 슬래시 앞은 접두어인 '10+등록자 코드', 뒤는 접미어인 '객체 식별 문자열'로 구분한다.

암기 포인트

URL은 위치, DOI는 지속 식별, OpenURL은 이용자 상황 기반 연결이다. DOI는 10+등록자 코드 / 객체 식별 문자열이다.

NoteBase

14. 다음은 RDF를 이용하여 레코드를 작성한 사례의 일부이다. ㉠, ㉡, ㉢에 들어갈 용어로 옳은 것은?

```
<?xml version="1.0"?>
<rdf:RDF
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:nsp="https://nsp.nanet.go.kr/rdf/">
    <㉠ ㉡="https://nsp.nanet.go.kr/">
    <nsp:title>국가전략정보</nsp:title>
    <nsp:author>국회도서관</nsp:author>
    </㉢>
</rdf:RDF>
```

	㉠	㉡	㉢
①	rdf:Description	rdf:about	rdf:Description
②	rdf:Description	rdf:resource	rdf:Description
③	rdf:Description	rdf:subject	rdf:Description
④	rdf:Description	rdf:about	rdf:resource
⑤	rdf:about	rdf:resource	rdf:about

정답 1번

출제 포인트

RDF/XML에서 자원을 기술하는 요소와 그 자원의 URI를 지정하는 속성을 구별하는 문제이다.

핵심 개념

rdf:Description은 하나의 자원에 관한 설명을 묶는 요소이다. rdf:about은 그 설명의 주어가 되는 자원의 URI를 지정하는 속성이다. 따라서 시작 태그는 <rdf:Description rdf:about="https://nsp.nanet.go.kr/">이고, 종료 태그는 시작 태그와 같은 요소 이름을 사용한 </rdf:Description>이 된다.
rdf:resource는 속성값이 다른 자원의 URI임을 나타낼 때 속성 요소에 사용하는 속성이다. rdf:subject는 이 코드에서 설명 대상 URI를 지정하는 속성이 아니다.

NoteBase

선지별 해설

① O — ㄱ의 rdf:Description은 자원을 기술하는 요소이고, ㄴ의 rdf:about은 설명 대상 자원의 URI를 지정하며, ㄷ은 시작 요소와 짝을 이루는 rdf:Description이므로 올바르다.
② X — ㄴ의 rdf:resource가 틀린 표현이다. 실제로는 설명 대상 자원의 URI를 지정하는 rdf:about이 들어가야 한다.
③ X — ㄴ의 rdf:subject가 틀린 표현이다. 실제로는 <rdf:Description>의 주어 자원을 지정하는 rdf:about이 들어가야 한다.
④ X — ㄷ의 rdf:resource가 틀린 표현이다. XML 종료 태그에는 시작 태그와 같은 요소명인 rdf:Description이 들어가야 한다.
⑤ X — ㄱ과 ㄷ의 rdf:about이 틀린 표현이다. rdf:about은 요소명이 아니라 rdf:Description에 붙는 속성이므로 ㄱ과 ㄷ에는 각각 rdf:Description이 들어가야 한다. 또한 ㄴ에는 rdf:resource가 아니라 rdf:about이 들어가야 한다.

오답 포인트

rdf:about과 rdf:resource가 모두 URI와 관련된 속성이라는 점이 함정이다. 이 문항에서는 '무엇을 설명하는가'를 지정하므로 rdf:about을 사용하며, XML의 종료 태그는 속성이 아니라 시작 요소의 이름인 rdf:Description을 반복한다.

다시 나오면 이렇게 판단

먼저 시작 태그와 종료 태그의 요소명이 같은지 확인한다. 이어 설명의 주어 URI이면 rdf:about, 속성값으로 다른 자원을 가리키는 URI이면 rdf:resource로 판단한다.

암기 포인트

Description은 자원을 설명하고, about은 설명할 자원을 정하며, resource는 값으로 다른 자원을 가리킨다.

NoteBase

15. 텍스트 범주화에 대한 설명으로 옳지 않은 것은?

① 텍스트 범주화 기법을 구성하는 기본 요소는 문헌집단, 분류자질, 분류기이다.

② 범주화의 효율성을 높이기 위해 특정한 자질 선정기준을 적용하여 자질 집합의 크기를 줄인다.

