

了解大數據對品質改善的影響以更佳地解決與處理客戶疑慮

◎楊沛昇 編譯

品質改善是整體品質管理的關鍵原則，而且對每一個組織都是一項永無止盡的過程。系統化的方法對於改善流程和產品品質對強化組織競爭力來說是至關重要的，

本文討論品質改善和大數據，強調大數據對品質改善的影響。此外，還介紹了一些相關技術和大數據分析舉例。

品質改善

許多組織嘗試發展一套系統化的方式，該方式使用特定的技術來改善品質並減少過程和產品中的浪費。在各個行業中，品質改善通常是透過團隊合作來執行。有三個要素對於成功執行品質改善至為重要（參見圖1）。

1. 品質的概念：了解各種品質概念可以有效地幫助處理品質問題。透過適當的在職訓練，許多企業引導員工學習品質大師的學說，以對品質改善有更好的態度。品質概念的例子包括：品質是由客戶定義的；品質來自預防；越少的表現變動能導致更

高的產品品質；並且應建立有吸引力的品質要素以滿足客戶的潛在需求。這些基本品質原則持續對品質改善的表現產生重大影響。

2. 管理模型：我們必須使用管理模型來解決品質問題，實務中經常使用的管理模型包括PDCA循環，品質管制歷程，福特的8D以及定義、量測、分析、改善與管制(DMAIC)法。其中，由於8D同時強調了矯正與預防問題的價值，因此在高科技公司中被廣泛使用。

3. 品質改善技術：常用的品質改善技術可分為三類。

- (1)統計方法，包括基本統計、假設檢定、迴歸和實驗設計(DoE)或田口方法。
- (2)品質工具，包括品質機能展開，品管七大手法、新品管七大手法、統計製程管制，製程能力分析、量測系統分析以及故障模式影響分析。
- (3)豐田生產系統和精益生產思想已廣泛用於消除浪費、降低成本和周期時間與品質改善在各行業中。

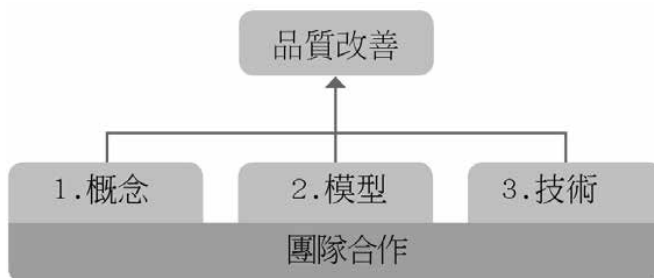


圖1 品質改善方法

品質改善要求解決問題，圖2展現了解決品質問題的數據驅動邏輯。對於給定的問題，您可以收集數據，使用合適的工具進行分析並決定理想的解決方案，修改理想的解決方案以獲得實際的解決方案。

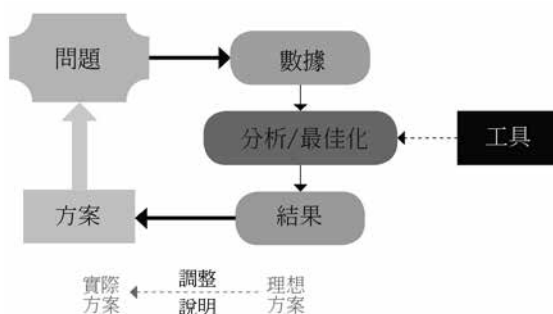


圖2 解決品質問題的數據驅動邏輯

大數據

大數據吸引了學者和從業人員的注意，而產生越來越多的數據是現代技術發展的必然趨勢。例如，您可以輕鬆地在機器和產品中安裝感應器與智慧晶片，以獲得相關資訊，像是產品特性和操作條件。從提供服務到製造，許多組織現在都必須利用大數據。

對於大數據沒有統一的定義，Doug Laney¹用了三個V來定義了大數據：

1. **容量(Volume)**-數據集的大小
2. **速度(Velocity)**-數據進出的速度
3. **種類(Variety)**-各種數據型式與來源

這三個V已成為描述大數據的普遍框架，除了這三個V外，大數據的其他兩個常見的特點包括準確性(Veracity)即數據的品質或可信賴性和價值(Value)即提取數據的價值。

