

學術論文

話中有話：以「自動化內容分析」解讀北韓的核子語言*

Reading between the Lines: Automated Content Analysis of North Korean Nuclear Rhetoric

芮宗泰 *Timothy Rich*

印地安納大學政治系博士候選人

*Doctoral Candidate of Department of Political Science
Indiana University*

劉泰廷 *Tai-Ting Liu*

中興大學國際政治研究所博士生

*Doctoral Student of Graduate Institute of International Politics
National Chung Hsing University*

摘要 / Abstract

雖然韓戰早已成為遙遠的歷史回憶，但兩韓關係自冷戰時期僅有些微改善，2010年朝鮮半島更因為天安艦事件、延坪島砲擊、北韓政權交接等，情勢再度緊張。六方會談在近期內似乎難以再啟動，北韓核武發展乃是抵制兩韓和美朝關係改善的主要因素。北韓政權的封閉性使外界無法透

* This research is supported by a Korean Studies Grant from the Academy of Korean Studies (AKS-2011-R-11). 本文初稿同時發表於2011年11月11-13日，政治大學舉辦之台灣政治學會年會暨「辛亥百年與兩岸政治發展」學術研討會，以及美國聖路易舉行之國際關係學會中西部分會年會(ISA Midwest Conference 2011)。兩位作者感謝主持人、評論人以及與會聽眾之寶貴意見，使完稿得以經修改後呈現。

視其內部發展，進一步解讀領導人的政策內涵和意圖。本文試透過自動化內容分析，透過以電腦軟體耙梳資訊的方式，追蹤和解讀平壤的政策意涵。1946年成立的朝鮮中央通信社為北韓政權的官方宣傳部門，其新聞報導在某種程度上代表官方立場，亦是外界觀察北韓發展的主要管道之一。以朝鮮中央通信社的報導為基礎，本文藉由分析對於核武研發進程、相關決策人士及2010年朝鮮半島情勢報導的比例和報導之間的連結，試說明北韓政府希望外界如何看待其發展。本文試指出平壤對於六方會談與會國在描述上的差異性，以及此差異對於爭端解決的暗示。

Although the Korean War may have ended in Western minds in 1953, North Korean views of inter-Korean relations have remained largely tied to its Cold War origins. Despite this challenge, little is known about how North Korean leadership views the outside world. While 2010 may be remembered more for two high profile military skirmishes (the sinking of the South Korean warship *Cheonan* and the shelling of Yeonpyeong Island) as well as the rise of the third generation of the Kim dynasty (Kim Jong Un), the lack of progress on nuclear talks should not be overshadowed. Pyongyang's nuclear program remains the largest obstruction to improving inter-Korean and US-DPRK relations. This paper uses 2010 as a case study of how North Korea attempts to frame its nuclear program vis-à-vis its neighbors and the US. Using daily news reports from the Korean Central News Agency (KCNA), this paper suggests that North Korea largely identifies nuclear issues with the US in contrast to the other parties in the Six Party talks. Furthermore, the nuclear rhetoric negative correlates with references to contemporary military conflicts or Kim Jong Un.

關鍵字：北韓核武危機、六方會談、朝鮮中央通信社、自動化內容分析

Keywords : North Korean Nuclear Crisis, Six-Party Talks, Korean Central News Agency (KCNA), Automated Content Analysis

壹、前言

四十年前，全世界僅有五個核武國家。此後，許多國家陸續投入內部研發或透過不同的交易管道試圖取得核子武器，而對美國而言，除了伊朗之外，北韓嚴重挑戰亞太區域安全及美國在東亞地區的利益。雖然韓戰早已成為遙遠的歷史回憶，但兩韓關係自冷戰時期僅有些微改善，2010年朝鮮半島更因為天安艦事件、延坪島砲擊、北韓政權交接等，情勢再度緊張。六方會談在近期內似乎難以再啟動，北韓核武發展乃是抵制兩韓和美朝關係改善的主要因素。本文以2010年為個案，觀察北韓如何藉由引述鄰國和美國來描述其核武計畫。固然2010年可能不具代表性，但是合理推論，該年度包括許多影響北韓形容核武計畫的時下議題。本文希望透過聚焦於該時間點，針對一個不常以理論或實證性架構檢視其行為的國家，建立理論。同時，此研究抱著探索的精神，企圖增進吾人對北韓如何看待外界的理解，並奠下日後涵蓋時間範圍更長的相關研究。

若將「北韓」輸入谷歌學術搜尋引擎（Google Scholar），我們可發現自2000年開始，北韓相關的社會人文類文獻有19,300筆，而自2010年開始，也有5,200筆。¹令作者感到意外的是，相較領土大小與北韓相當的十個國家，²就文獻的數量而言，並未呈現如北韓一般的巨幅增長。自2000年起關於北韓的文獻，數量上低於其他十個國家的中間值（median），而自2010年開始，其更比其他國家文獻的點擊數少1000筆左右。雖然近年世人對北韓的理解比過去都還多，但資料的質量是否也隨之提升，仍值得進一步探討。尤其是言及北韓領導人的認知和政權的內部運作，目前仍存在相當大的探索空間。逐年擴大的文獻量雖然供世人透視北韓的機會，³但大

¹ 搜尋日期：2011年6月21日。

² 此十個國家分別是：尼泊爾、孟加拉、塔吉克、希臘、尼加拉瓜、馬拉威、厄利垂亞、貝南、宏都拉斯、利比瑞亞。

³ 參見：Marcus Noland, *Avoiding the Apocalypse: The Future of the Two Koreas* (Washington DC: Institute for International Economics, 2000); Han Park, *North Korea: The Politics of*