③ 입력문헌과 모든 학습문헌 간의 유사도 산출결과에 따라 가장 유사한 학습문헌을 선택한다.

④ 각 문헌을 표현하는 특성들을 비교하여 문헌 간의 유사성을 측정하고, 비슷한 내용의 문헌들을 동일한 집단에 포함되도록 군집화하는 기법이다.

⑤ 텍스트 범주화 기법은 전자우편이나 뉴스기사 필터링, 뉴스기사 선정, 웹 문서 분류, 단어의 중의성 해소 등의 영역에서 다양하게 응용된다.

정답 4번

출제 포인트

이 문제는 미리 정해진 범주에 문헌을 배정하는 텍스트 범주화와, 내용이 비슷한 문헌끼리 집단을 새로 형성하는 텍스트 군집화의 차이를 묻는 문제이다.

핵심 개념

텍스트 범주화는 범주가 부여된 학습문헌을 이용하여 새 문헌을 미리 정해진 범주 중 하나에 분류하는 지도학습 기법이다. 기본적으로 문헌집단, 각 문헌을 수치로 나타낸 특성벡터, 범주를 결정하는 분류기가 필요하다.
텍스트 군집화는 미리 정한 범주나 정답표 없이 문헌 간 유사성을 바탕으로 비슷한 문헌끼리 집단을 형성하는 비지도학습 기법이다.
결정 차이축은 '기존 범주에 배정하는가'와 '유사한 문헌끼리 새로운 집단을 만드는가'이다.

NoteBase

선지별 해설

① O — 문헌집단은 학습 자료가 되고, 특성벡터는 문헌을 분류기가 처리할 수 있는 수치 형태로 표현하며, 분류기는 그 특성을 바탕으로 범주를 결정한다. 따라서 세 요소는 텍스트 범주화의 기본 구성 요소이다.
② O — 텍스트의 특성 차원은 매우 클 수 있으므로 분류에 유용한 특성만 고르는 특성 선택을 적용하면 특성 집합의 크기와 계산 부담을 줄여 범주화의 효율성을 높일 수 있다.
③ O — 입력문헌과 학습문헌들의 유사도를 구한 뒤 가장 유사한 학습문헌을 기준으로 범주를 판단하는 것은 최근접 이웃 방식, 특히 하나의 최근접 문헌을 이용하는 1-최근접 이웃 방식의 원리이다.
④ X — 틀린 표현은 '비슷한 내용의 문헌들을 동일한 집단에 포함되도록 군집화하는 기법'이다. 이는 미리 정해진 범주로 분류하는 텍스트 범주화가 아니라, 문헌 간 유사성에 따라 집단을 형성하는 텍스트 군집화의 설명이다. 올바르게는 '텍스트 범주화는 학습문헌에서 얻은 범주 정보를 이용하여 입력문헌을 미리 정해진 범주에 배정하는 기법'이다.
⑤ O — 전자우편이나 뉴스기사 필터링, 관심 뉴스기사 선정, 웹 문서 분류, 문맥에 따른 단어 의미 판별은 모두 텍스트의 특성을 이용해 정해진 부류나 의미 범주를 선택하는 응용에 해당한다.

오답 포인트

④의 '유사성을 측정한다'는 표현만 보면 범주화 기법처럼 보일 수 있다. 그러나 결정적인 함정은 '비슷한 문헌들을 동일한 집단에 포함되도록 군집화한다'는 부분이다. 유사도는 범주화에서도 활용될 수 있지만, 유사한 자료끼리 집단 자체를 형성하는 것은 군집화이다.

다시 나오면 이렇게 판단

문헌을 이미 정해진 범주에 배정하고 학습문헌의 범주 정보를 사용하면 범주화로 판단한다. 정답 범주 없이 문헌 간 유사성만으로 비슷한 문헌끼리 집단을 만들면 군집화로 판단한다.

암기 포인트

범주화는 '정해진 범주에 배정', 군집화는 '비슷한 것끼리 집단 형성'이다.

NoteBase

16. MODS에 대한 설명으로 옳은 것만을 <보기>에서 모두 고르면?

< 보 기 >

ㄱ. MODS의 주된 목적은 MARC를 대체하기 위한 것이다.

ㄴ. MODS에서 제시하는 상위요소, 하위요소, 속성 등은 모두 XML의 구문규칙에 따라서 구현된다.

