

中國鋼鐵股份有限公司 104 年新進人員甄試(II)試題

甄試職位／類組【代碼】：師級職位／圖書資訊【H9611】

專業科目：1.資訊系統與資訊檢索 2.參考資源與服務 3.資訊服務加值運用

*請填寫入場通知書編號：_____

注意：①作答前須檢查答案卡、入場通知書號碼、座位標籤號碼、甄試類別是否相符，如有不同應立即請監試人員處理，否則不予計分。

②本試卷一份共 8 頁，測驗題型為【四選一單選選擇題 30 題，每題 1.5 分，複選題 22 題，每題 2.5 分】，限用 2B 鉛筆在「答案卡」上作答，請選出最適當答案，全部答對才給分，答錯不倒扣；未作答者，不予計分。

③請勿於答案卡上書寫姓名、入場通知書號碼或與答案無關之任何文字或符號。

④本項測驗僅得使用簡易型電子計算器(不具任何財務函數、工程函數功能、儲存程式功能)，但不得發出聲響；若應考人於測驗時將不符規定之電子計算器放置於桌面或使用，經勸阻無效，仍執意使用者，該節以零分計；該電子計算器並由監試人員保管至該節測驗結束後歸還。

⑤答案卡務必繳回，未繳回者該科以零分計算。

壹、四選一單選選擇題 30 題（每題 1.5 分，答錯不倒扣；未作答者，不予計分）

【3】1.下列何者可以用來評估有序的檢索結果(ranked output)？

- ①專指性(specificity)
- ②準確性(accuracy)
- ③平均求準率(average precision)
- ④求全性(recall)

【2】2.根據齊夫定律(Zipf's Law)，若詞頻最高的詞佔了所有詞彙出現次數的 7%，則詞頻次高的詞將佔所有詞彙出現次數的：

- ① 2%
- ② 3.5%
- ③ 4%
- ④ 5.5%

【2】3. XML（可擴展標記語言）的英文全稱為：

- ① eXpendable Markup Language
- ② eXtensible Markup Language
- ③ eXclusive Markup Language
- ④ eXemplar Markup Language

【3】4.一個詞的詞項頻率(term frequency) $tf_{t,d}$ 所代表的是：

- ①該詞在整個文件集的出現頻率
- ②該詞在相關文件的出現頻率
- ③該詞在某一文件中的出現頻率
- ④該詞出現的文件數

【1】5.一個詞的文件頻率倒數（Inverse Document Frequency，又譯「逆向文件頻率」）可以用來判斷該詞的：

- ①專指性(specificity)
- ②窮盡性(exhaustivity)
- ③可得性(availability)
- ④中心性(centrality)

【1】6.根據 tf-idf 權重計算，下列哪一個詞的權重最高？

- ①當該詞只在少數文件中多次出現時
- ②當該詞在一篇文件中但是出現次數很少時
- ③當該詞在很多文件中出現
- ④當該詞在所有文件中都出現

【3】7.下列何者不是建構反向索引檔(inverted index)的步驟？

- ①詞條化(tokenization)
- ②詞條正規化(normalization)
- ③編碼轉換(encoding)
- ④去除停用詞(stop words removal)

【1】8.檢索評估中融合了求準率(precision)和求全率(recall)的調和平均值指標稱為：

- ① F measure
- ② T measure
- ③ P measure
- ④ X measure

【2】9.根據 PageRank，下列哪一網頁會被給予較高的排名？

- ①網頁連結出許多鏈結(outlink)到權威網頁
- ②網頁收到許多來自權威網頁鏈結(inlink)，而這些網頁連出的鏈結數很低
- ③網頁收到許多來自權威網頁的鏈結，且這些網頁連出的鏈結數很高
- ④網頁連出及收到的鏈結數很高

【2】10.系統評估中所使用的 ROC curve (Receiver Operating Characteristics)是根據正確正例率 (true positive，或譯真陽性) 及下列何者所繪製成的？

