

# 資料科學於我國食安風險偵測之應用概況

張格樵 蘇怡鳳 王兆儀

食品藥物管理署食藥戰情中心

## 摘要

隨著巨量資料時代的來臨，世界各國家政府皆面臨數位轉型，運用資料分析將資料轉化為有用的資訊，進而輔助決策。本文彙整我國食品藥物管理署導入資料科學輔助各面向食品安全管理之應用策略，說明食品自邊境輸入到後市場流通之現況與挑戰，透過導入機器學習方法，精準鎖定高風險產品及業者，並強化食品追溯追蹤風險控制；另針對國人高關注食安議題，介紹如何擬定風險偵測方式，輔助實務稽查應用，防堵相關食安事件。隨著各式新興科技日益蓬勃發展，未來可透過拓展資料來源、引入相關技術、進行人才培育等方式，持續強化食品巨量資料之加值應用，提升政府精準決策之能力。

**關鍵詞：**巨量資料、食品安全管理、風險偵測

## 前言

食品安全攸關我國人民健康與國家經濟發展，相關議題深受國人重視，每當食安事件發生往往會造成消費者人心惶惶。為落實食安風險管理，確保國人吃的安全、吃的健康，衛生福利部食品藥物管理署(下稱食藥署)戮力於建構完善的食品安全防護網，惟面對琳琅滿目的產品類別及眾多的食品業者，在有限的經費及人力成本下，如何擇定並優先控管高風險對象，是食安管理的一大課題。

因應巨量資料時代來臨，網路科技發展提供生活的便利性，資料蒐集及流通之速度也隨之提升，使得資料科學領域更加蓬勃發展，資料科學是利用科學方法、流程、演算法和系統，從中提取有用資訊的跨學科領域。食藥署自建置食品雲以來，持續精進食品資訊管理系統，除了整合核心五非系統<sup>A</sup>，並積極串聯跨部會資料，至今食品雲已完成6大部會之資訊系統介接<sup>(1)</sup>，累積大量資料可供應用。為有效發掘政府資料之價值，強化食品安全控管，降低食安事件之發生，食藥署自104年起導入資

<sup>A</sup> 食藥署核心五非系統包括非登不可、非報不可、非追不可、非驗不可、非稽不可等系統，非登不可為「食品業者登錄平臺」，提供食品及藥物業者線上登錄業者及產品資料；非報不可為「邊境查驗自動化資訊系統」，整合進口食品相關廠商報驗及邊境抽驗結果資料；非追不可為「食品追溯追蹤管理系統」，提供食品業者線上申報產品流向資料；非驗不可主要有3大系統，分別為「實驗室資訊管理系統」、「實驗室認證資訊網」及「檢驗管理系統」，提供檢驗案件檢驗資料登錄，整合全國食品稽查檢驗結果；非稽不可為「產品通路管理資訊系統」，整併食品、藥品、化妝品及醫療器材相關衛生稽查、裁處與抽驗紀錄。



料科學方法，進行巨量資料整合分析，鎖定高風險產品及業者，提供政府精準決策之能力，並達到減少抽驗或稽查成本浪費之效。

## 從邊境輸入到後市場流通的食安風險偵測

隨著巨量資料時代的來臨，世界各國皆面臨政府數位轉型挑戰，資料收集方法變得多元且快速，為提升政府巨量資料之運用價值，資料整合分析輔助決策成為資料科學時代之趨勢，然而資料分析之良窳對於後續決策具有重大且關鍵的影響，有效的資料分析能將資料轉化為有用的資訊，進而形成決策支援參考。食藥署為強化食品安全管理，導入巨量資料分析及統計方法輔助食安風險管理，自食品從邊境輸入到後市場流通，採取主動監管方式為民眾食安把關。

### 一、輸入食品智能風險預測，輔助邊境查驗作業

伴隨全球食品貿易蓬勃發展，經濟全球化與食品供應鏈擴大，食安議題已不再受國境拘束，我國食品高度仰賴進口，依據食安法第30條，輸入經衛生福利部公告的食品、食品添加物、食品器具、食品容器，應向食藥署申報產品有關資訊並申請查驗<sup>(2)</sup>。我國過往邊境抽驗是依「食品及相關產品輸入查驗辦法」<sup>(3)</sup>，依報驗義務人過往輸入產品之不合格情形，決定抽驗機率並進行抽樣，而近年輸入食品之重量及批數有成長的趨勢，不合格產品之違規態樣五花八門，儼然成為邊境管理人員的負擔，僅