ㄷ. 자연어에 기반한 데이터요소를 사용하고 있어 전문적인 지식이 없더라도 MODS를 이용하여 메타데이터 레코드를 생성할 수 있다.

ㄹ. 메타데이터뿐만 아니라 정보자원의 원문까지도 포함하여 기술할 수 있는 확장성을 지니고 있다.

ㅁ. 구문이 MARC 포맷보다는 복잡하지만 유연하고 명확하게 서지적 계층구조를 표현하고 있다.

① ㄱ, ㄴ, ㄷ

② ㄱ, ㄴ, ㅁ

③ ㄱ, ㄹ, ㅁ

④ ㄴ, ㄷ, ㄹ

⑤ ㄷ, ㄹ, ㅁ

정답 4번

출제 포인트

MODS의 개발 목적, XML 기반 구조, 자연어 태그, 디지털 자원 기술의 확장성, MARC와의 복잡도 차이를 구별하는 문제이다.

핵심 개념

MODS(Metadata Object Description Schema)는 MARC 21의 의미와 주요 요소를 바탕으로 만든 XML 기반 서지 메타데이터 스키마이다. 기존 MARC 데이터를 선택적으로 수용하면서 새로운 자원 기술 레코드도 만들 수 있으며, 숫자 태그 대신 의미를 알기 쉬운 자연어 기반 요소명을 사용한다. MARC보다 단순하고 유연하지만 Dublin Core보다는 풍부한 요소 집합을 제공한다.

NoteBase

선지별 해설

① X — ㄱ·ㄴ·ㄷ 가운데 ㄱ이 틀렸다. MODS는 MARC를 대체하려는 형식이 아니라 MARC의 주요 데이터를 XML 환경에서 활용하고 다른 메타데이터 형식을 보완하려는 스키마이다. 올바른 조합은 ㄴ·ㄷ·ㄹ이다.
② X — ㄱ·ㄴ·ㅁ 가운데 ㄱ과 ㅁ이 틀렸고, 옳은 ㄷ과 ㄹ이 빠졌다. MODS는 MARC 대체가 주목적이 아니며 MARC보다 복잡하지도 않다. 올바른 조합은 ㄴ·ㄷ·ㄹ이다.
③ X — ㄱ·ㄹ·ㅁ 가운데 ㄱ과 ㅁ이 틀렸고, 옳은 ㄴ과 ㄷ이 빠졌다. MODS는 MARC의 대체물이 아니라 MARC보다 단순한 XML 기반 스키마이다. 올바른 조합은 ㄴ·ㄷ·ㄹ이다.
④ O — ㄴ·ㄷ·ㄹ은 모두 옳다. MODS의 상위요소·하위요소·속성은 XML 스키마로 구현되고, 자연어 기반 태그를 사용해 숫자 태그 중심의 MARC보다 이해하기 쉬우며, 전자자원과 함께 패키징되거나 METS의 확장 스키마로 활용되는 등 원문 자원과 연계해 기술할 수 있는 확장성이 있다.
⑤ X — ㄷ과 ㄹ은 옳지만 ㅁ이 틀렸고, 옳은 ㄴ이 빠졌다. ‘MARC보다 복잡하다’가 아니라 ‘MARC보다 단순하고 유연하다’가 올바른 표현이다. 올바른 조합은 ㄴ·ㄷ·ㄹ이다.

오답 포인트

ㄱ의 ‘MARC를 대체’와 ㅁ의 ‘MARC보다 복잡’이 핵심 함정이다. MODS는 MARC의 의미를 이어받되 XML 환경에 맞게 단순화한 보완적 형식이므로, 대체 관계나 복잡도 방향을 거꾸로 판단하면 안 된다.

다시 나오면 이렇게 판단

MODS를 보면 ‘MARC 기반·XML 구현·자연어 태그·MARC보다 단순·디지털 자원과의 연계 확장성’의 다섯 기준으로 각 문장을 판별한다. ‘MARC 대체’ 또는 ‘MARC보다 복잡’이라는 표현은 틀린 진술로 판단한다.

암기 포인트

MODS는 MARC를 없애는 형식이 아니라, MARC를 XML로 더 단순하고 읽기 쉽게 활용하는 형식이다.