部分的文獻皆未能表示北韓政府如何理解外界發展。另一方面，由於缺乏可供檢證的一手資料，許多關於平壤「核武野心」(nuclear ambition)的探討也不無缺陷。⁴Hymans 將世人對此議題的理解喻為「一團無知」(a cloud of ignorance)。⁵無論如何，根據俄羅斯和匈牙利的解密檔案，北韓自 1960 年代初期即開始透露對發展核武的興趣。⁶當其他議題在 2010 年取代北韓成為頭條新聞的同時，平壤的核武發展計畫依舊悄悄進行，北韓政權更在 11 月 30 日承認推動濃縮鈾計畫 (uranium enrichment program)。

相較於透過「揣測式」地試圖組織零碎的資訊，本文試透過「自動化內容分析」(automated content analysis)，以電腦軟體耙梳資訊的方式，追蹤和解读平壤的政策意涵。作者以朝鮮中央通信社 (Korean Central News Agency, KCNA) 的每日新聞報導作為分析平壤如何理解核武議題和其國際身分的實證基礎。1946 年成立的朝鮮中央通信社為共產黨的官方舌喉，其自 1997 年透過網頁提供英文報導。作者試透過分析提及「核武計畫」、「核武會談的主要與會者」以及 2010 年重要國際事件的頻率和三者之間的連結，呈現平壤希望外界如何看待其發展核武的野心。此研究試展現與迄今相關韓文研究之間的區隔，以及與刻意扭曲北韓原文的英譯研究

Unconventional Wisdom (Boulder: Lynne Reinner, 2002); Victor Cha and David Kang, *Nuclear North Korea: A Debate on Engagement Strategies* (New York: Columbia University Press, 2003); Stephan Haggard and Marcus Noland, "Economic Crime and Punishment in North Korea," *EAI Asia Security Initiative Working Paper*, No.5 (2010); Ralph Hassing and Kongdan Oh, *The Hidden People of North Korea: Everyday Life in the Hermit Kingdom* (New York: Rowman and Littlefield, 2009).

⁴ 參見：Ralph Hassing and Kongdan Oh, *North Korea through the Looking Glass* (Washington DC: Brookings Institution Press, 2000); Andrew Scobell, "North Korea's Nuclear Intentions," in James Lister ed., *Challenges Posed by the DPRK for the Alliance and the Region* (Washington DC: Korea Economic Institute), pp.78-95; Cha and Kang, *op. cit.*

⁵ Jacques Hymans, "Assessing North Korean Nuclear Intentions and Capacities: A New Approach," *Journal of East Asian Studies*, Vol.8, No.2 (2009), p.259.

⁶ Balasz Szalontai and Sergey Radchenko, "North Korea's Efforts to Acquire Nuclear Technology and Nuclear Weapons: Evidence from Russian and Hungarian Archives," *Cold War International History Project Working Paper*, No.53 (2006).

之間的落差。⁷作者希望此舉能釐清迄今一些關於北韓議題的模糊連結和假設，並期盼能為未來的核武談判提供啟示。

本文首先說明自動化內容分析如何可以運用在北韓發展核武的議題，其次簡單地回顧北韓核武計畫的發展歷程。本文將接續以多層次統計（multilevel statistics）呈現2010年平壤在公開聲明中的政治意涵（political rhetoric）。此分析最終將歸納北韓對六方會談各國的認知差異，以及此差異對重啟談判和解決北韓核武問題的影響。

貳、從字裡行間取得資料

就含義而言，「溝通」包括「意義」（meaning），而解讀意義（尤其在文本內）並非易事。社會科學學者常以不同內容分析的手法試圖直接呈現文本的意義。⁸然而，「人為的符號分析」（human coded analysis）有兩大限制。首先，以人力記錄符號是頗耗費時間和精力的工作，通常要求研究者必須共同執行一項計畫或限縮計畫範圍，例如專注在幾個月以內的新聞報導或報紙的某一部分，而非一段時間內所有坊間報章雜誌的相關報導。其次，即使個人可以符合時間要求，維持資料解讀的一致性（consistency）也是一大難題，因為（1）符號記錄者隨時間進程而產生的非一致性（例如，歸納條件的更動）或（2）兩個以上記錄者之間的非一致性。⁹以人為設定的字句庫或更複雜的演算法，電腦輔助的記錄方法（computer assisted coding）可克服前述限制，並且實際上將可被分析的文本擴增到無限大。嘗

⁷ D.B. Poneman et al., *Going Critical: The First North Korean Nuclear Crisis* (Washington DC: Brookings Institution Press, 2004).

⁸ Bernard Berelson, *Content Analysis in Communication Research* (Glencoe, IL: Free Press, 1952); Ole Holsti, *Content Analysis for the Social Sciences and Humanities* (Reading: Addison-Wesley, 1969).

⁹ Robert Weber, *Basic Content Analysis* (Newbury Park: Sage, 1990); Kimberly Neuendorf, *The Content Analysis Guidebook* (Thousand Oaks: Sage, 2002); Klaus Krippendorf, *Content Analysis: An Introduction to its Methodology* (Thousand Oaks: Sage, 2004).