靠此機制恐難以顧全所有違規情形。

為強化食品邊境管理，輔助邊境食品抽驗機制，食藥署自107年導入人工智慧分析技術，建置「邊境預測智能系統」，應用食品雲、跨部會資料及國際輿情指標等風險因子，結合羅吉斯迴歸、決策樹(CART)、決策樹(C5.0)、樸素貝氏分類、隨機森林、Elastic net與Gradient boosting machine等7種機器學習演算法<sup>(4)</sup>，並採取投票法進行集成，找出邊境高風險報驗案件，輔助查驗人員進行決策，藉以強化邊境食品安全管理，為國內流通之食品建立防線。

110年1月12日已全面導入78個產品中分類<sup>B</sup>，即時掌握風險資訊輔助抽驗決策，截至110年12月31日與導入預測模型前同期效益比較，抽驗查獲命中率提升1.23倍<sup>C</sup>，有效強化食品邊境抽驗管理機制，防堵邊境不合格產品。

### 二、資料科學導入後市場風險偵測，啟動稽查防範未然

民以食為天，食品的安全衛生管理與人民的健康息息相關，然而在成本與獲利的考量下，食品的安全及生產作業環境衛生等恐淪為不肖業者節省成本的下的犧牲品。為使消費者吃得安心，食品安全管理不可獲缺，除了仰賴業者的自主管理，食品安全衛生稽查亦是當中重要的一環，面對五大食品業別眾多之業者，要如何有效運用有限的稽查量能是後市場食安監管的一大挑戰。

我國後市場稽查之標的主要可分為兩大種類，其一是針對食品業者自主管理能力進行稽查，檢視業者是否落實良好衛生管理、及

<sup>B</sup> 為利輸入食品管理、統計及查驗，依產品屬性將輸入食品概分為肉品類、水產品類、農產品類、加工產品類及包括塑膠製奶瓶、免洗筷和瓷製餐具在內的食品器具類等共5大類，每一大類中再依據種類與特性共細分為78個中分類。

<sup>C</sup> 抽驗查獲命中率計算方式：比較模型導入前後之不合格率，模型導入後之不合格案件命中率為1.99%，模型導入前為1.62%，模型導入後之抽驗查獲命中率提升約1.23倍 (1.99%/1.62% ≈ 1.23)

具備預防危害發生的管制措施，確認業者符合食品良好衛生規範準則(Good hygienic practice, GHP)以及食品安全管制系統準則(Hazard analysis and critical control point, HACCP)；另一則是針對食品品質進行檢驗與產品標示正確性之稽查，兩者所針對之目標不同，前者為各家業者，後者為各項產品，為更精進後市場食安監控效能，食藥署自109年導入統計科學及人工智慧技術，結合食藥署、經濟部、財政部及農委會等跨部會資料，針對兩項不同稽查標的及資料特性，擬定合適之風險偵測方式，篩選關鍵特徵因子，透過建模預測找出GHP之高風險業者，並使用統計科學方法推論高風險產品。

110年度建立後市場監測風險預判機制，應用普瓦松迴歸、負二項迴歸及羅吉斯迴歸等統計模型<sup>(5)</sup>，及隨機森林分類器、隨機森林迴歸及Adaptive boosting等機器學習演算法<sup>(4)</sup>，並結合平均序位法進行集成，完成五大業別之後市場監測模型，偵測食安有虞之高風險業者，並產製高風險業者清單，提供予第一線稽查人員實地稽查與抽驗作業使用，強化後市場稽查之效能。

### 三、串聯跨部會食品交易流向資料，強化食品追溯追蹤風險控制

歷年來食安事件層出不窮，如黑心油品、芬普尼汙染蛋品等，造成民眾恐慌。食藥署戮力運用食品雲巨量資料庫之流向資料，藉由強化溯源管理與流向追蹤資訊等巨量資料應用，透過事件前的偵測及監控，以及事件後的追溯追蹤尋找問題，降低食安事件對民眾之衝擊。

