

檔案脈絡描述標準之發展與可能應用

Records in Contexts: A New Way of Archival Description

黃玟臻 Huang, Mei-Jhen

國家發展委員會檔案管理局檔案徵集組科長

Section Chief, Acquisition Division, National Archives Administration, National Development Council

壹、前言

檔案脈絡是理解檔案之重要基礎，能提供有關檔案內容、結構、實體特性及外部環境等之輔助資料，以協助使用者瞭解檔案產生、保管及使用之過程，並從中建構檔案意義，及評估檔案價值與可信度（Theimer, K., 2012；Yeo, G., 2015；Sheridan, J., 2018）。

檔案脈絡透過檔案描述具體呈現，檔案描述係就檔案內容與形式加以分析，據以摘錄事項，藉由檔案描述可描繪檔案特徵，揭示主題內容與價值，指明出處及區別異同（檔案管理局，2002）。檔案脈絡描述則進一步強調檔案來源與保管資訊、檔案與相關人事物關係、檔案與其他檔案關係等（Yeo, G., 2015），相較於一般檔案描述更重視關係之表現。

為促進檔案描述後續之整合交換及查詢應用，檔案描述需建立在一致性基礎上，遵循相關標準規範，然而隨著時代演進，現行檔案描述標準已面臨挑戰，既有的層級描述方式難與其他資料連結（International Council on Archives, 2016），亦難以滿足組織整理及使用者檢索利用等需求，

因此結合資訊科技建立一套較現代化之描述標準有其必要。

新描述標準提出於 2016 年，稱為檔案脈絡描述（Records in Contexts, RiC），RiC 在實體、屬性及關係元素上，以開放架構描述及連結相關實體，呈現出檔案之脈絡關連，雖然 RiC 尚在草案修正階段，但其內容發展與影響層面值得關注。本文即以 RiC 為主軸，先探討 RiC 發展背景，再說明其概念模式組成部分，及對應之知識本體架構，最後提出 RiC 之可能應用。

貳、檔案脈絡描述發展背景

19 世紀中葉後，在新興傳播科技驅動下，檔案學專家開始思考如何設定及擷取基本描述單元，如產生者、典藏地等，以建立更有效的檔案檢索工具（Pitti, D., McCarthy, G., & Popovici, B. F., 2017），後續提出之標準多從這些描述單元予以規範，但各標準間並不一致。2012 年最佳實務模式及標準委員會（Committee on Best Practices and Standards）評估檔案描述標準修訂做法，在評估報告中指出，為解決描述標準不一致問題，建議

檢視檔案關係描述後直接發展新模式 (Popovici, B. F., 2016)。

根據評估報告結果，國際檔案理事會 (International Council on Archives) 之計畫委員會 (Programme Commission) 於 2012 年底成立檔案描述專家小組 (Experts Group on Archival Description)，交付該小組發展新檔案描述標準之任務。該小組所擬定之任務目標，包括確認描述基本元件、凝聚檔案描述共識、充分運用資訊技術、促進檔案及文史界合作、提升檔案使用與理解，及尊重傳統描述原則 (Pitti, D., Popovici, B. F., Stock, B.-F., & Clavaud, F., 2014)，期許新描述標準突破以往，能獲得檔案界普遍認同。

2016 年，檔案描述專家小組完成新描述標準 RiC 之草案，RiC 結合國際檔案理事會 4 項既有標準，包括：一、檔案描述國際標準 (General International Standard Archival Description, ISAD (G))；二、檔案權威紀錄國際標準：團體、個人與家族 (International Standard Archival Authority Record for Corporate Bodies, Persons, and Families, ISAAR (CPF))；三、檔案典藏機構描述國際標準 (International Standard for Describing Institutions with Archival Holdings, ISDIAH) 及四、描述功能國際標準 (International Standard for Describing Functions, ISDF) (International Council on Archives, 2016)。

上述 4 項標準中，以 ISAD (G) 最廣為使用及受國際採納，ISAD (G) 是檔案資料結構標準，提供一般性檔案描述方針，ISAD (G) 發展時間早於其他標準，1994 年發表第一版，1999 年發表第二版。ISAAR (CPF) 之使用亦相當普遍，主要用於檔案產生來源說明，規範檔案產生者之權威紀錄，目前有兩個版本，分別發表於 1996 及 2004 年。ISDIAH 和 ISDF 使用相對較少，兩者為 ISAAR (CPF) 之延伸，皆於 2008 年發表，發表時間較晚，其中 ISDIAH 提供有關檔案典藏機構之描述架構，ISDF 則提供檔案產生來源職