NoteBase

17. 디지털도서관 세대 구분에 대한 설명으로 옳지 않은 것은?

① 제1세대에는 도서관자동화시스템의 확장기능에 대한 요구가 강조되었다.

② 제2세대에는 다양한 메타데이터 변환에 대한 요구가 증가하였다.

③ 제2세대에는 더블린 코어 등 다양한 메타데이터 표준이 등장하였다.

④ 제3세대에는 분산된 디지털도서관들을 연동 및 통합하였다.

⑤ 제3세대에는 메타데이터 상호운용성을 향상하였다.

정답 3번

출제 포인트

디지털도서관의 발전을 제1세대의 개별 구축, 제2세대의 연계 확대, 제3세대의 분산 자원 통합과 상호운용성 강화로 구분할 수 있는지를 묻는 문제이다.

핵심 개념

제1세대: 기존 도서관자동화시스템을 바탕으로 디지털 자료를 다루는 기능을 확장하고 여러 메타데이터 표준이 등장한 단계이다.
제2세대: 서로 다른 디지털도서관과 메타데이터를 연계하면서 메타데이터 간 변환 요구가 커진 단계이다.
제3세대: 분산된 디지털도서관을 연동·통합하고 메타데이터 상호운용성을 높인 단계이다.

선지별 해설

① O — 제1세대에는 기존 도서관자동화시스템을 디지털 자료의 구축·검색·서비스에 맞게 확장하려는 요구가 강조되었다.
② O — 제2세대에는 서로 다른 시스템과 메타데이터 형식을 연계해야 했으므로 메타데이터 변환 요구가 증가하였다.
③ X — '제2세대에 더블린 코어 등 다양한 메타데이터 표준이 등장하였다'는 세대 연결이 잘못되었다.
올바르게는 더블린 코어 등 다양한 메타데이터 표준의 등장은 제1세대의 특징으로 본다. 제2세대의 핵심은 이미 다양해진 메타데이터 사이의 변환과 연계 요구가 커졌다는 점이다.
④ O — 제3세대에는 개별적으로 분산된 디지털도서관들을 네트워크로 연동하여 통합적으로 이용할 수 있게 하는 방향이 강조되었다.
⑤ O — 제3세대의 연동·통합을 위해서는 서로 다른 시스템이 메타데이터를 교환하고 이해할 수 있어야 하므로 메타데이터 상호운용성이 향상되었다.

NoteBase

오답 포인트

③은 '메타데이터 표준의 등장'과 '서로 다른 메타데이터의 변환'을 바꾸어 놓은 선지이다. 제1세대의 다양한 표준 등장과 제2세대의 변환 요구 증가를 구별해야 한다.

다시 나오면 이렇게 판단

세대별 중심 과제를 연결해 판단한다. 제1세대는 기능 확장과 표준의 등장, 제2세대는 이질적인 메타데이터의 변환과 연계, 제3세대는 분산 디지털도서관의 통합과 상호운용성이다.

암기 포인트

1세대는 확장·표준 등장, 2세대는 변환, 3세대는 통합·상호운용이다.

NoteBase

18. 다음에서 설명하는 분류기에 해당하는 것은?

> ○ 문헌 내에서 각 단어의 출현은 상호 독립적이라고 가정한다.
> ○ 학습문헌 집합 내 분류대상 문헌들에서 출현한 단어가 포함되어 있지 않으면 문제가 발생한다.
> ○ 학습문헌에 나타난 각 단어가 특정한 범주를 대표할 확률을 계산한 다음, 분류대상인 입력문헌에 출현한 단어들을 단서어로 삼아 문헌의 범주를 예측한다.

① 나이브 베이즈 분류기(Naive Bayes Classifier)
② 선형 분류기(Linear Classifier)
③ 신경망 분류기(Neural Network Classifier)
④ 의사결정나무 분류기(Decision Tree Classifier)
⑤ SVM 분류기(Support Vector Machine Classifier)

정답 1번

출제 포인트

문헌의 단어 출현을 서로 독립으로 가정하고, 범주별 단어 출현확률로 문헌의 범주를 예측하는 분류기를 식별하는 문제이다.

