【문 1】 다음 <보기>에서 다크웹(Dark Web)에 대한 설명으로 옳은 것을 모두 고른 것은?

< 보 기 >
ㄱ. 본래 미 해군 연구소에서 익명성을 목적으로 만들어진 기술이다.
ㄴ. 익명성 보장은 물론 IP주소 추적이 불가능하도록 고안된 인터넷 영역이다.
ㄷ. 일반적인 검색엔진으로는 찾을 수 없기 때문에 해킹으로 얻은 개인정보, 경쟁사의 영업비밀 등 주로 불법적인 정보가 거래되는 웹이다.
ㄹ. 웹페이지를 찾아다니는 웹 크롤러에 의해 수집되지 않아 검색 등으로는 접근이 어려운 웹이다.

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

【문 2】 정보검색시스템의 성능 평가를 위한 대표적인 척도로 재현율과 정확률이 있다. 재현율과 정확률은 일반적으로 반비례 관계에 있으며 시스템의 검색 효율을 판단하기 위해서는 함께 사용된다. 하지만 두 개 이상의 시스템 성능을 비교할 경우 어느 시스템 성능이 더 나은지를 판단하기 어렵기 때문에 F-척도와 같은 단일 척도가 개발되기도 했다.

가상적으로 어떤 검색시스템의 재현율이 50%, 정확률이 75%라고 한다면 해당 시스템의 F-척도값은 다음 중 어떤 것인가?

① 0.45 ② 0.50 ③ 0.60 ④ 0.75

【문 3】 디지털도서관의 발전에 따라 사서의 도서관 자료와 관련된 저작권에 대한 이해와 적용이 요구된다. 공정이용(Fair Use)은 저작권자의 배타적인 권리를 제한하는 규정으로 1976년 미국 저작권법에 의해 성문화된 바 있다. 공정이용은 저작권법의 침해를 받지 않는 언론보도, 교육, 연구 등을 목적으로 하는 저작물의 이용을 뜻하며 전통적으로 도서관 자료 서비스의 근거가 되어왔다.

다음 중 공정이용을 결정하는 요건으로 보기 힘든 것은?

① 사용의 목적과 성격
② 저작권이 있는 저작물의 성격
③ 공정이용과 관련된 정보의 분량
④ 저작자의 허락 또는 승인

【문 4】 벡터 스페이스 모델(Vector Space Model, VSM)에 관한 다음 설명 중 가장 옳지 않은 것은?

① 벡터 스페이스 모델은 문서와 쿼리를 벡터로 표현하여 유사도를 계산하는 모델이다.
② 벡터 스페이스 모델에서 문서와 쿼리는 단어의 빈도를 기반으로 벡터 공간에 위치하게 된다.
③ 벡터 스페이스 모델에서 문서와 쿼리 간의 유사도는 주로 코사인 유사도(Cosine Similarity)로 측정된다.
④ 벡터 스페이스 모델은 문서 간의 순서나 구조를 고려하여 문서의 의미를 파악하는 데 중점을 둔다.

【문 5】 다음 중 정보 사이클에서 분류, 편목, 색인 초록 작업을 수행하는 단계는 무엇인가?

① 정보생산 ② 정보조직 ③ 정보축적 ④ 정보검색

【문 6】 용어-문헌 인접 행렬(Term-Document Adjacency Matrix)에 관한 다음 설명 중 가장 옳지 않은 것은?

① 용어-문헌 인접 행렬은 문서 컬렉션에서 각 문서에 등장하는 각 용어의 빈도를 행렬 형태로 나타낸 것이다.
② 행렬의 행은 문서를, 열은 용어를 나타내며, 셀의 값은 해당 문서에서 해당 용어의 빈도이다.
③ 용어-문헌 인접 행렬은 문서 간의 유사도를 계산하기 위한 기초 자료로 사용될 수 있다.
④ 용어-문헌 인접 행렬은 문서가 추가되거나 삭제될 때마다 행렬을 갱신해야 하는 단점이 있다.

【문 7】 윌슨(Wilson)의 정보행위모형에 관한 다음 설명 중 가장 옳지 않은 것은?

① 심리학 등 다른 분야의 이론을 접목하여 정보추구의 복잡한 맥락을 강조하였다.
② 스트레스 · 대처이론을 이용하여 왜 어떤 정보원은 다른 정보원보다 더 많이 이용되는지 설명하고자 하였다.
③ 사회학습이론을 이용하여 왜 어떤 목표는 성취하고 다른 목표는 성취하지 못하는지 설명하고자 하였다.
④ 이용자로 하여금 정보추구과정을 시작하게 하는 동기요인인 촉진 메카니즘을 제시하였다.

【문 8】 인터넷에 연결된 모든 기기는 고유의 IP 주소를 가지고 있다. IP 주소 체계에 관한 다음 설명 중 가장 옳지 않은 것은?