試以電腦記錄符號並非創新之舉，¹⁰但是數十年來，演算法和記憶體的相關技術使終停留在基礎階段，未能有所突破。直到十五年前，相關技術的一致性和準確性的提升，方允許電算擁有更廣泛的用途。¹¹

雖然克服人為的限制，政治學領域內仍罕見自動化內容分析，該方法主要運用在解讀西方民主政體的政策傾向以及部分條約的發展。¹²在谷歌學術搜尋引擎搜索「自動化內容分析」一詞，可發現在 1990-2010 年之間，政治學期刊僅收錄 19 筆相關文章，相較同一時期，鍵入「內容分析」所回覆的 2000 筆。¹³同樣地，針對北韓的傳統內容分析為數頗多，自動化內容分析則相對稀少。Malici and Malici 曾以自動化內容分析比較金日成（Kim Il Sung）和卡斯楚（Fidel Castro）兩人的演說。¹⁴Choi 關注 2001 年 1 月至 2002 年 12 月之間紐約時報（New York Times）和華盛頓郵報（Washington

¹⁰ 參見：Philip Stone et al., *The General Inquirer: A Computer Approach to Content Analysis* (Cambridge, MA: MIT Press, 1966).

¹¹ Tom Mitchell, *Machine Learning* (New York: McGraw Hill, 1997); Christopher Manning and Hinrich Schütze, *Foundations of Statistical Natural Language Processing* (Cambridge, MA: MIT Press, 1999); Bo Pang et al., "Thumbs Up? Sentiment Classification Using Machine Learning Techniques," in *Proceedings of EMNLP* (2002), pp.79-86; Matt Thomas, "Get Out the Vote: Determining Support or Opposition from Congressional Floor-Debate Transcripts," in *Proceedings of EMNLP* (2006), pp.327-335; Mark Schaefer and Stephen Walker, *Beliefs and Leadership in World Politics: Methods and Application of Operational Code Analysis* (New York: Palgrave, 2006).

¹² 前者請參見：Michael Laver et al., "Extracting Policy Positions from Political Texts Using Words as Data," *American Political Science Review*, Vol.97, No.2 (2003), pp.311-331; Kenneth Benoit and Michael Laver, "Estimating Irish Party Positions Using Computer Wordscoring: The 2002 Elections," *Irish Political Studies*, Vol.18, No.1 (2003): 97-107; Ian Budge et al., *Mapping Policy Preferences: Estimates for Parties, Electors and Governments 1945-1998* (Oxford: Oxford University Press, 2001)。後者見：Arthur Spirling, "Bargaining Power in Practice: U.S. Treaty-making with American Indians, 1784-1911," working paper (2010).

¹³ 作者在名稱包括「政治」(politics)和「政治學」(political science)的所有學術期刊搜尋「自動化內容分析」(automated content analysis)一詞。搜尋日期：2011 年 2 月 22 日。

¹⁴ Akan Malici and Johnna Malici, "The Operational Codes of Fidel Castro and Kim Il Sung: The Last Cold Warriors?" *Political Psychology*, Vol.26, No.3 (2005), pp.387-412.

Post) 關於北韓的報導。¹⁵Seo 同樣分析前述兩份報紙，並納入南韓的朝鮮日報 (Chosun Ilbo) 和中央日報 (JoongAng Daily)，關注 2003-2007 年間每次六方會談召開前後的一星期。¹⁶Hymans 針對 1975-2008 年間北韓領導人的元旦演說進行分析，以金日成和金正日 (Kim Jong Il) 父子的英譯宣言為文本。¹⁷Kwon 僅分析「勞動新聞」(Rodong Sinmun) 的社論；「勞動新聞」是 1980-2004 年間朝鮮勞動黨 (共產黨) 的官方報刊。¹⁸Chung et al. 以傳統手法分析美國、中國、南韓、日本、俄羅斯等五國的原版報紙，評估 2006 年北韓核試爆的報導。¹⁹ Zuell 雖然以「自動化內容分析」分析西方報紙，但其僅關注 2006 年 10 月的某週，北韓也僅是該研究的不同重點之一。²⁰Haynes 可能最接近本文的研究取向，以「自動化內容分析」解讀 1997 年 12 月至 1999 年 2 月以及 2000 年 1 月至 2001 年 1 月 KCNA 英文報導。²¹然而，其分析大多加總至單月頻率 (monthly frequencies)，並未應用更複雜的自動化分析手法。雖然上述研究皆不失為「朝正確方向邁進」的佳作，但是研究者的自我設限使其未能充分利用「自動化內容分析」的功能，進而應用在北韓研究。

¹⁵ Jinbong Choi, "The Representation of North Korean National Image in National Newspapers in the United States," *Public Relations Review*, Vol.36, No.4 (2010), pp.392-394.

¹⁶ Hyunjin Seo, "International Media Coverage of North Korea: Study of Journalists and News Reports on the Six-Party Nuclear Talks," *Asian Journal of Communication*, Vol.19, No.1 (2009), pp.1-17.

¹⁷ Hymans, *op. cit.*

¹⁸ Soyoung Kwon, "Survival of the North Korean Regime and Changing Legitimation Modes," *Mario Einaudi Center for International Studies Working Paper*, No.07-05.

¹⁹ Mun-Yung Chung et al., "Framing the Nuke: How News Media in Six-Party-Talk Countries Framed North Korea's Nuclear Test," paper presented at the annual meeting of the International Communication Association, Montreal, Canada, 2008.

²⁰ Cornelia Zuell, "Using Computer-Assisted Text Analysis to Identify Media Reported Events," paper presented at the 10th International Conference on Statistical Analysis of Textual Data, Sapienza University, Rome, Italy, June 9-11, 2010.

²¹ Joseph Haynes, *Red Journalism as a Keyhole: Evaluating Discrepancies in New Systems and Inferring the Political Direction of North Korea Based on a Quantitative Analysis of the Korean Central News Agency Website*, MA Thesis (2001), University of North Carolina.