能描述 (Popovici, B. F., 2016)。以這些標準為基礎，RiC 運用資訊模型建構出新概念模式。

參、檔案脈絡描述概念模式

檔案脈絡描述概念模式 (Records in Contexts-Conceptual Model, RiC-CM) 主要提供檔案描述系統之語意與結構基礎，能符合檔案描述與使用需求，呈現社會複雜現象，並支援圖形與語意科技 (International Council on Archives, 2016)。RiC-CM 包含實體 (Entity)、屬性 (Attribute) 及關係 (Relation) 三部分，屬性能描述實體特性，關係可呈現實體之關連架構，三者分述如下。

一、實體

RiC-CM 定義之檔案實體，可概分為特定實體及共同實體。特定實體含代理人、檔案資料，及代理人與檔案互動所界定之實體，主要包括下列 8 項 (Pitti, D., Stocking, B., & Clavaud, F., 2016a；Pitti, D., Stocking, B., & Clavaud, F., 2016b；Popovici, B. F., 2016)：

- (一) 代理人 (Agent)：負責行動及引發影響之本體，除了個人、家族、團體外，亦包含機器人、軟體、探測器等，由個人或團體開發製造，以自動或半自動方式運作產生資料之實物。
- (二) 檔案 (Record)：代理人在生活或工作事件及活動中以任何形式或透過任何載體呈現之資訊。
- (三) 檔案元件 (Record component)：具有零散資訊內容之檔案組成元素，有助於檔案實體或其理解之完整性。
- (四) 檔案集合 (Record set)：代理人聚集一個或多個具共通屬性或關係之檔案，如具共通來源、主題、類型、地點等；一個檔案可同時分屬於多個檔案集合。

(五) 功能 (Function)：代理人之目的、用途或目標，可能隨時間改變。

(六) 行動 (Activity)：代理人為達成其功能所採取之行為或處置。

(七) 規定 (Rule)：對達成功能或採取行動之方式予以規範，如行動之期程等。

(八) 授權 (Mandate)：代理人將職務、行動或工作之責任委託給一個或多個代理人。

共同實體則包括：地點 (Place)、日期 (Date)、工作 (Occupation)、功能-摘要 (Function-Abstract)、文件形式 (Documentary form)、主題-摘要 (Subject-Abstract)、外型/特徵/本質 (Figure/Character/Being)、著名事件 (Named event)、著名時期 (Named period)、活動 (Movement)，共同實體可運用控制詞彙作為檢索點，並與其他資源連結。

二、屬性

RiC-CM 之屬性大致可分為兩類，一類為適用於所有實體之描述，不論是特定或共同實體，如識別號、名稱、一般註等；另一類為根據不同實體性質之個別描述，如語言、取用條件、範圍與內容等。以檔案集合為例，其屬性包括 (Pitti, D. et al., 2016a; Popovici, B. F., 2016)：

(一) 檔案集合整體屬性：檔案集合類型、成長情況、日期、歷史、排列、可靠性、完整性等。

(二) 所屬檔案或檔案集合屬性：範圍及內容註、檔案長度等。

(三) 後續檔案之共通屬性：檔案狀態、分類、文件形式、內容類型、編碼形式、媒體類型、語言、實體特徵、產出技術、取用情況、權利等。

三、關係

關係用以呈現實體之關連，如檔案主題關

係、同實體中不同事物之時間先後與層級關係、工作內容與工作時間地點關係、家族成員關係等，並可區分過去與當前關係或固定特質。關係之屬性包括：相關代理人註、相關地點註、相關檔案註、相關檔案元件註、相關檔案集合註、相關功能或行動註、相關規定註、相關授權註等 (Popovici, B. F., 2016)。

實體、屬性及關係之範例如圖 1 所示，檔案 1、檔案集合 1-4、代理人 1-2、功能 1 及行動 1 皆為實體，以檔案與代理人兩實體來看，檔案實體之屬性包括識別號、名稱、語言、取用條件、範圍與內容等，代理人實體之屬性包括識別號、類別、名稱、語言、歷史等。所有實體之關係說明了檔案 1 同時是檔案集合 1、2、3 之一部份，檔案集合 1、2、3 分別為案卷、系列、全宗，檔案集合 1 過去是檔案集合 4 之一部份；此外，檔案 1 分別由代理人 1 建立及代理人 2 徵集，檔案 1 來自執行功能 1 之行動，可作為功能 1 之稽憑資料。