大數據與數據本身無關，相反地，它是關於能更有效解決問題之策略的決定。在本文中，大數據被認為是探討五個V(容

在過程中，確定適當的改善機會，解構問題並說明分析結果是至關重要的。此外，品質改善項目必須與業務戰略目的、客戶意見(VOC)、過程與工程相互聯繫在一起。

量、速度、種類、準確性和價值)的整體方法，從而能夠獲取可行的知識以增強企業的競爭力。科技將隨著時間的推移不斷發展，未來我們將對大數據提出不同的看法。

大數據對品質改善的影響

大數據對品質改善的三個關鍵影響是：

1. **客戶意見**：了解客戶意見（包括內部和外部客戶）是非常重要的，因為在決定有關品質改善方向時，這個資訊對於組織而言有非常有價值。照往例，組織利用調查、訪談、聚焦小組、保固數據、現場數據和抱怨來確定客戶的需求。但是，大數據具有精確評估客戶意見的優秀能力。龐大的數據集和完善的工具讓您能夠確定實際的客戶採購與動機。根據客戶意見可以確定最重要客戶的要求，從而為品質改善提供了充足的機會。
2. **數據收集和分析**：數據收集在品質改善中起有著重要的角色。傳統上，依據不同

情況使用各式數據收集方式。²數據有三種類型：

- + 實驗數據-數據是從設計的實驗中收集，經常使用實驗設計或田口方法來處理此類數據。
- + 觀測數據-透過有計畫的觀測研究對數據進行採樣，通常使用迴歸分析或因果分析來分析此類數據。
- + 歷史數據-已收集的數據，計算智能和數據挖掘方法對於解決歷史數據問題相當重要。

例如當使用DMAIC解決問題時，通常您可以應用實驗設計或田口方法來優化程序。主要原因是，六西格瑪(Six Sigma)很好地支持工程工作，您可以完整地了解問題背景。同時，可以執行妥善規劃的實驗以收集解決問題所需的數據。

另一方面，在處理定義不明(或無法理解)的問題時，您可能只有歷史數據，結果是某些智能方法可以有效地調查數據。

在大數據的時代，可以從如社群媒體、交易、公共數據和機器對機器數據等來源來收集數據。在製造業中，除了歷史數據外，還有大量的即時現場數據構成了大數據。對於這些大數據，您通常不知道要分析什麼，您可能要進行許多次的嘗試與錯誤，但每次的嘗試都可以幫助您進一步了解問題的背景。

大數據可能會挑戰並鼓勵您應用更先進的技術來分析已收集到的大量結構化和非結構化數據集，使用正確的方法可能會揭示大數據中所隱藏的含義。

例如，一家披薩餐廳監控社群媒體並分析張貼的文字和圖片，以決定引起客戶不滿的根本原因。該餐廳設計了一個系統來解決因送餐人員引起的主要問題。³另一個例子是應用關聯規則演算法來決定鑄造廠中不同機器的組合對產量的影響。

3. 預測：大數據可以使您更準確地預測未來，透過大數據的預測，您可以增強流程

和產品性能，同時有卓越的風險管理。您可以發展一個預測模型，以在運作機器容易出現故障之前識別出主要的品質問題，例如，一家鑄造公司藉由工廠製程參數使用神經網路來預測拉伸強度。

預防性的維護已廣泛應用於製造業，許多組織在大數據的幫助下更願意實施預測性的維護，該維護目的在幫助決定運作中設備的狀況以預測何時需要維護。

例如，台灣的一家半導體製造公司利用設備降級模型來預測設備的使用壽命。一家筆記型電腦製造公司嚐試利用客戶要求維修的原因來預測維修筆記型電腦所需的零件。

預測是大數據分析的實際結果，但是，構建有效的預測模型並不容易，因為大數據通常都是非常複雜的，通常很難理解其完整的圖像，並且始終存在著未知之處，過去沒有發生的事情並不能保證將來就不會發生。此外，由於大數據是即時生成地，因此如何準確、快速地預先預測變化點是很重要的。

根據W. 愛德華茲·戴明的教程，⁴如果某個過程處於統計控制下(即處於穩定狀態)，則可以預期未來的變化。如果過程不穩定，則表現是無法預測的。預測大數據需要仔細的評估過程是否處於穩定狀態，從不穩定的過程中獲得的數據可能無法提供對未來的可靠預測。