此外，為更進一步提升食品安全管理之效能，食藥署自109年建置「食品鏈追溯追蹤異常流向智能監管系統」，透過導入SNA社群網絡分析<sup>(6)</sup>、時序性流向分析及Sankey異常流向監控等分析技術，建立視覺化、可互動式之查詢平台，掌握產品或原料流向、食品業者之

易流向及上下游關係，以網絡密度與網絡中心性作為指標，偵測食品業者於交易網絡所擔任之角色，發掘對消費者影響力較高之業者，優先加強管控，並配合實務需求進行輔導查核。當有食安事件突發時，立即釐清上下游業者交易行為，觀察交易量變化，偵測問題產品之流向，即時下架及回收相關產品，針對問題產品及廠商進行控管，以降低問題發生時對消費者的影響幅度。

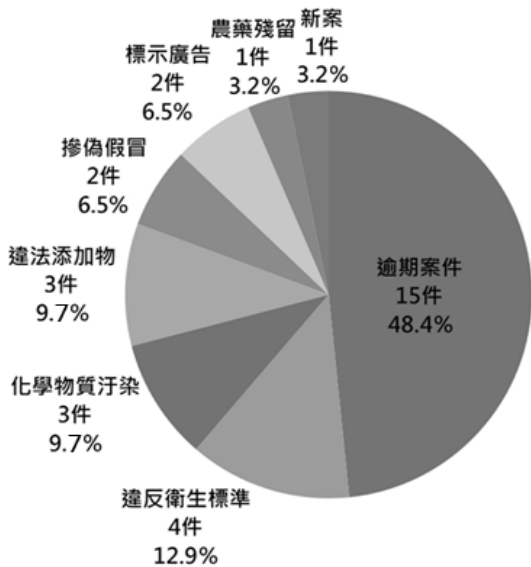
以越南茶葉混充台灣茶葉衍生之食安事件為例，不肖業者利用茶葉不易用外觀辨識原產地之特性，將較低價之越南茶葉混充台灣茶葉，再以提高價出售謀取暴利，食藥署透過核心五非系統資料，結合跨部會資料，溯源產品來源及下游業者鎖定高風險業者，輔以實地稽查，遏止不肖廠商違反行為，強化茶葉標示管理。

## 鎖定高關注食安議題

食品安全是保護健康及提高生活品質之基礎，同時也是社會經濟穩定發展的重要基石，對國家及國際貿易有重要的影響性。107年至109年重大食品違規事件統計(圖一)發現，逾期事件、違反衛生標準、化學物質汙染及違規添加物等事件占比高達80.7%。針對國人高關注食安議題，食藥署透過風險偵測方式，開發風險路徑模式(Risk pathways modeling)，輔助實務稽查，藉以防堵相關食安事件。此外，因應近年來網路科技蓬勃發展，後市場新興產業崛起，食安風險議題層出不窮，食藥署拓廣風險分析目標範圍，依各類風險型態篩選合適之分析手法，提升後市場風險分析之效能。

### 一、食品添加物暨化學物質流向鏈結及風險監控

隨著科技進步，人們除了飲食上越趨精緻，亦追求產品保存與感官上的滿足，市場為



圖一、107年至109年重大食品違規事件統計

順應民眾需求，快速發展多種用於提升風味、增加色澤及防止腐敗等目的之化學成分及其使用方式，該類合法添加於食品上的化學物質即為食品添加物。食藥署頒布多項食品添加物相關法規，期完善我國食品添加物使用規範，惟部分不肖業者不當使用食品添加物或非法添加化學物質於產品製程，進而對人體健康造成危害。

為完善我國食品添加物管理及監控機制，食藥署自108年便著手研擬相關監控分析方法，並於109年建置「化學物質流向鏈結跨部會資訊系統」進行系統性監控分析，透過彙整並參酌歷年國內外相關食安輿情資訊，建立易不當添加之食品級/非食品級化學物質、食品及其對應之違規態樣(如：超標、非法添加、不得檢出、不得殘留)等監控組合，並結合食藥署、經濟部、行政院環境保護署與財政部等跨機關資料，研析化學物質交易流向、食品業者登錄資料及稽查單位稽查紀錄等面向之交易資料，橫縱向追蹤食品業者持有及使用食品添加物及化學物質之狀況，篩選出同時持有組合