- ①正確負例率(true negative rate)
- ②錯誤正例率(false positive rate)
- ③錯誤負例率(false negative rate)
- ④以上皆非

【2】11.假設某系統根據一資訊需求所給予的前 10 筆結果的相關性判定如下 (左邊結果排名較高，R 代表相關，N 代表不相關)

RNNRNRNNNR

求其 AP (average precision)為：

- ① 2.2
- ② 2.4
- ③ 3.6
- ④ 4.4

【2】12. NDCG (Normalized Discounted Cumulative Gain，譯作正規化折損累積增益) 近年來常被用於搜尋引擎的檢索評估，其「折損」累積增益的考量為：

- ①文件的增益(gain)應扣除重複性
- ②文件的增益應考慮其排名
- ③文件的增益應考慮其新穎性
- ④文件的增益應考慮其權威性

- 【2】13.** 加值服務(Value-added service, VAS)係將下列何者利用新方式加以提昇改善，以創造更高的價值？
- ① 核心技術、產品或服務
 - ② 非核心技術、產品或服務
 - ③ 競爭性技術、產品或服務
 - ④ 領先的技術、產品或服務
- 【2】14.** 虛擬國際化權威檔(Virtual International Authority File, VIAF)係由美德法三國的國家圖書館與 OCLC 合作，利用下列何種技術，建立國際性的名稱權威檔服務？
- ① 視覺化資料
 - ② 鏈結資料
 - ③ 索引資料
 - ④ 關聯資料
- 【3】15.** 下列何項資訊檢索策略係將檢索問題分解為 3-4 個主題層面，分別界定檢索詞彙後，再以布林邏輯連結，進行檢索？
- ① 引用文獻滾雪球式檢索(Citation Pearl Growing)
 - ② 滾雪球式檢索(Snowball Search Strategy)
 - ③ 分區組合檢索(Block Building)
 - ④ 引用文獻檢索(Citation Search Strategy)
- 【4】16.** 「2014 年台灣共核發多少張綠建築標章？」此係屬於下列何種參考諮詢問題？
- ① 探索型問題(exploratory question)
 - ② 書目查證問題(bibliographic verification question)
 - ③ 主題研究問題(research question)
 - ④ 快速參考問題(ready reference question)
- 【2】17.** 受到 Google 搜尋引擎的衝擊，新一代圖書館目錄(Next generation catalog)係以下列何種服務型式呈現，方便使用者可以一次性地查詢圖書館自製、訂購和公開取用的館藏資源與資料庫？
- ① 開放探索倡議(Open Discovery Initiative, ODI)
 - ② 資源探索服務(Resource Discovery Service, RDS)
 - ③ 統一檢索引引(Unified search index)
 - ④ 開放鏈結資料(open linked data)
- 【1】18.** 當組織未訂購資料庫時，可以下列何種工具搜尋權威性可信賴的醫學資訊？
- ① MedlinePlus
 - ② Embase
 - ③ Medline
 - ④ Up To Date
- 【4】19.** 主動針對使用者感興趣的主題提供最新資訊，是圖書館常見的資訊加值服務，下列何項服務係提供使用者新到館的期刊或書籍的相關訊息（如：目次）？
- ① 專題選粹(selected dissemination of information, SDI)
 - ② 引用通報(citation alert)
 - ③ 出版通報(Publishing alert)
 - ④ 新知通報服務(Current Awareness Services, CAS)

【1】 20.下列何者係從資料的生命週期及其在學術、科學與教育各面向的效用，主動持續地管理，以使資料易於被發現與檢索，管理其品質，進行加值，並提供長時間的再利用？

- ①資料度用(Data curation)
- ②資料視覺化(Data visualization)
- ③資料展示(Data exhibition)
- ④大科學(big science)