綜上所述，實體、屬性、關係為 RiC-CM 基本組合元素，運用三者描述實體及關連網絡，能彈性呈現檔案狀況。此外，RiC-CM 亦試圖降低實體描述之侷限，如以代理人表示檔案產生、徵集、典藏等之行動者，不採用傳統之產生者 (Creator)，又如建立檔案集合實體，聚集有共通屬性或關係之檔案等，與傳統作法相當不同。

肆、檔案脈絡描述知識本體

檔案脈絡描述知識本體 (Records in Contexts-Ontology, RiC-O) 是將 RiC-CM 轉換為機器可處理之詮釋資料，使用資源描述架構 (Resource Description Framework, RDF)、資源描述綱要 (Resource Description Framework Schema, RDFS)、網路本體語言 (Web Ontology Language, OWL) 等標準，將分散於各系統之檔案與相關資料連結，

Record 1	is part of	Record Set 1 (title = file1)
Record 1	is part of	Record Set 2 (title = series1)
Record 1	is part of	Record Set 3 (title = fond1)
Record Set 1	was part of	Record Set 4 (title = series2)
Record 1	is evidence of	Function 1
Record 1	results from	Activity 1
Activity 1	is performed to fulfil	Function 1
Record 1	is created by	Agent 1
Record 1	is collected by	Agent 2

→ ID
Name
Language Information
Conditions of Access
Scope and Content
...

→ ID
Type
Name
Language Information
History
...

圖 1 RiC 實體、屬性及關係範例

資料來源：Popovici, B. F. (2016)；Pitti, D. et al. (2016a)

置於語意網 (Pitti, D. et al., 2017)，以增進資源整合及取用。

語意網是 1998 年 Tim Berners Lee 提出之概念，是將全球資訊網的巨量網頁賦予機器可存取之詮釋資料，再藉由標準、標記語言及軟體程式，讓機器可以處理網頁內容及網頁間錯綜複雜之關係，使網際網路成為一個整合的資訊交換媒體，代理程式也能更具智慧，執行使用者意志或滿足使用者需求 (陳光華，2012)，希望透過相關規則及網頁內容結構化，促進機器對網頁之理解。

在語意網中，資料有其獨一無二的統一資源識別字 (Uniform Resource Identifier, URI)，可用以標示及連結相關資料，形成資料集合，並以圖形化呈現，提供較完整資訊。如欲瞭解某位歷史學家之背景，可從各機構蒐集他的出生紀錄、居住成長資料、不同名字與筆名、參與之活動、各類著作及政府機關檔案等。當愈來愈多機構將

保存於封閉資料庫之資料開放使用，並以相同知識本體建立詮釋資料，將有助於提升查詢結果，增進資料可及性與重複利用機會 (Pitti, D. et al., 2016a)。

語意網的實現需運用知識本體對資料內容進行描述，知識本體是透過一個詞表達一個已存在之概念，並延伸至相關詞，及建立詞間關係，透過這些詞彙與關係，建構出特定領域的知識 (張慧銖，2011)。RiC-O 建構知識本體之原則如下 (Pitti, D. et al., 2016a)：

- 一、完整性：知識本體需將實體相關概念均納入，包含較高層次抽象概念及實務功能需求。
- 二、明確性：明確界定概念及關係，加強呈現檔案特性、來源、歷史、重要事件等，並朝多語言發展。
- 三、彈性：概念方面，可選擇呈現部分或完整之知識本體，關係方面，可選擇呈現一組關

係，或對關係進一步描述，並能維持知識本體之延展性，增加新類別和屬性。

四、對應：需能對應其他知識本體之概念與關係，如對應文化資產領域知識本體 CIDOC-CRM 與 FRBRoo、一般性知識本體 FOAF 與 Dublin Core Metadata Initiative 等。

為了將資料加入語意網，RiC-O 採用 RDF、RDFS 等全球資訊網協會（World Wide Web Consortium）語意網資料格式標準，RDF 是描述網路資源內容及結構之語法，RDFS 是 RDF 之詞彙描述語言，用以建立概念架構及定義屬性（Llanes-Padrón & Pastor-Sánchez, 2017）。RiC-O 並依 RDF 三元素詞組（Triple）標準形式建立機器可處理之詮釋資料，三元素詞組包括主詞（Subject）、述語（Predicate）及受詞（Object），主詞用以表示動作者、資源等事物，述語用以定義關係或描述資源屬性，受詞用以表示接受動作之對象或屬性的值，加上屬性與關係之描述，形成「資源－關係－資源」及「資源－屬性－值」語意模型（Pitti,