中的化學物質及食品之業者，並於排除業者合理持有情形後，產製建議稽查之高風險業者清單，提供第一線稽查人員實地稽查參考。

## 二、逾期食品風險偵測暨管理

近年來屢次發生不肖業者使用逾期原料或改標販售逾期產品等事件，危害國人健康，例如108年知名零食製造業者添加過期原料以降低成本販售牟利，109年烘焙行及早午餐店儲放逾期原料，111年食品原料行將過期食品竄改製造日期及有效日期後，出售給下游廠商等，顯示逾期事件層出不窮遍及各食品產業，107年至109年重大食品違規事件統計發現，逾期事件占比高達48.4%。

政府單位戮力快速追溯食品原料源頭及產品流向，衛生福利部於102年11月19日發布「食品及相關產品追溯追蹤系統管理辦法」，明確規範部分食品業者從事食品及其相關產品製造、輸入、販售業務時應建立流向資料。然雖可運用食品業者輸入、製造及交易等資訊，導入資料分析技術進行風險偵測，然各食品類別(如：冷凍肉品、乳品加工及調味製品)皆有其獨特性，食品效期長短、運輸及保存方式等特性皆有所不同，追蹤產品流向及效期雖能有部分成效，卻難以再更精進；此外，部分業者甫於資料面出現規避之情形，如有業者於產品有效期限將至時，將該產品申報為自用，此舉雖無違法之情事，但存在逾期原料加工再製之風險。

為更有效控管食品逾期風險，食藥署除透過業者申報之收交貨資料進行監管外，更進一步蒐集國內外逾期資料庫、新聞及輿情事件之情資，釐清逾期違規態樣種類，並導入巨量資料分析、統計科學及機器學習等技術，研析逾期食品風險因子，落實巨量資料之加值應用，產出疑似逾期風險產品及業者清單，供第一線人員實地稽查之參考，藉以防範逾期事件發生。

### 三、拓廣後市場風險監測範圍，挖掘新興食安議題潛勢風險

近年來網路科技蓬勃發展，各種線上商業模式興起，在新冠肺炎疫情持續蔓延影響之下，民眾減少外出消費，轉往線上消費及訂購美食餐飲，「宅經濟」衍然已成為時下的風潮，大量的美食外送平台與網購平台如雨後春筍般湧現，小型商家及大型商場為存活於疫情下，開始與平台合作，新興產業及相關食安風險議題亦隨之發酵。

食藥署訂有美食外送及網購食品業者「以通訊交易方式訂定之食品或餐飲服務定型化契約應記載及不得記載事項」<sup>(7)</sup>等相關規範，期保障消費者的權益，但由於網路的便利與無國界等特性，掌握業者追查食品來源，以發掘潛藏於網路販售劣質食品之投機業者，是現行食品安全管理首要目標。

面對新興消費型態之崛起，政府的食安管理之手法亦有與時俱進之必要，為綿密與強化網路通路之食安管理效能，把關「宅經濟」所延伸之食品安全，食藥署利用網路自動化資料擷取技術，以食品雲勾稽國內各大線上平台，挖掘高關注之食品與販售業者資訊，結合資料模糊及精確比對並應用文字探勘技術，針對網路上相關業者資訊、地址、關鍵字等，挖掘潛在風險業者產出清單，提供衛生相關單位加強輔導改善，開發新興消費模式之潛在食品安全風險，落實網路電商平台或網站販售食品之監督管理。

### 討論與未來展望

資訊科技之發展日新月異，人工智慧時代的到來是契機也是挑戰，各項技術領域不斷實現突破性的創新，面臨如此高科技蓬勃發展的時代，政府數位轉型是刻不容緩的要務，食安風險管理亦應與時俱進，對於未來發展可從巨