【4】 21.科技圖書館可以下列何種指標做為過期期刊典藏之決策考量？

- ①影響係數(Impact Factor)
- ②即引指數(Immediacy Index)
- ③引用半衰期(Citing Half-Life)
- ④被引用半衰期(Cited Half-Life)

【4】 22.圖書館參考館員與讀者進行參考晤談的主要目的是：

- ①協助圖書館館員進行研究
- ②掌握讀者基本資料
- ③瞭解讀者人數與使用頻率
- ④確認讀者的問題與資訊需求

【2】 23.下列何者不是參考館員的工作項目？

- ①協助選擇讀者所需的資訊資源
- ②為參考工具書進行分類編目
- ③指導讀者使用館藏資源
- ④解決讀者的資訊問題

【2】 24.下列何種參考工具書不是用以回答簡要的事實性問題？

- ①百科全書
- ②索引
- ③年鑑
- ④名錄(directory)

【2】 25.聯合目錄的功能為：

- ①查詢最低圖書定價的出版社
- ②查詢多個圖書館館藏資料
- ③查詢即將出版的圖書
- ④查詢絕版的圖書

【2】 26. WorldCat 是：

- ①傳記資料
- ②聯合目錄
- ③百科全書
- ④索引

【2】 27.索引係屬於何種資料類型？

- ①第一手資料 ②第二手資料 ③第三手資料 ④第四手資料

【4】 28.根據 UNESCO 定義，索引典(thesaurus)為：

- ①一種年鑑
②一種索引
③一種百科全書
④一種控制詞彙工具，可應用於資料庫，協助讀者查詢資料

【1】 29.當本身圖書館資源無法解決讀者問題時，參考館員應如何處理？

- ①詢問讀者是否願意接受轉介服務
②請讀者放棄問題
③請讀者自行尋求專家協助
④請讀者參加圖書館利用教育課程

【1】 30.部分紙本參考工具書提供多種輔助索引的原因為：

- ①提供讀者多種查檢資料方式
②出版規定
③可提高定價
④圖書館要求

貳、複選題 22 題（每題 2.5 分，全部答對才給分，答錯不倒扣；未作答者，不予計分）

【1,2,4】 31.資訊檢索的三個主要模型為：

- ①向量空間模型(Vector Space Model) ②布林模型(Boolean Model)
③推理網絡模型(Inference Network Model) ④機率模型(Probabilistic Model)

【1,2】 32.下列哪些檢索功能可以增加求全率(recall)？

- ①切截(truncation) ②增加同義詞，並以“OR”連結
③增加同義詞，並以“AND”連結 ④增加限定“Limits”欄位條件

【1,2,3】 33.下列何者為人工索引典的詞項關係？

- ①廣義詞(Broader terms)
②關聯詞(Related terms)
③狹義詞(Narrow terms)
④頻率詞(Frequent terms)

【1,2,3,4】 34.下列哪些資料可以用來作為網頁索引及排序的根據？

- ①出現在網頁內的字詞
- ②利用“anchor text”來索引網頁
- ③使用者的檢索詞(query)與點擊連結(clickstream)的關連
- ④網頁得到的連結(inlinks)

【2,3,4】 35.下列哪些檢索技巧可以提供檢索結果的求準率(precision)？

- ①增加同義詞並以“OR”連結
- ②以控制字彙取代自然語言
- ③使用“NOT”除去不相關文獻
- ④使用專指性較高的詞彙

【1,2,3】 36.下列哪三種是相關反饋的基本方法？

- ①準相關回饋(Pseudo feedback)
- ②互動式相關反饋(Interactive feedback)
- ③隱式相關反饋(Implicit feedback)
- ④機器相關反饋(Machine feedback)

【1,3,4】 37.下列何者為都柏林核心集(Dublin Core)的基本元素？

- ①與其他作品關聯(Relation)
- ②摘要(Abstract)
- ③主題(Subject)
- ④簡述(Description)