D. et al., 2016a; Llanes-Padrón, D., & Pastor-Sánchez, J., 2017），此模型能讓資料內容詮釋較具規則性，有助於機器自動處理及邏輯推演。

運用 RDF 組織後之 RiC 圖形如圖 2 所示，其內容表示 Lucien Febvre 建立編號 591AP 之全宗，其中包含信件及課程兩個系列檔案集合，信件檔案集合之編號為 591AP 52-55，其下尚有個別信件檔案，課程檔案集合之編號為 591AP 2-3，課程檔案集合來自 Lucien Febvre 講授之歷史課程。

圖 3 為 RiC 語法範例，其描述內容與圖 2 呈現內容相同，資源屬性與關係為「資源－關係－資源」者，表示列於主詞之資源，對列於受詞之資源有所動作，形成連結主詞與受詞之關係，如「anf:LucienFebvre/ RiC-O:created/ anf:fonds591AP」，表示 LucienFebvre 建立了編號 591AP 之全宗。資源屬性與關係屬「資源－屬性－值」者，用以說明資源本身屬性內涵，如「anf:LucienFebvre/ rdf:type/ RiC-O:Person」，說明 LucienFebvre 類型為個人。此外，主詞、述語及受詞之 URI 字串，包含 anf、rdf、owl

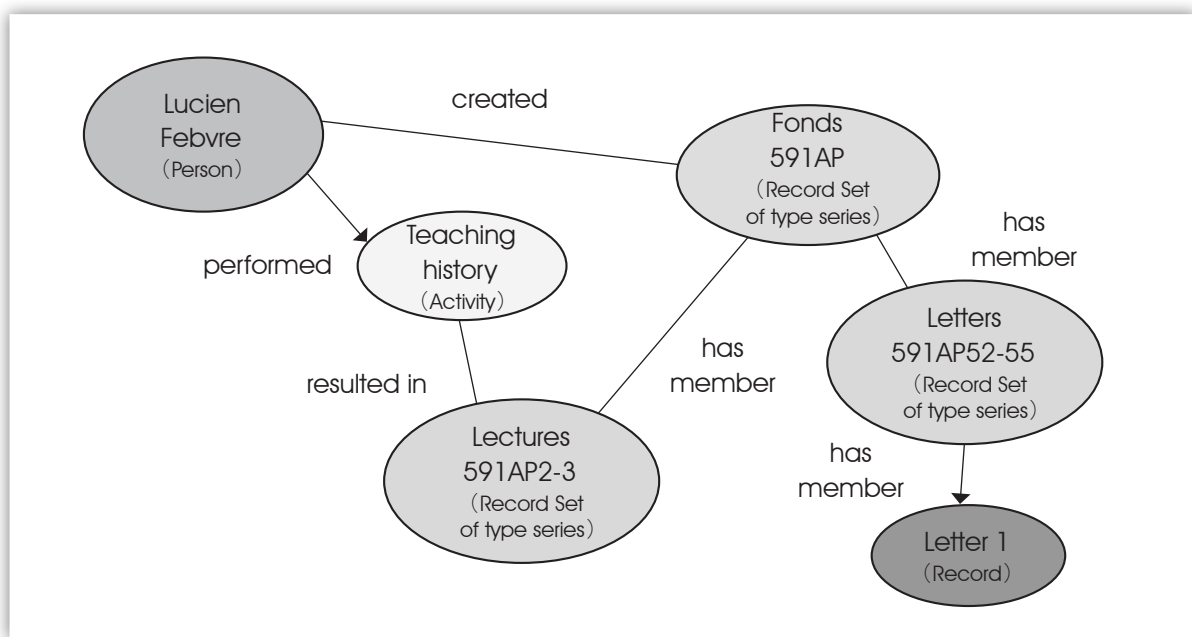


圖 2 RiC 圖形範例

資料來源：Pitti, D. et al. (2016a)

等命名空間 (Namespace) 及識別字 (Identifier)，如「anf:LucienFebvre/ owl:sameAs/ <http://data.bnf.fr/ark:/12148/cb119024609#foaf:Person>」，即指引有關 LucienFebvre 之生平介紹，可連結至 <http://data.bnf.fr/ark:/12148/cb119024609#foaf:Person>。