量資料的5V特點<sup>(8)</sup>為架構進行探討：

- 一、資料大小將決定資料的價值和潛在的訊息，除了持續取得現有資料源不斷增加的資料，亦須開拓更多資料來源，如可透過介接更多跨部會資料，或以爬蟲方式擷取網路資料，藉以補足資料的大量性(Volume)，完備資料分析之全面性。
- 二、食安事件層出不窮，民眾關注之食安議題瞬息萬變，掌握資料分析的快速性(Velocity)，才能夠提供有效的資訊，如食安事件突發時，若未能及時追蹤流向，過長的延遲將會造成更多問題產品被消費者吃下肚。
- 三、資料分析不應侷限於傳統的資料型態，因應科技發展，食品的訊息可能被儲存在網路圖片或影像中，多元類型的資料類型，包括結構化、半結構化和非結構化資料，應對資料的多樣性(Variety)，須提升對各式資料的處理能力。
- 四、由於資料可能來自不同來源端或系統，且資料會隨時間變化，妨礙串聯、處理及管理資料，常見的情形如進行網路資訊擷取時，因原始資料包含網路流行用語，造成同一單詞有不同意義之狀況，導致資料擷取及解讀之結果有誤，面對資料的可變性(Variability)，須不斷釐清定義及彈性調整資料處理方法，並提升所需之技術。
- 五、近年來的科技突破，已大幅降低資料的儲存及運算成本，惟巨量資料裡充斥著雜訊，考驗著分析人員是否有能力洞察出資料背後所存在的巨大價值(Value)，錯誤的提取資料進行分析，會產生糟糕或是無用的結論，合理運用巨量資料分析方法，方能以低成本創造高價值。

綜上所述，巨量資料探勘是對資料進行擷取、挖掘、分析，反映出大量資料背後的規律，建立邏輯關係為決策提供支持。現今資料科學導入我國食安風險偵測，雖具備一定之規

模，仍有可精進之空間，未來可透過拓展資料來源、引入相關技術、進行人才培育等方式，實踐食品巨量資料之加值應用，節省抽驗或稽查成本並聚焦關鍵業者，提升政府精準決策之能力。

### 參考文獻

1. 林文洲。2017。食安相關巨量資料之應用－e起串起來。[[https://www.ndc.gov.tw/Content\\_List.aspx?n=4F76BD4E1FD15647](https://www.ndc.gov.tw/Content_List.aspx?n=4F76BD4E1FD15647)]。
2. 總統府公報。2019。食品安全衛生管理法。108.06.12華總一義字第10800059261號令。
3. 衛生福利部。2019。食品及相關產品輸入查驗辦法。108.06.10衛授食字第1082002676號公告。
4. GÉRON, Aurélien. 2019. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc.
5. Faraway, Julian J. 2016. Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. Chapman and Hall/CRC.
6. Knoke, David, and Song Yang. 2019. Social network analysis. SAGE publications.
7. 衛生福利部。2017。以通訊交易方式訂定之食品或餐飲服務定型化契約應記載及不得記載事項。106.12.25衛授食字第1061303394號公告。
8. IBM. 2016. The 5 V's of Big data - Watson Health Perspectives. [<https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/>].



# An Overview of Data Science Application on Food Safety Risk Identification in Taiwan.

GE-QIAO ZHANG, YI-FENG SU AND CHAO-YI WANG

Decision Support Center, TFDA

## ABSTRACT

As the emergence of the Big data, the governments all over the world are facing the challenge of digital transformation. With the aid of Big data technology, primary data can be transformed into useful information to assist decision making. In this report, we summarized the preliminary achievement by applying data science technology on food safety management by Taiwan Food and Drug Administration (TFDA). Current situations and challenges of food safety issues were demonstrated from border inspection to post-market surveillance by implementing the machine learning methodology, precisely searching for potential high risk food products and food product dealers to enhance food products traceability, and risk control. In order to strengthen the ability to prevent the food safety events, we focused on food safety issues with high public attentions and introduced methods to establish risk detection systems. As variety of new technologies come into existence, the Big data applications on food safety management can thus be improved via expanding data sources, incorporating relevant technical skills and talent cultivation, etc., to enhance the government policy-making capabilities.

Key words: Big data, food safety management, risk identification