① 현재 사용되고 있는 IPv4는 점(.)으로 구분된 4개의 옥텟(Octet)으로 구성된 32비트 주소 체계로 표현된다.
② 차세대 인터넷 주소 체계인 IPv6는 IPv4의 제한된 주소 공간을 해소하기 위해 개발되었고, 256비트의 길이로 주소 공간이 확대되었다.
③ 현재 IPv4와 IPv6가 혼용되고 있다.
④ IPv6는 인터넷 표준으로 채택되었다.

【문 9】 디지털도서관에서는 도서관 내외의 정보자료를 이용자들이 보다 손쉽게 이용할 수 있도록 다양한 이용자 인증 방법을 활용하고 있다.

다음 중 오픈소스 기반의 이용자 인증 솔루션으로 기관 간의 인증을 통해 기관이 자신의 이용자가 정보를 안정적으로 교환할 수 있도록 하는 프로토콜은 무엇인가?

① LDAP(Lightweight Directory Access Protocol)
② DRM(Digital Rights Management)
③ PKI(Public Key Infrastructure)
④ Shibboleth

【문10】 네트워크 분석은 구조화를 통하여 개체 간 상호관계를 체계적으로 표현하고, 관계성을 파악하기 위한 분석 방법으로 정보학 분야에서 다양하게 활용되고 있다.

다음 중 네트워크에서 중심에 위치하는 정보를 표현하는 중심성 지표가 아닌 것은?

① 연결중심성
② 매개중심성
③ 근접중심성
④ 포괄중심성

【문11】 데이터사이언스의 체계에 관한 다음 설명 중 가장 옳지 않은 것은?

① 데이터 관리영역과 기계학습영역으로 구분할 수 있다.
② 데이터사이언스 프로세스에는 사람이 개입하여 통제와 조절이 가능하다.
③인공지능의 구현은 지도학습 방식의 기계학습에 의존한다.
④대용량, 다양성, 빠른 속도의 특징을 가지는 빅데이터를 포함한다.

【문12】 다음 <보기>에서 메타 검색엔진에 대한 설명으로 옳은 것을 모두 고른 것은?

< 보 기 >
ㄱ. 자체 검색기능 없이 여러 개의 검색엔진을 연계하여 이를 검색하여 보여준다.
ㄴ. 많은 양의 정보 검색이 가능하다.
ㄷ. 데이터의 중복도가 적다.
ㄹ. 검색 속도가 빠르다.

① ㄱ, ㄴ ② ㄱ, ㄷ ③ ㄴ, ㄷ ④ ㄷ, ㄹ

【문13】 다음 <보기>에서 생성형 인공지능에 대한 설명으로 옳은 것을 모두 고른 것은?

< 보 기 >
ㄱ. 머신러닝 알고리즘을 활용한다.
ㄴ. 학습 데이터를 기반으로 유사한 콘텐츠를 새롭게 만들어내는 기술을 가리킨다.
ㄷ. 대형 언어모델이나 이미지 생성 모델을 활용한다.
ㄹ. 이미지 생성 모델로는 AE, GAN, BERT 등이 있다.

① ㄱ ② ㄱ, ㄴ ③ ㄱ, ㄴ, ㄷ ④ ㄱ, ㄴ, ㄷ, ㄹ

【문14】 다음 <보기>에 해당하는 기술은 무엇인가?

< 보 기 >
- 거대언어모델에 쌓인 데이터와 별개 외부 데이터를 이용해 답변의 정확도를 높여주는 기술이다.
-거대언어모델의 단점인 환각현상을 보완하는 역할을 한다.
-이용자 질문인 프롬프트에 맞는 정확한 정보를 제공한다.

① 자연어 기반 코드 생성 모델
② 언어 생성 모델
③ 검색 증강 생성
④ 트랜스포머 모델

【문15】 과학기술정보, 연구데이터, 정보분석서비스 및 연구인프라를 연계·융합하여 연구자가 필요로 하는 지식인프라를 제공하는 한국과학기술정보연구원(KISTI)의 서비스는 무엇인가?

① ScienceON
② DataON
③ AccessON
④ NTIS

【문16】 문헌 클러스터링(Document Clustering)에 관한 다음 설명 중 가장 옳지 않은 것은?

① 문헌 클러스터링은 유사한 주제나 내용을 가진 문서들을 그룹화하는 기법이다.
② 문헌 클러스터링은 주로 텍스트 마이닝과 정보 검색에서 사용된다.
③ 문헌 클러스터링 기법에는 K-평균 클러스터링(K-means clustering)과 계층적 클러스터링(Hierarchical Clustering)이 포함된다.
④ 문헌 클러스터링은 클러스터의 수를 사전에 지정하지 않고 자동으로 결정된다.