雖然北韓 KCNA 報導常被西方媒體視為政治宣傳，但是 KCNA 的英文報導仍有可供觀察之處。作為摒除外界詮釋並以國外觀眾為主的官方新聞來源，KCNA 的報導相當程度地捕捉北韓政府所希望呈現的形象。誠如 Kim 指出，北韓社論提供理解該國的最佳途徑，針對 KCNA 的更廣泛分析可望揭發直接和間接的訊息，並同時避免專注於社論可能產生的選擇性偏差 (selection bias)。²²眾所皆知，「呈現手法」(framing) 影響大眾的理解以及個人對新聞報導的認知，²³但是缺乏來自北韓的一手資料，往往難以分辨報導是北韓政府的刻意扭曲或西方新聞媒體的詮釋。Cumings 指出，由於美國媒體常常充斥錯誤消息，因此較好作法是直接分析北韓新聞以取得對美朝關係的見解，儘管新聞內容經常誇大，並且如大部分的共產黨的發聲器一般，公然地「指鹿為馬」。²⁴然而，無論刻意與否，此消息來源避免將國外新聞以美國外交利益的視角呈現。就實際表達的訊息和語調而言，KCNA 呈現現況的手法應與針對國內大眾的政治宣傳略有不同。²⁵由於 KCNA 的英文報導針對國外觀眾，不易察覺其如何塑造議題以博取同情。從這個角度而言，可以此對照一般對北韓政權及其立場的認知。公諸於世和不被報導的新聞同樣重要，因為多數人認為北韓將忽視或低調處理可能有損其形象的事件。²⁶再者，KCNA 的報導不失為可供量化的公開資訊來源。相較多數其他北韓相關研究，本文的作法有助於後繼研究者引以

²² Daniel Kim, *North Korean Foreign Relations in the Post-Cold War Era* (New York: Oxford University Press, 1998), p.44.

²³ Gaye Tuchman, *Making News: A Study in the Construction of Reality* (Beverly Hills, CA: Sage, 1978); D.A. Scheufele, "Agenda-Setting, Priming, and Framing Revisited: Another Look at Cognitive Effects of Political Communication," *Mass Communication and Society*, Vol.3, No.2/3 (2000), pp.297-316.

²⁴ Bruce Cumings, *Parallax Vision* (Durham: Duke University Press, 1999), pp.148-149.

²⁵ Edward Herman, "The Media's Role in U.S. Foreign Policy," *Journal of International Affairs*, Vol.47 (Summer 1993), p.23; Michael Hachten, *The World News Prism* (Ames: Iowa State University Press, 1999), p.11.

²⁶ Bong-uk Chong ed., *North Korea: Uneasy, Shaky Kim Jong-Il Regime* (Seoul: Naewoe Press, 1997).

為基礎，再向前推進。

普遍認為，核武相關議題與六方會談的與會國有關，但迄今的證據顯示，我們無法清楚劃分哪項議題的重要性較高，而兩者的重要性是否恆常不變。傳統上，美國被認為是北韓核武宣示的主要目標，但是缺乏一個可與其他談判國進行比較的廣泛架構，我們無法正確地評估美國對北韓的重要性。平壤的負面話語通常針對美日韓三國。美國對北韓核武計畫的關注是否造成平壤以同類型的語言攻擊美國在東亞的盟友，需要一個各項變數被更嚴格控制的比較。當引述美國和引述核武議題呈正相關時，尤其是當其他因素不變時，可能表示北韓刻意將美國描述為敵人，並間接地拒絕六方會談作為談判的適當途徑。同理，假如在平壤對中俄兩國的言詞中發現類似的模式，其可能暗示著不同的想法與架構。如與會五國皆和核五議題呈正向關係，則可能暗示六方會談為合適的談判架構。然而，如果其他國家皆未在北韓的引述中出現，可能代表平壤某種程度上希望阻絕外界的注意力，並將核武計畫內化為國內議題。相對平壤對美國、南韓和日本的負面宣示，對中國可能以不同語氣呈現，表達中國作為北韓的經濟支柱和政治盟友的特殊角色。中國顧及政權崩潰的隱憂，長期以來不願針對核武計畫施壓北韓。雖然俄羅斯亦明顯關注北韓核武計畫，同樣再三斟酌，並將注意力投注在不以美國利益為首的區域集體安全。²⁷

最後，由於現有資料顯示北韓軍事挑釁的主要對象是南韓，次要對象為美國，將注意力集中在核武發展似乎不甚妥當，尤其當我們納入平壤希望將自己塑造為被害者的可能性。因此較合適的期待應該是，天安艦和延坪島事件將取代核武談判作為新聞報導的重點。同樣地，基於權力交接的考量，政治菁英應該會偏好將金正恩（Kim Jong Un）與爭議性議題之間的連結斬斷。以現有資料為基礎，本文將檢證以下三項假設：

²⁷ Stephan Blank, "Nonproliferation, Russian Style," *Journal of International Security Affairs*, Vol.19 (Fall/Winter 2010), pp.15-22.