핵심 개념

나이브 베이즈 분류기는 베이즈 정리를 바탕으로 각 특징이 범주가 주어졌을 때 서로 조건부 독립이라고 단순화하고, 범주별 사후확률을 비교하여 분류한다. 문헌 분류에서는 단어 출현을 특징으로 삼아 범주별 단어 확률을 결합한다. 학습문헌의 특정 범주에 한 번도 나오지 않은 단어는 추정확률이 0이 되어 전체 확률을 0으로 만들 수 있으므로, 보통 라플라스 평활화를 적용한다.
나이브 베이즈는 확률모형, 선형 분류기는 선형 결정경계, 신경망은 층과 가중치의 비선형 학습, 의사결정나무는 특징별 분기 규칙, SVM은 클래스 사이의 마진을 최대화하는 결정경계를 핵심으로 한다.

NoteBase

선지별 해설

① O — 단어 출현의 독립 가정, 범주별 단어 출현확률 계산, 입력문헌의 단어들을 이용한 범주 예측은 나이브 베이즈 문헌 분류의 전형적인 특징이다. 또한 학습문헌의 특정 범주에 없던 단어 때문에 확률의 곱이 0이 되는 문제도 나이브 베이즈에서 발생한다.
② X — 틀린 부분은 '선형 분류기'이다. 선형 분류기는 특징의 가중합으로 선형 결정경계를 구성한다. 제시문처럼 단어의 조건부 독립을 가정하고 범주별 단어 확률을 결합하는 분류기는 나이브 베이즈 분류기이다.
③ X — 틀린 부분은 '신경망 분류기'이다. 신경망 분류기는 여러 층의 뉴런과 가중치를 학습하여 입력과 출력의 관계를 모델링한다. 제시문의 독립 가정과 미출현 단어의 영확률 문제는 나이브 베이즈 분류기의 특징이다.
④ X — 틀린 부분은 '의사결정나무 분류기'이다. 의사결정나무 분류기는 특징값에 관한 질문을 노드에 배치하고 분기를 반복해 범주를 결정한다. 제시문처럼 범주별 단어 출현확률을 계산해 분류하는 방식은 나이브 베이즈 분류기이다.
⑤ X — 틀린 부분은 'SVM 분류기'이다. SVM 분류기는 클래스 사이의 마진이 최대가 되는 초평면을 찾는 방식이다. 단어의 독립 가정 아래 확률을 결합하는 제시문의 방식은 나이브 베이즈 분류기이다.

오답 포인트

'입력문헌의 단어를 이용해 범주를 예측한다'는 표현만 보면 여러 텍스트 분류기가 모두 가능해 보인다. 그러나 '각 단어의 출현은 상호 독립'이라는 가정과 '학습문헌의 특정 범주에 단어가 없으면 문제가 발생한다'는 문구는 나이브 베이즈의 조건부 독립 가정과 영확률 문제를 직접 가리킨다.

다시 나오면 이렇게 판단

단어별 범주 조건부확률을 곱하고 특징의 독립을 가정하면 나이브 베이즈로 판단한다. 미출현 단어 때문에 전체 확률이 0이 된다는 단서가 있으면 라플라스 평활화가 필요한 나이브 베이즈 문제로 연결한다.

암기 포인트

독립 가정·확률의 곱·영확률 문제는 나이브 베이즈이고, 영확률 문제의 대표적 보완은 라플라스 평활화이다.

NoteBase

19. 다음 복수문헌의 요약 작성과정을 순서대로 바르게 나열한 것은?

> ㄱ. 중복성을 제거한다.
> ㄴ. 요약문을 생성하거나 시각적으로 표현한다.
> ㄷ. 유사성 척도를 사용하여 텍스트 요소들을 비교한다.
> ㄹ. 복수의 문헌들로부터 추출할 텍스트 요소를 식별한다.
> ㅁ. 대표되는 텍스트 요소(문헌, 문장, 문단 등)를 선정한다.

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

정답 2번

출제 포인트

복수 문헌을 요약할 때 텍스트 요소를 추출하고 비교한 뒤 대표 요소를 중심으로 중복을 제거하여 요약문을 만드는 절차의 선후관계를 묻는 문제이다.