綜上所述，RiC-O 之詮釋資料以 RiC-CM 為基礎，所建立之知識本體能展現檔案脈絡中的實體與關係，並與其他知識本體對應。再者，RiC-O 將各系統之檔案描述資料開放在網路上，連結相關資源，能讓檔案脈絡呈現更形完整，提升檔案處理與使用效益。

伍、檔案脈絡描述之可能應用

RiC 之概念及架構與以往檔案描述標準相當不同，其差異也為後續應用帶來更多可能。在原則概念方面之主要差異，在於以往的檔案描述標準大多奠基在來源原則 (Principle of Provenance) 之傳統理解，通常聚集同一機構或個人之檔案，

對單一全宗進行獨立描述，但 RiC 將來源予以擴大詮釋，認為檔案並非單獨存在，而是建立於過去、當前與未來之脈絡中，與檔案產生者、管有者及使用者等相互連結，相較以往，RiC 更為開放也較能反映檔案來源複雜性。在編排描述架構方面，以往的檔案描述標準多為層級式，從全宗到其組成部分進行各層級描述，但 RiC 在資訊科技輔助下，採取較開放多面向之描述方式，除了能含括原先各層級描述內容，亦可對檔案及其相關資料、人員、活動、法令等所形成之網絡提供較完整描述 (Darnall, K. L., 2018)，相較於層級式描述，更能彈性呈現檔案實際關連狀況。

雖然 RiC 可能有資訊系統複雜度較高、目錄品質較難控制、關係定義尚不完整等問題，但系統操作及著錄問題能透過培訓逐漸克服，關係定義也將隨 RiC 發展愈來愈完備；加上 RiC 展現之特性，包括重視檔案產生與存在之脈絡、強調檔案關係建構等，不論在資料檢索範圍、目錄呈現方式、檔案比對推薦，或使用者理解分析上都

```
anf:LucienFebvre      RiC-O:created      anf:fonds591AP ;
                      rdf:type          RiC-O:Person ;
                      RiC-O:performed   anf:teachingHistory ;
                      owl:sameAs
<http://data.bnf.fr/ark:/12148/cb119024609#foaf:Person> .

anf:fonds591AP        rdf:type          RiC-O:RecordSet ;
                      RiC-O:hasMember   anf:letters591AP52To55,
                                          anf:Lectures591AP2to3 .

anf:letters591AP52To55  rdf:type          RiC-O:RecordSet ;
                      RiC-O:hasMember   anf:letter1 .

anf:letter1           rdf:type          RiC-O:Record.

anf:teachingHistory    rdf:type          RiC-O:Activity ;
                      RiC-O:resultedIn  anf:Lectures591AP2to3 .

anf:Lectures591AP2to3  rdf:type          RiC-O:RecordSet.
```

圖 3 RiC 語法範例

資料來源：Pitti, D., Stocking, B., & Clavaud, F. (2016b)

有相當發揮空間，其後續應用值得期待。整體而言，RiC 可能應用面向如下：

一、連結相關檔案，促進整合檢索

RiC 運用網路技術標準，能將各系統檔案資料開放與連結。目前各單位大多為獨立系統，詮釋資料封閉，較難與其他單位之資料整合運用；RiC 將檔案資料開放至網路上，能與文史機構檔案及數位典藏等相關資料建立連結，對使用者而言，可一次檢索到較完整脈絡之資料，原先未被注意之檔案也能藉此獲得關注與使用。

二、結合圖形化技術，協助檔案辨識

RiC 建構在實體、屬性及關係模式上，其檔案描述架構並非傳統層級形式，而是趨向開放網絡結構。相關研究顯示，層級形式架構與使用者資訊行為未必相符（Darnall, K. L., 2015），對於較複雜之檔案關係亦難清楚呈現；相較之下，RiC 結合圖形化技術，加強脈絡關係之視覺辨識，運用於檔案查詢系統瀏覽及檢索結果呈現，能協助使用者快速辨識與選擇所需檔案。

三、擴增資料比對，推薦相關檔案

當使用者對欲探索之議題缺乏基本概念，查詢結果不盡理想時，檔案查詢系統應適時推薦相關檔案輔助使用者，傳統上可比對主題及全宗等產生推薦；RiC 之應用則擴展了比對來源，除了主題與全宗外，亦可從實體間有無關係連結推測相關性，兩實體之關係連結愈多，或可表示兩實體緊密度愈高，再配合使用行為紀錄分析，應能提高推薦檔案之相關度。