假設一：提及「美國」與提及核武議題的次數呈正向關係

假設二：提及「中國」和「俄羅斯」與提及核武議題次數呈反向關係

假設三：提及「天安艦」、「延坪島」和「金正恩」皆與提及核武議題的次數呈反向關係

參、研究設計

作者從 KCNA 英文網蒐集每日新聞，並加以敘述統計（descriptive statistics）和迴歸分析（regression）。²⁸ 儘管作者分別就一段時間和單一日分析文本，每日以一單元計算，呈現多層次分析並在統計資料上突顯更大差異性。除一般字數統計以外，Wordstat 可指出文件中語句的重複性和語詞的相似性；敘述統計和統計推論（inferential statistics）可用於準確評估文本中的意涵（rhetoric）。作者透過整裡自動化內容分析和北韓歷史的相關文獻，將語句和語詞輸入「字庫」（content dictionary）以指認重要的概念和事件。作者直接引述（explicit reference）各類別下的主題，並透過電腦（diagnostic test）發覺可能隱含於其中的間接意涵（implicit reference）。此設計雖然稱不上完美，但作者期望能為後繼研究提供基準。表一列出各類別下的語詞。再者，為避免錯誤分類（inaccurate classification），作者在原測試中指出關鍵字，必要時再進行重新分類（recoding）。²⁹

表一：內容分析的類別

類別	語詞
核武議題	Denuclearization, Denuclearize, NPT, Nuclear, Nuclearization, Nuclear War, Nuke

²⁸ 本文運用 Provalis 的 WordStat 軟體，見：
<http://www.provalisresearch.com/wordstat/WordstatFeatures.html>。

²⁹ 例如，許多關於「美國」（America）和「美國的」（American）的引述與「拉丁美洲」（Latin American）有關。這些引述被「重新分類」（recode）以確定僅有「美國」（U.S.）相關的引述被納入分析中。

天安艦	Cheonan
延坪島	Yonphyong
金正恩	Kim Jong Un
戰爭	War, Warfare, Warhawks, Warlike, Warman, Warmaniac, Warmonger, Wartime
和平	Peace, Peaceful, Peacefully, Peacekeeping
友誼	Friend, Friendly, Friendship
對於北韓	Against the DPRK
主權	Sovereign, Sovereignty
北韓	North Korea, DPRK
南韓	South Korea, South Koreans
日本	Japan, Japanese
中國	China, Chinese
俄羅斯	Russia, Russians, Soviet Union
美國	America, Americans, The United States, The US, USA

資料來源：作者自行整理。

註：作者保持原英文語詞，以表達原意。為求簡單明瞭，複數等其他語詞上的變化含括在分析裡，但並未收錄在表格。

肆、北韓核武發展和政權轉移

平壤的「核武野心」並非新鮮事，早在 1960 年代即在延邊 (Yongbyon) 成立核能研究中心，並編制於蘇聯受訓的專業技術人員進行研發。在接續的二十年，平壤開發重點在核能，直到核武發展計畫的相關證據於 1970 年代末曝光，北韓才在國際壓力下於 1985 年簽署「防止核武器擴散條約」(Treaty on the Non-Proliferation of Nuclear Weapons, NPT)。然而，國際原子能機構 (International Atomic Energy Agency, IAEA) 直到 1992 年才得以進入北韓實行檢查。1993 年，北韓拒絕 IAEA 人員檢視之前未透露的核相關設施，更在同年宣布退出 NPT，成為迄今唯一退出 NPT 的國家。

在聯合國要求北韓開放安檢的同時，美國與北韓在同年展開雙邊談判，並在 1994 年達成框架協議。透過此協議，美國將援助北韓興建輕水核能發電場，以換取北韓中止核發展計畫以及允許 IAEA 人員入內進行檢查。然而，在雙方相互猜忌的情形下，美國將來自北韓官員關於有權擁有

核武的模糊證詞，解讀為平壤持續執行核武計畫，同時不排除平壤已擁有核子武器。2003年，北韓官員首次對中美代表坦承擁有核武，而平壤相繼在2006年和2009年進行核試爆。六方會談等多邊協商自2006年起開始舉步艱辛，更在2009年開始完全終止。在外界壓力有限的情形下，北韓於2010年宣佈完成核融合反應。

關於核武的政治語言絕非無中生有，反之受環境因素和重大事件的影響。例如，我們可以期待南韓領導人更替將影響北韓如何表達。相較保守強調對抗的李明博政府，或許因為南韓在金大中和盧武鉉時期提出「陽光政策」並鼓勵南北雙方直接對話，平壤對首爾在2008年以前的言詞相對不具敵意。由於本文僅關注2010年，更廣泛的環境因素不是討論的重點。就2010年而言，東北亞情勢劇變，有三項變數使國際社群聚焦朝鮮半島：南韓天安艦事件、延坪島砲擊及北韓政權轉移。3月26日，南韓軍艦天安艦在白翎島（Baengnyeong-do）和北方界線（Northern Limit Line, NLL）西岸處沉沒。國際輿論和證據皆指向北韓為滋事者，但平壤否認與事件有關，雖然其在對外宣言中仍間接透露肇事的跡象。事發八個月後，北韓砲擊位於北方界線南邊的延坪島，造成兩名南韓海軍軍官死亡、十五名軍官和兩名當地居民受傷。相對天安艦事件，平壤遂馬上坦承其攻擊乃是針對美韓海上演習以及南韓軍方率先開火的直接反擊。雖然兩韓之間小規模的駁火並非新鮮事（1998、1999、2002、2009年皆曾發生），但延坪島事件的規模和平壤展現的態度皆指向衝突升高的可能性。

前述兩起事件與金正恩預定繼承北韓政權有直接關係。傳統上，獨裁政權內部的權力交接總是充滿許多可能危及政權穩定的不確定性。McEachern指出，黨和軍之間似乎發生政策偏好分離的情形，可能直接影響未來領導人的產出。³⁰外界對「摯愛的領導人」金正日之子的認識不多，僅能從內部的政治宣傳中獲得零碎資訊，KCNA在9月27日以前

³⁰ Patrick McEachern, "North Korea's Policy Process: Assessing Institutional Policy Preferences," *Asian Survey*, Vol.39, No.3 (2009), pp.528-552.