올바른 순서

1. ㄹ 텍스트 요소 식별 — 복수의 문헌에서 요약에 사용할 문헌, 문장, 문단 등의 요소를 추출한다.
2. ㄷ 유사성 비교 — 추출한 텍스트 요소들을 유사성 척도로 비교한다.
3. ㅁ 대표 요소 선정 — 비교 결과를 바탕으로 대표가 되는 텍스트 요소를 고른다.
4. ㄱ 중복성 제거 — 대표 요소를 중심으로 겹치는 내용을 제거한다.
5. ㄴ 요약문 생성·시각화 — 정리된 내용을 최종 요약문으로 만들거나 시각적으로 표현한다.

왜 이 순서인가: 유사성을 비교하려면 먼저 비교할 텍스트 요소가 식별되어야 하고, 대표 요소를 정하려면 유사성 비교 결과가 필요하다. 또한 중복이 제거된 뒤에야 간결한 최종 요약을 생성할 수 있다.

NoteBase

선지별 해설

① X — ㄷ으로 시작하여 비교 대상인 텍스트 요소를 식별하기 전에 유사성을 비교하고, ㅁ 뒤에 ㄹ을 두어 대표 요소를 선정한 후 텍스트 요소를 식별하는 순서가 된다. 올바르게는 ㄹ → ㄷ → ㅁ → ㄱ → ㄴ이다.
② O — ㄹ에서 텍스트 요소를 식별하고, ㄷ에서 유사성을 비교한 다음, ㅁ에서 대표 요소를 선정한다. 이어 ㄱ에서 중복을 제거하고 ㄴ에서 최종 요약을 생성하거나 시각화하므로 절차가 모두 맞다.
③ X — ㄹ 다음에 ㅁ을 두어 유사성 비교 전에 대표 요소를 선정한다는 점이 틀리다. 올바르게는 ㄹ → ㄷ → ㅁ → ㄱ → ㄴ이다.
④ X — ㅁ으로 시작하여 텍스트 요소의 식별과 유사성 비교 전에 대표 요소를 선정하고, ㄴ의 요약 생성도 중복 제거보다 앞에 둔 점이 틀리다. 올바르게는 ㄹ → ㄷ → ㅁ → ㄱ → ㄴ이다.
⑤ X — ㅁ으로 시작하여 식별과 비교 전에 대표 요소를 선정하고, ㄴ의 요약 생성도 중복 제거보다 앞에 둔 점이 틀리다. 올바르게는 ㄹ → ㄷ → ㅁ → ㄱ → ㄴ이다.

오답 포인트

가장 큰 함정은 '대표 요소 선정'과 '유사성 비교'의 순서를 바꾸는 것이다. 대표성을 판단하려면 먼저 추출된 요소 사이의 유사성을 비교해야 한다. 또한 요약문 생성·시각화는 중복 내용을 제거한 뒤 수행하는 최종 단계이다.

다시 나오면 이렇게 판단

절차형 문제는 각 단계의 입력과 산출물을 연결해 판단한다. 비교에는 식별된 요소가 필요하고, 대표 선정에는 비교 결과가 필요하며, 최종 요약에는 중복이 제거된 내용이 필요하므로 '식별 → 비교 → 선정 → 정제 → 생성'의 흐름이 된다.

암기 포인트

식별하고 비교한 뒤 대표를 선정하며, 중복을 제거한 다음 요약을 생성한다: 식-비-대-중-요.

NoteBase

20. 다음 설명에 대한 디지털도서관 구축기술을 바르게 연결한 것은?

> ㄱ. 자동인식기술의 하나로 IC칩을 내장한 태그를 부착하고 태그에 내장된 정보를 무선주파수를 통해 비접촉식으로 자동인식, 처리, 활용하는 기술이다.
>
> ㄴ. 현실의 이미지와 가상의 이미지를 결합한 것이고, 실시간으로 상호 작용이 가능한 것이며, 3차원의 공간 안에 놓인 것으로서 무한한 정보들을 현실 속으로 가져와 원하는 시간과 장소에서 활용할 수 있다.
>
> ㄷ. 각종 IT자원을 언제, 어디서나 필요에 따라 가변적으로 제공하고, 네트워크에 연결된 디바이스를 통해 통일한 콘텐츠 및 서비스를 제공받을 수 있게 한다.

	ㄱ	ㄴ	ㄷ
①	NFC	가상현실	클라우드 컴퓨팅
②	NFC	증강현실	상황인식기술
③	RFID	가상현실	클라우드 컴퓨팅
④	RFID	증강현실	상황인식기술
⑤	RFID	증강현실	클라우드 컴퓨팅

정답 5번

출제 포인트

디지털도서관 구축기술의 설명을 RFID, 증강현실, 클라우드 컴퓨팅의 핵심 특징과 정확히 연결하는 문제이다.

