

## 搜尋引擎如何運作？

### 09-02

**寶莉在網咖碰到她的同事哈維：**

哈維：天啊，這些搜尋引擎好快。我在網路系統上搜尋只花了 0.09 秒。

寶莉：事實上，搜尋引擎並不搜尋整個網際網路，而只搜尋它們自己的資料庫。例如，號稱有最大索引的 Google 有超過 30 億頁的資料庫。

哈維：他們如何建立起這麼驚人的資料庫？一定是自動化。

寶莉：有人工的指南，但大部分搜尋引擎的資料庫是由稱為「蜘蛛」的程式所建立的。這些自動程式在網路中四處遊走，每「閱讀」一個網頁後，循著網頁的連結到其它的網頁。然後這些網頁便交由另一個程式做成索引。

哈維：蜘蛛程式不可能爬完整個網路，所以一定有很多網站沒被列出來。

寶莉：沒錯。搜尋引擎提供的是「看得到」的網路系統，但是據估計，「看不到」的網路系統有兩三倍之譜。

哈維：好神秘。回到看得見的網路系統，我猜得出後面的部分：搜尋引擎軟體藉由關鍵字或是什麼的在索引裡篩選數十億個網頁以找尋配對，然後以關聯性將結果排序。

寶莉：然而，各家引擎如何決定關聯性則是最高機密。

哈維：我只希望妳的電話號碼不是最高機密。