更是對金正恩隻字未提。普遍認為，將初生之犢金正恩推上權力寶座需要高度的政治運作，其中包括對南韓進行挑釁。³¹在金正恩交接為中央軍委會副主席以後，類似的武力展示被認為是凝聚內部支持的必要手段。

2010年的情勢發展將如何影響北韓核武議題目前仍不甚清楚，但是我們也不必然需要將區域情勢的發展與核武議題做連結。首先，當下事件可能只會打散北韓當局形塑國際輿論的注意力；其次，北韓當局更可能為了不加深外界對平壤採取對立性政策的認知，而故意將現況與核武議題脫鉤。無論如何，前述的三項變數不應該與提及「核武」的次數呈正向關係。

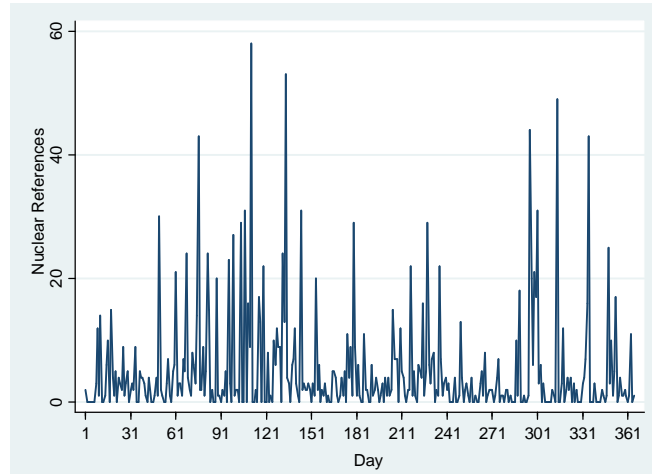
伍、北韓核武議題的自動化內容分析

內容分析的基礎假設是語詞和語句出現的頻率與作者希望在文本中表露的意圖有關。表二呈現 2010 年提及「核武議題」的單日頻率（daily rate）。迅速掃描後可發現，「核武議題」絕非孤立的個別事件，關於「核武」的引述在 259 天內出現 1832 次，每天平均出現五次，中間值是 2。有十個日子，KCNA 引述「核武議題」逾 30 次（依時間順序，分別是：2 月 19 日、3 月 17 日、4 月 17 日、4 月 21 日、5 月 14 日、5 月 24 日、10 月 23 日、10 月 28 日、11 月 10 日、12 月 1 日），其中 4 月 21 日和 5 月 14 日分別引述 58 和 53 次。雖然 KCNA 關於天安艦的首次報導落於五月的「高峰期」之間（5 月 19 日），前述峰期似乎與特定事件無關。表三呈列核武、天安艦、延坪島、³²金正恩等相關報導的頻率。

³¹ 同樣地，北韓 2009 年的挑釁行為可能是金正日試圖贏得軍方支持其指定接班人而採取的行動。Bruce Klingner, "Leadership Change in North Korea—What it Means for the U.S.," http://www.heritage.org/research/reports/2010/04/leadership-change-in-north-korea-what-it-means-for-the-us#_ftnref10。

³² 在中央通信社的報導裡，「Yeonpyeong」（延坪島）一詞從未被引用。然而，依據北韓拼法的「Yonphyong」一詞基本上指延坪島事件，在事件發生前僅出現兩次。在兩次案例中，中央通信社皆指涉延坪島附近聚集的敵軍。本文將關於「Yonphyong」的分析視為「Yeonpyeong」以求一致。

表二：「核武議題」的單日引述次數



值得一提的是，就出現的頻率以及日期而言，KCNA 投注在天安艦、延坪島、金正恩等的整體注意力顯然比核武議題來的低（見表三）。

表三：KCNA 報導各項議題的頻率

	頻率	時間(日)	TF-IDF
核武議題	1832	259	273
天安艦	560	134	243.7
延坪島	105	30	113.9
金正恩	37	27	41.8
戰爭	3754	338	125.3
和平	2458	352	38.7
友誼	2095	342	59.2
對於北韓	720	238	133.7
主權	521	252	83.8

資料來源：作者自行整理。

就西方報導的角度而言，或許此發現出乎意料，但是若將北韓核武的發展歷程以及其他三項變數納入考量，結果不全然值得驚訝。例如，KCNA 雖然在早期的政治宣傳中曾提及金正恩，但是直到 2010 年 9 月 27 日以

前，新聞社對北韓政權的預定接班人隻字未提。「語詞頻率 – 反向文本頻率」(term frequency-inverse document frequency, TD-IDF) 的比重被用於準確度量語詞在一系列文本中的重要性；以此頻率評估，同樣顯示「核武議題」在文本中的重要性。TD-IDF 以「反向頻率」度量語詞在文本中的重要性，隨著語詞出現頻率的上升而降低其重要性(比重)：

$$TF-IDF = (N_{ij}/N^*_j) * \log (D/D_i)$$

N_{ij} 代表 i 在 j 文本中出現的次數， N^*_j 表示文本 j 的字數， D 代表文本的數量， D_i 則表示文本中出現 i 的數量。此公式賦予 i 在 j 文本中固定比重 (assigned weight)；當 i 在少數文本中出現時，則該比重上升；當 i 在文本中出現的次數減少或散佈在許多文本中，比重遂降低。當 i 出現的頻率高並散佈在許多文本中，比重則最低。前述方法控制天安艦等事件可透過其他管道報導，換言之，天安艦事件的相關引述不會在事件發生前出現。以前述公式計算，僅天安艦事件接近核武議題相關報導的重要性。