핵심 개념

RFID: IC칩이 내장된 태그의 정보를 무선주파수로 비접촉 인식하는 기술이다. 키워드는 태그, 무선주파수, 자동인식이다.
증강현실: 현실 세계에 가상 정보를 결합하여 실시간 상호작용이 가능하게 하는 기술이다. 키워드는 현실과 가상의 결합이다.
클라우드 컴퓨팅: 네트워크를 통해 서버·저장공간·소프트웨어 등의 IT 자원을 필요에 따라 제공받아 이용하는 방식이다. 키워드는 네트워크 기반 자원 제공과 확장성이다.
NFC는 가까운 거리에서 기기 간 통신을 수행하는 근거리 무선통신 기술이고, 가상현실은 사용자를 가상 환경에 몰입시키는 기술이며, 상황인식기술은 사용자의 위치·시간·환경 등의 맥락을 파악하여 적절한 서비스를 제공하는 기술이다.

NoteBase

선지별 해설

① X
ㄷ의 클라우드 컴퓨팅은 맞지만, ㄱ의 'NFC'와 ㄴ의 '가상현실'이 틀렸다. ㄱ은 IC칩 태그를 무선주파수로 자동인식하므로 RFID이고, ㄴ은 현실 이미지와 가상 이미지를 결합하므로 증강현실이다. 따라서 올바른 연결은 RFID－증강현실－클라우드 컴퓨팅이다.

② X
ㄴ의 증강현실은 맞지만, ㄱ의 'NFC'와 ㄷ의 '상황인식기술'이 틀렸다. ㄱ은 RFID이며, ㄷ은 IT 자원을 필요에 따라 제공하고 네트워크 연결 기기에 동일한 콘텐츠와 서비스를 제공하므로 클라우드 컴퓨팅이다. 따라서 올바른 연결은 RFID－증강현실－클라우드 컴퓨팅이다.

③ X
ㄱ의 RFID와 ㄷ의 클라우드 컴퓨팅은 맞지만, ㄴ의 '가상현실'이 틀렸다. 현실의 이미지와 가상의 이미지를 결합하여 실시간 상호작용을 가능하게 하는 기술은 증강현실이다. 따라서 ㄴ을 증강현실로 고쳐야 한다.

④ X
ㄱ의 RFID와 ㄴ의 증강현실은 맞지만, ㄷ의 '상황인식기술'이 틀렸다. IT 자원을 필요에 따라 가변적으로 제공하고 네트워크를 통해 콘텐츠와 서비스를 이용하게 하는 기술은 클라우드 컴퓨팅이다. 따라서 ㄷ을 클라우드 컴퓨팅으로 고쳐야 한다.

⑤ O
ㄱ은 태그와 무선주파수를 이용하는 RFID, ㄴ은 현실과 가상을 결합하는 증강현실, ㄷ은 네트워크를 통해 IT 자원과 서비스를 제공하는 클라우드 컴퓨팅에 각각 해당한다. 세 연결이 모두 정확하다.

오답 포인트

ㄱ의 '비접촉 무선통신'만 보고 NFC를 고르거나, ㄴ의 '가상의 이미지'만 보고 가상현실을 고르기 쉽다. 그러나 ㄱ의 결정적 표현은 'IC칩 내장 태그의 자동인식'이고, ㄴ의 결정적 표현은 '현실과 가상의 결합'이다. ㄷ도 단순히 주변 상황을 인식하는 기술이 아니라 IT 자원을 네트워크로 제공하는 기술이라는 점이 핵심이다.

다시 나오면 이렇게 판단

설명의 결정적 키워드를 먼저 찾는다. '태그＋무선주파수＋자동인식'이면 RFID, '현실＋가상＋실시간 상호작용'이면 증강현실, 'IT 자원＋필요에 따른 제공＋네트워크 서비스'이면 클라우드 컴퓨팅으로 판단한다.

암기 포인트

태그는 RFID, 현실 위 가상은 증강현실, 빌려 쓰는 IT 자원은 클라우드 컴퓨팅이다.

NoteBase