相關技術

大數據項目通常與大而復雜的問題相關，因此，問題通常是模糊且不清的。所以，您必須首先要決定所要面對的問題以及誰應該加入這個團隊。大數據實現方式類似於圖2所示的架構，但是，需要一些更先進的技術和工具來分析大數據。

大數據分析：分析是利用數學和統計學以發現數據中有意義的模式，使用常規



數據分析方式來處理大數據可能有困難。因此，除了數學和統計學之外，通常還建議使用人工智慧(artificial intelligence, AI)、機器學習(machine learning, ML)和數據

探勘法來處理這些大量的數據。分析大數據以發現大數據的潛在價值並獲得有用的見解以做出更好的商業決策的專業稱之為大數據分析(圖3)。

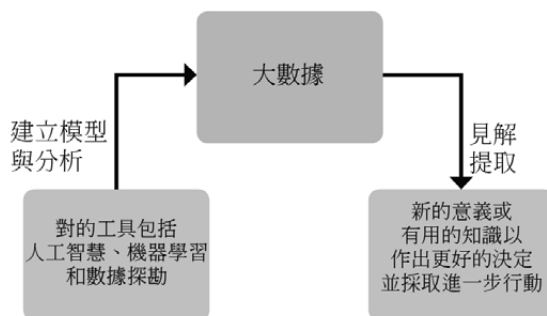


圖3 大數據分析

人工智慧：人工智慧涉及使用計算機解決有關於感知或智力的問題，通過數據處理和演算法運算，人工智慧試圖產生有意義的信息並讓機器比人更聰明。人工智慧技術包括自然語言處理、基於邏輯的推理、計算機視覺、搜尋演算法和機器學習。

機器學習：機器學習是AI的一個分支，對於給定的問題，機器學習收集訓練的數據，從數據中選擇特徵並構建模型。該模型可被視為學習的結果，並用於作出預測或解決高度特定的問題。機器學習方法包括迴歸、神經網路、支援向量機、決策樹和Naive Bayes。

數據探勘：數據探勘是發現模式並在大量數據集中建立關係以更好地了解所研究系統的過程，典型的過程涉及問題定義、數據收集和準備、建立模型、驗證和應用，常見的數據探勘任務包括分類、叢聚、關聯和預測。文本探勘，也稱為文本數據探勘，是從非結構化文本中提取有用

信息的過程。

數據處理：大數據通常是非結構化的，為了提高原始數據的品質與執行結果，需要花費更多時間進行數據的預先處理是

- 1.數據清理：數據可能是不完整、嘈雜且不一致，因此，您應該使用領域知識來提供遺漏值、識別異常值、消除嘈雜數據並修正不一致的數據。
- 2.預先處理分類數據：必須將分類數據轉換為適合的數值，One-hot編碼是一種常用於處理分類數據的技術。
- 3.數據轉換：必須將數據轉換為適當的比例以進行處理，常態化和標準化是眾所周知用於數據轉換的的技術。

特徵選取(FS)：特徵選取(feature selection, FS)是選擇具有有價值特徵或屬性之數字的過程，這些特徵或屬性有助於預測或識別由輸入(Xs)得到的輸出(Y)。特徵選取可以用來簡化、提高準確性並加強對

模型的理解和解釋。常用的特徵選取方法包括相關分析、神經網路中輸入變數相對重要性、元啟發式演算法與決策樹。

舉例

在網際網路上執行數據分析可能需要一個品質資訊平台，該平台可以幫助：

1. **監控**：了解當前情況。
2. **分析**：決定問題的原因。
3. **預測**：預測可能的結果或衍生問題，以準備或預防問題的再次發生。
4. **最佳化**：最佳化目標。

舉例1：一家鑄造公司執行了一項計畫，以收集現場數據來改進流程，基於工程知識，計畫團隊定義了17種可能影響過程輸出-抗拉強度(y)的因子。

首先，團隊試著確定影響y的重要過程因素，執行了五種特徵選取技術，包括神經網路、隨機森林、支援向量機、粗集合理論與回歸分析。根據多數規則，團隊選擇了九個關鍵過程因子進行進一步研究。

下一步，使用神經網路在九個控制因子和相應(y)之間建立非線性的關係，經過訓練的網路被用作遺傳演算法(genetic algorithm, GA)中的適應度函數，將控制因子值轉換為向量(染色體)以表示可能的解答，然後使用遺傳算法最佳化解答。