其次，作者分析相關重要概念出現的頻率。就某種程度而言，概念之間的取舍純屬主觀，因此作者以「相當低的主觀認知」將一些實際上與核武脫離不了關係的概念納入，包括「戰爭」、「和平」、「友誼」和「主權」。此外，常被引用的「對於北韓」(against North Korea) 一詞也包含在其中。內容分析呈現兩個有趣的結果：(一)「戰爭」、「和平」和「友誼」等詞被引述的頻率相近，但是就文本整體而言，「戰爭」的重要性遠高於其他；(二) 前述所有概念的重要性皆遠低於「核武議題」和「天安艦」的 TD-IDF。

就某個角度而言，「概念」本身僅呈現 KCNA 所透露的訊息的一個面向。另一個更直接的面向是關於其他國家的引述，因為其反應北韓政權予與其他國家的重要性。作者將所有的國家納入字庫中，而如同普遍認知，六方會談的其他成員國佔據分析結果的前五名。表四呈現每日和加總頻率 (total frequency)，國家之間擁有未被預期的大幅度差異 (unexpected

variation)。對比南韓的結果，「日本」的數據出乎預料地高，因為普遍的認知是北韓的政治宣傳主要針對南韓。此外，中俄之間引述頻率和時間的落差似乎也和中朝兩國關係的升溫不無關係，恰好與 Haynes 認為「北韓投注在中俄兩國的注意力幾近相同」的結論相反。³³再者，或許更重要的是「美國」的低頻率。相較其他五國，只有「美國」被引述的時間低於半年（180 天），其引述頻率（334 次，頻率值 116.9）也僅達第四位俄羅斯的一半。若排除北韓對內政治宣傳猛打美國所造成的刻板印象，此發現與一般認為「美國對北韓極為重要」的高估呈現極大落差。³⁴

表四：KCNA 引述各國的頻率

	頻率	時間（日）	TF-IDF
北韓	9330	364	11.1
南韓	4090	353	59.4
日本	4240	323	225.1
中國	2184	243	385.9
俄羅斯	833	203	212.2
美國	334	163	116.9

資料來源：作者自行整理。

「迴歸分析」(regression analysis) 可供我們作出更完整的分析，讓我們可以控制個別國家和議題因素的影響，以揭露可能受「多面向特質」(multi-dimensionality) 影響的各項因素的重要性。雖然以「日」作為分析單位有其侷限，但是此作法允許作者在統計上擁有更大的控制空間，並同時度量時間因素 (temporal factor)。就單一模型而言，主要的獨立變數是對於各國的引述次數，以及對於「天安艦」、「延坪島」和「金正恩」所作的控制引述 (controlled reference)。為幫助讀者詮釋，本文包括

³³ Haynes, *op. cit.*

³⁴ 美國的「低頻率」(lower raw frequency)在某種程度上可歸咎於間接引述的缺少。然而，將「洋基」(Yankee)一詞(常見於北韓的文本中)納入美國的類別下對數據並無顯著影響。「洋基」一詞僅出現兩次，兩次皆在美國已被提及的日子，僅將 TF-IDF 變為 117.6。

對於「軍事衝突期間」及「首次引述金正恩」兩者的「控制點」(control)。由於 KCNA 在 1 月 1 日才提及「核武議題」，作者認為不需要在此之前納入「控制點」。表五呈現「核武議題」引述次數的「負二項迴歸」(negative binomial regression) 結果。為測試模型，作者另外加入兩項「規範」(specification)。模型二顧及單日新聞報導可能受近期報導的影響，因此納入「時間差」(lag) 因素 (t-1) (前一日對於「核武議題」的引述)。針對各國，模型三納入更多的「時間差」規範，以顧及可能的「時間性作用」(time effect)。

表五：KCNA 報導引述事件的「負二項迴歸」結果

獨立變數	模型一	模型二	模型三
北韓	0.0518186*** (0.006957)	0.0517579*** (0.0069746)	0.0533795*** (0.0069796)
中國	-0.0112965 (0.0070039)	-0.011976* (0.0070692)	-0.0198417*** (0.00761)
日本	0.0111562* (0.0060844)	0.0112145* (0.0061012)	0.0100526 (0.0066206)
俄羅斯	-0.0387361** (0.0172446)	-0.0382876** (0.0173524)	-0.0476944*** (0.017436)
南韓	-0.0069258 (0.0112132)	-0.0069667 (0.0112655)	-0.0018276 (0.0116173)
美國	0.1325741** (0.056933)	0.1326649** (0.0569696)	0.1259936** (0.0601237)
天安艦	-0.1283027*** (0.03285)	-0.1275228*** (0.0329035)	-0.117557*** (0.0338262)
延坪島	-0.1069031 (0.0848166)	-0.1033597 (0.0853492)	-0.123896 (0.0873029)
金正恩	-0.5078801*** (0.1901373)	-0.5015564*** (0.1910323)	-0.6356042*** (0.1977229)
天安艦事件以前 (pre-Choenan)	-0.035728 (0.1963118)	-0.0271285 (0.197268)	0.0696362 (0.1978603)
延坪島事件以前	-0.2873215 (0.3724293)	-0.268418 (0.3746545)	-0.5190472 (0.4019154)