【2,4】 38.下列哪兩種情境容易造成檢索結果過多？

- ①在自動索引下，因索引窮盡性低，且使用者檢索詞(query)短
- ②在自動索引下，因索引窮盡性高，且使用者檢索詞短
- ③在人工索引下，控制詞彙多，導致索引專指性高
- ④在人工索引下，控制詞彙少，導致索引專指性低

【2,3,4】 39.一般常將參考問題分成下列哪三大類？

- ①事實型問題(facts question)
- ②快速參考問題(ready reference question)
- ③主題研究問題(research question)
- ④書目查證問題(bibliographic verification question)

【1,3,4】40.巨量資料具備下列哪 3V 的特性？

- ① Volume 大量
- ② Value 價值
- ③ Velocity 持續快速產生
- ④ Variety 多樣龐雜

【1,2,4】41.數位出版時代來臨，面對傳播更為容易的數位內容，為保護智慧財產權，數位版權管理(Digital Right Management, DRM)的概念應運而生，DRM 權限的表達是以分權方式進行，分為下列哪三個獨立運作的機制，以完成各種可運作的商業機制？

- ①組織
- ②內容
- ③結構
- ④權限

【1,3】42.有關 Google Scholar™ (GS)的敘述，下列哪些正確？

- ① GS 是很好的跨領域研究工具
- ② GS 是全文資料庫，每筆搜尋結果都可取得全文
- ③ GS 要連結到機構所屬圖書館的付費資源/資料庫，始能取得全文
- ④透過 GS，所有的資訊都可以免費從網路取得

【1,2】43.如果研究人員希望其研究成果能發表在高影響力的期刊，而且希望能及早為其他學者引用；請問您會建議依下列哪兩項指標為其選擇投稿期刊？

- ①影響係數(Impact Factor)
- ②即引指數(Immediacy Index)
- ③引用半衰期(Citing Half-Life)
- ④被引用半衰期(Cited Half-Life)

【1,2,4】44.下列何者為第一手資訊(Primary Sources)？

- ①學術性期刊論文(Scholarly periodical articles)
- ②技術報告(Technical Reports)
- ③手冊(handbooks)
- ④專利說明書(patents)

【2,3】45.有關 ORCID(Open Researchers and Contributor ID)的敘述，下列何者正確？

- ①對於個人的服務是開放且收費
- ②是一開放型研究者與投稿者的數位識別碼
- ③可以有效區辨同名同姓的作者的科研成果辨識問題
- ④個人作者目前尚無法申請

【1,2,3】46.評估參考工具書的標準包括：

- ①編撰者權威性
- ②資料正確性
- ③內容編排方式
- ④定價及出版社每年銷售金額

【1,2】47.數位參考服務的特色包括：

- ①方便服務交通不便的讀者
- ②讀者可隨時於線上系統提出問題，不受圖書館開放時間限制
- ③付費才能使用的服務
- ④讀者可隨時與參考館員進行即時線上晤談

【1,2,3】48.館員進行參考晤談時，應注意哪些事項？

- ①保密讀者的個人資料與問題
- ②專注傾聽讀者問題
- ③鼓勵讀者陳述問題與需求
- ④由讀者決定採用開放式或封閉式問題進行參考晤談

【1,2】49.類書的功能包括：

- ①查找事物的起源與沿革
- ②查找詞藻典故
- ③查找圖書的典藏地點
- ④查找作家的筆名

【1,2,3】50.維基百科的特點不包括：

- ①由各領域專家學者撰寫
- ②內容由各領域專家學者審查
- ③第一手資訊
- ④由全球使用者共同撰寫

【1,2】51.我國政府統計資料的查找來源包括：

- ①各政府機關網站
- ②中華民國統計資訊網
- ③中華民國人名大辭典
- ④中國民國工商名錄

【1,2,4】52.電子版百科全書與紙本版百科全書比較，前者具備哪些特性？

- ①資料更新速度較快
- ②能以多媒體方式呈現內容
- ③撰寫者的權威性更高
- ④提供更多種檢索方式