【문17】 새로운 창작과 합법적인 정보의 공유를 촉진시키려는 목적으로 만들어진 크리에이티브 커먼즈 라이선스(Creative Commons License)가 인터넷에서 활용되고 있다.
다음 중 [저작자표시-변경금지]를 나타내는 크리에이티브 커먼즈 라이선스 유형을 올바르게 표시한 것은?

① CC BY-NC
② CC BY-ND
③ CC BY-SA
④ CC BY-NC-SA

【문18】 크롤러(Crawler)에 관한 다음 설명 중 가장 옳지 않은 것은?

① 크롤러는 주기적으로 검색된 정보를 데이터베이스화하는 역할을 담당한다.
② 크롤러는 웹상의 문서나 이미지, 영상 등을 수집하는 데 사용되는 검색 엔진 자체이다.
③크롤러는 정보를 수집하기 위해 자동화된 방식으로 제어된다.
④ 크롤러는 봇(Bot)이라고도 불리며, 웹상의 정보를 자동으로 검색하고 수집하는 프로그램이다.

【문19】 다음 중 IF(영향력 지수)를 계산하는 식으로 가장 올바른 것은?

① $IF = \frac{\text{논문이 전체 인용된 횟수}}{\text{해당 저널에 발표된 전체 논문의 수}}$

② $IF = \frac{\text{해당 저널에 발표된 논문의 수}}{\text{논문이 전체 인용된 횟수}}$

③ $IF = \frac{\text{지난 2년간 해당 저널에 발표된 논문의 수}}{\text{지난2년간 해당 저널에 발표된 논문의 인용 횟수}}$

④ $IF = \frac{\text{지난2년간 해당 저널에 발표된 논문의 인용 횟수}}{\text{지난 2년간 해당 저널에 발표된 논문의 수}}$

【문20】 다음 중 2차 정보원에 해당하지 않는 것은?

① 서지(Bibliography)
② 데이터베이스(Database)
③ 학술잡지(Scholarly Journal)
④ 색인(Index)

【문21】 더빈(Dervin)의 의미형성(Sense-making)이론의 세 개의 핵심 개념이 바르게 짝지어진 것은?

① 상황 - 이용·도움 - 격차
② 상황 - 이용·도움 - 탐색
③ 정보 - 이용·도움 - 격차
④ 정보 - 이용·도움 - 탐색

【문22】 다음 중 더블린 코어 메타데이터 요소와 해당 요소에 대응하는 인코딩 스킴(Encoding Scheme)의 연결이 올바르지 않은 것은?

① date - W3C Date and Time Formats(W3CDTF)
② format - IMT
③ identifier - IP
④ subject - LCSH, MeSH

【문23】 오픈액세스 추진방법에 관한 다음 설명 중 가장 옳지 않은 것은?

① 골드 OA는 저자가 논문출판비용을 부담한다.
② 실버 OA는 저자가 학술논문을 디지털 저장소에 셀프 아카이빙한다.
③ 골드 OA는 출판 후 원문을 즉시 이용 할 수 있다.
④ 학술지 원문은 PDF 이외에 기계가독형 포맷으로 생산·공개되는 추세이다.

【문24】 브라우징 검색에 관한 다음 설명 중 가장 옳지 않은 것은?

① 브라우징 검색은 주제가 명확히 정의되지 않았을 때 정보를 얻기 위해 활용한다.
② 브라우징 검색은 명확하게 설명하기 어려운 정보를 찾기 위해 활용한다.
③ 브라우징 검색은 적합정보와 부적합정보가 섞여있는 것으로부터 올바른 정보를 선택하기 위해 활용한다.
④ 브라우징 검색은 정보의 '우연한 발견(Serendipity)'이 가능하지 않는다.

【문25】 다음 표는 정보자원의 기술에 중점을 두는 전통적인 메타데이터와, 다양한 분야에서 생성되는 여러 가지 형태의 정보자원을 체계화된 구조를 통해 상세하게 기술하여 보다 종합적인 정보활동을 지원하는 차세대 메타데이터의 특징을 비교한 것이다. 빈 칸에 들어갈 내용으로 바르게 짝지어진 것은?

구 분	전통적메타데이터	차세대메타데이터
해석의다양성	객관적	주관적
발생	정적	동적
요소의 비중	ㄱ	ㄴ
유연성	ㄷ	ㄹ
인코딩 기술	XML 기반	RDF 기반
콘텐츠 관리	중앙집중식	분산식

	ㄱ	ㄴ	ㄷ	ㄹ
①	자원중심적	상황중심적	구문의존적	의미의존적
②	상황중심적	자원중심적	구문의존적	의미의존적
③	자원중심적	상황중심적	의미의존적	구문의존적
④	상황중심적	자원중심적	의미의존적	구문의존적