(Pre-Yeonpyeong)			
引述「金正恩」以前 (Pre-KJU)	0.1534735 (0.2267511)	0.1542674 (0.2269467)	0.1515265 (0.2286201)
時間差 (t-1)			
核武		0.0051684 (0.0097979)	0.0069974 (0.0105626)
北韓			-0.0098745 (0.0071244)
中國			0.0207771** (0.0084479)
日本			0.00501 (0.0066596)
俄羅斯			0.0191766 (0.0150967)
南韓			0.0010848 (0.0113849)
美國			-0.0191071 (0.0594067)
常數 (constant)	0.3631438 (0.3340728)	0.3202466 (0.3439954)	0.5324988 (0.391295)
N	365	364	364
Pseudo R2	0.0439	0.044	0.0488

***p<.01, **p<.05, *p<.10

註：括號裡表示標準誤 (standard error)

在模型一，僅有北韓、美國、日本等三個國家與核武相關引述成正向關係，並具有統計上的意義。雖然與日本的聯結可能出乎預料，此聯結也僅達到.10，可能是所謂的「統計假象」(statistical artifact)。相較之下，北韓、美國兩者與核武的聯結達到 p<.05 以上。美國與核武的關係符合假設一；俄羅斯符合假設二，與核武相關引述在.05 的標準上呈反向關係。中國雖然也和相關論述呈反向關係，其統計意義不明顯。就時事議題而言，關於天安艦和金正恩的引述，兩者皆與核武的引述在.01 的標準上呈反向關係，並符合假設三。雖然係數的變動的方向符合預期，但是相較假設三，關

於延坪島的引述未達到統計意義。

模型二與模型一的結果大致上一致，其中唯一的差異在於中國達到具統計意義的.10標準。模型三的結果引進幾項值得關注的差異。在模型三，中國與核武引述呈反向關係並在.01的標準上具統計意義，日本則未達到統計意義。然而，「時間差變數」(lagged variable) --- 中國 ---呈正向關係，並在.05的標準上具統計意義。換言之，引述「中國」的次數增加與同日引述「核武」的次數減少，但隔日次數隨即上升有關。雖然出乎作者的預期以外，無論是否刻意，此項結果與平壤不認同談判架構並試圖將中國推離六方會談有關。為了控制可能的誤差，作者為引述「核武」逾40次的六個日期植入「虛擬變數」(dummy controls)，重新檢視各項模型，並得到大致上一致的結果。³⁵

就所有的模型而言，可發現幾項明顯因素。各模型的系數指出，就影響力而言，「美國」無疑是最重要的變數。其他變數維持不變，引述「美國」的次數增加，是最能有效推測其他相關引述的指標。此發現符合美國對北韓的認知，即使引述美國的總數相較引述其他與會國來的低。日本同樣在兩個模型取得重要性，但是其影響不明顯，僅達到美國的十分之一。就全面的影響力(系數)而言，呈現最明顯的反向關係的兩個變數不在於「國家」，而在「天安艦」和「金正恩」的引述；此發現與平壤企圖將時事和核武脫勾的假設一致。俄羅斯和中國作為核武引述中的配角則符合作者的預期。相較他國，中俄兩國對北韓保持相對友善的態度，而兩者更期望將會談的本質形容為雙邊性質。然而，此項影響亦不大。最後，值得注意的是，南韓雖然在各模型中呈反向關係，其並未達到任何統計上的意義。雖然南韓是北韓言詞中的主角，平壤似乎未將首爾和核武議題作聯結。

無疑地，此研究僅提供一個北韓如何描述核武議題的截影，其不甚完

³⁵ 關於這些虛擬變數，其中5月14日和12月1日兩個日期是正相關，並在.05的標準上達到統計上的意義。

美，2010年也因為當下因素而不具象徵性。一個時間較長的分析框架能讓我們更清楚地評估核武意涵（nuclear rhetoric）的模式，並指出隨時間演變而轉移的重心發展，究竟多邊會談是否有效，以及究竟其他與會國的領導人更替對平壤的語言是否有影響。然而，此研究發現美國依然是提及核武議題時的焦點，而提及美國時，平壤似乎間接地與其他國家和該年度其他重要事件做區隔。相較傳統上將北韓的新聞報導視為「地毯式宣傳」（blanket propaganda）的觀點，此發現指出平壤有更進一步的盤算。

陸、結論

或許和蘇聯與東德崩解的情形一般，外界可能需等到平壤政權瓦解、內部相關資料開始流出後，才可能一窺北韓內部發展情形。也因為資訊不透明，外界分析常常處於「猜測勝於實證」的情形。自動化內容分析雖然仍無法完全透視平壤政權的認知，但是相較之前研究，不失為一種能更清楚呈現「平壤對核武認知」的方法。研究發現，「核武議題」的引述情況基本上與外界普遍的認知吻合。就其他方面而言，本文試挑戰一般關於北韓的基本假設，並指出外界對於北韓的理解甚缺。若考量美國在核武談判中扮演的角色，以及北韓常試圖將核武協定與雙方外交關係綁在一起，回歸分析的結果並不令人感到意外。

就統計上的差異（「南韓」偏低的 T-F 值）而言，結果可能暗示北韓並不認同六方會談作為解決核武問題的主要管道。如同平壤嘗試將其他議題的本質設定在兩韓或國內範圍，KCNA 的報導似乎表示北韓希望能與美國開啟直接對話。對美方而言，此舉無疑將是對北韓的重大讓步，一個美國尚無意願採取的動作。相對地，KCNA 的報導也透露平壤當局對六方會談的不信任。