在這項研究中，執行遺傳演算法20次，執行結果說明，20次運行的標準偏差很小，證明了所獲得解決方案的堅實性。從這20種可能的解決方案中選擇最佳解決方案(具有最高的抗拉強度)，使用這種最佳組合使抗拉強度提高約13.5%。

流程最佳化的中心思維如圖4所示。⁵

舉例2：一家公司經營高速公路電子收費系統(ETC)，該系統使用感應器發射無線電波並偵測附著在汽車上的射頻識別(RFID)標籤。

平均每日交易金額約為1500萬，相對

較低的車輛偵測準確率會給公司造成可觀的財務損失。該公司成立了一個大數據計畫，用於分析進行車輛偵測的ETC數據，以決定影響射頻識別標籤偵測與改善射頻識別標籤偵測率的重要特徵

在這五種車輛型式中，以一輛大型卡車為例，大型卡車的車輛偵測率準確度約為83.4%，從數據庫中取樣了170,500輛大型卡車的記錄數據。

在原始數據中，可能有190個變數影響了車輛偵測率。經過數據預先處理後，剩餘170,000個紀錄偵測到的有141,700條，未偵測到的28,300條。將數據分為訓練集(120,000)和測試集(50,000)，在訓練集中，偵測到100,000個樣本，未偵測到20,000個樣本，表示數據不平衡。

因此，藉由從較小的集合中加入更多例子來執行採樣技術。也就是說，提取的量、未偵測到的車輛數據與偵測到的車輛數據的相匹配。因此，最終的訓練集包括100,000個檢測到的樣本和100,000個未檢測到的樣本。該訓練集用於進一步分析，用於驗證未執行過採樣的測試數據集。

採用五種特徵選擇算法來選出最關鍵的特徵，執行結果決定了29個關鍵變數，從決策樹/C4.5演算法產生了一些有用的規則。

此外，計畫團隊從29個變數中選擇了一些可控制變數，並應用神經網路和遺傳演算法來決定可控制變數的最佳設定。

基於這些分析，可以獲得一些具價值的見解，例如車速控制、RFID標籤放置位置和使用時間以及交通流量的確定，以提高車輛偵測率。

需記得的關鍵原則

大數據已大大影響了品質改善的工作，在使用大數據改善品質時，了解工程問題本身至關重要。但是，三個關鍵原則對於成功進行大數據分析也很有用：



- 1.數據品質：數據是否存在問題，例如與測量、不正確的紀錄或缺少數值等有關的問題。
- 2.分析數據的方法：如何有效、快速地處理大量數據，並透過適當的工具提取可行的資訊。
- 3.客戶觀點：解決的問題是否是客戶所關注的。
要取得成功，您必須擁有高品質的素材與精通的技巧，才能創造出滿足客戶需求的正確味道。

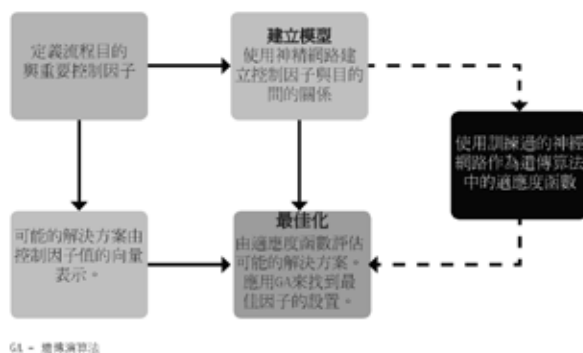


圖4最佳化的中心思維(舉例1)

REFERENCES AND NOTES

- 1.Doug Laney, "3D Data Management: Controlling Data Volume, Velocity and Variety," Application Delivery Strategies, Meta Group, File 949, Feb. 6, 2001.
- 2.Geoff Vining, "Technical Advice: Scientific Method and Approaches for Collecting Data," Quality Engineering, Vol. 25, No. 2, 2013, pp. 194-201.
- 3.Jim Duarte, "Data Disruption," Quality Progress, September 2017, pp. 20-24.
- 4.W. Edwards Deming, The New Economics, Massachusetts Institute of Technology Press, 1993.
5. For similar examples, see Chao-Ton Su's,

Quality Engineering: Off-Line Methods and Applications, CRC Press/Taylor & Francis Group, 2013.

資料來源：

Quality Progress Oct 2019, pp.22-29

Reprinted with permission from Quality Progress ©2020 ASQ, www.asq.org All rights reserved. No further distribution allowed without permission.