四、以脈絡延伸，增進檔案理解詮釋

檔案並非獨立存在，而是存在歷史脈絡中，與其他檔案及其建立者、使用者等息息相關。在特定議題或事件上，過去以主題為基礎之分析，內容易片段，使用者需花較多時間理解（Simões, M. G., Freitas, M. C. V., & Rodríguez-Bravo, B., 2016）；運用 RiC 能讓使用者循檔案脈絡延伸，有條理地理解相關檔案，甚至從來源廣泛的檔案中觀察各方觀點，提升檔案內容詮釋。

陸、結語

RiC 為新一代檔案描述標準，其以實體、屬性及關係進行檔案描述，並透過知識本體轉換詮釋資料與連結語意網。相較於傳統檔案描述方法，RiC 由層級架構轉向多面向架構，在檔案來源與脈絡之描述上更具彈性，又其架構能協助使用者從廣泛的脈絡資訊中進行觀察，獲得更深入之理解詮釋；此外，RiC 連結相關資源，既能促進資源整合檢索利用，又能結合圖形化技術協助使用者快速辨識，進一步還可將連結緊密度作為檔案推薦之比對來源。

目前 RiC 仍持續徵詢各方意見進行調整，其中概念模式部分，雖已有雛形但其關係特性與使用細節還待訂定，至於知識本體及應用系統部分，由於設計複雜度高，至今尚未建置完成。期許未來有更多學界及實務界專業人士投入，共同研究與開發，及早實現 RiC 理想為檔案描述帶來新面貌。

參考文獻

- 張慧銖 (2011)。《圖書館電子資源組織：從書架到網路》。新北市：Airiti Press。
- 陳光華 (2012)。語意網。《圖書館學與資訊科學大辭典》。檢自 <http://terms.naer.edu.tw/detail/1679015/>
- 檔案管理局 (2002)。《檔案管理名詞彙編》。新北市：檔案管理局。
- Darnall, K. L. (2015). *Un/Mediated: Access to human rights records in context* (Master's thesis, The University of Texas at Austin). Retrieved from <https://repositories.lib.utexas.edu/handle/2152/31948>
- Douglas, J. (2018). A call to rethink archival creation: Exploring types of creation in personal archives. *Archival Science*, 18(1), 29-49.
- International Council on Archives. (2016). *Records in Contexts: A conceptual model for archival description*. Retrieved from <https://www.ica.org/sites/default/files/RiC-CM-0.1.pdf>
- Llanes-Padrón, D., & Pastor-Sánchez, J. (2017). Records in contexts: The road of archives to semantic interoperability. *Program*, 51(4), 387-405
- Pitti, D., Popovici, B. F., Stock, B.-F., & Clavaud, F. (2014). *Expert group on archival description: Interim report*. Retrieved from <http://www.girona.cat/web/ica2014/ponents/textos/id56.pdf>
- Popovici, B. F. (2016). *Records in Contexts: Towards a new level in archival description?* Paper presented at the International Conference Technical and Field Related Problems of Traditional and Electronic Archiving, Radenci. Retrieved from http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci_2016/013-031_popovici_2016.pdf
- Pitti, D., Stocking, B., & Clavaud, F. (2016a). *Records in Contexts (RiC)*. Retrieved from <https://www.ica.org/sites/default/files/session-7.8-ica-egad-ric-congress2016.pdf>
- Pitti, D., Stocking, B., & Clavaud, F. (2016b). *An introduction to "Records in Contexts": An archival description draft standard*. Retrieved from <https://online.liverpooluniversitypress.co.uk/doi/pdf/10.3828/comma.2016.18>
- Pitti, D., McCarthy, G., & Popovici, B. F. (2017). *Records in Contexts (RiC)*. Retrieved from <http://www.alaarchives.org/wp-content/uploads/2018/01/1.-Daniel-V.-Pitti.pdf>
- Sheridan, J. (2018). *Digital archiving: Context is everything*. Retrieved from <https://blog.nationalarchives.gov.uk/blog/digital-archiving-context-everything/>
- Simões, M. G., Freitas, M. C. V., & Rodríguez-Bravo, B. (2016). Theory of classification and classification in libraries and archives: Convergences and divergences. *Knowledge Organization*, 43(7), 530-538.
- Theimer, K. (2012). Archives in context and as context. *Journal of Digital Humanities*, 1(2). Retrieved from <http://journalofdigitalhumanities.org/1-2/archives-in-context-and-as-context-by-kate-theimer/>
- Yeo, G. (2015). Contexts, original orders, and item-level orientation: Responding creatively to users' needs and technological change. *Journal of Archival Organization*, 12(3-4), 170-185.