

不論你同不同意現在是 AI 時代，但你不能反對現在是大資料時代。因為現代人的一舉一動都會產生資料，例如：進行消費時，無論使用實體貨幣、塑膠貨幣或線上支付系統進行支付，都會記錄在資料庫裡。搭乘大眾交通工具，使用刷悠遊卡支付費用，此時悠遊卡也將產生相關紀錄。上街閒晃時，手機會自動向基地台註冊你的位置。無聊時，使用社群軟體交流，運動時運動手環、運動 AP 紀錄路徑軌跡，在家裡休閒時觀看影音平台節目或上網瀏覽等，這些動作不斷地產生資料。

根據 ID 最新發佈的“Data Age 2025”報告，預估在 2025 年時，當年全球所產生的新增資料量將達到 175ZB。其中亞太區(APIxC，不包含中國)預估將由 2018 年的 5.9ZB 增加到 33.8ZB 的新增資料量，位居第三位(第一位為中國:48.6ZB，第二名為 EMEA 區域:48.3ZB)。請注意這僅是新增資料，並未包含歷史資料。可是進行資料分析、探勘時，需要大量的歷史資料。所以資料儲存將變得越來越重要，而這些資料依照不同的結構化程度，分為結構化、半結構化與無結構化資料，這些資料將分別儲存在不同模型的資料庫。

下圖為 [www.db-engines.com](http://www.db-engines.com)(2020/06)的資料庫排行榜(前 20 名)

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	Oracle +	Relational, Multi-model ⓘ	1343.59	-1.85	+44.37
2.	2.	2.	MySQL +	Relational, Multi-model ⓘ	1277.89	-4.75	+54.26
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model ⓘ	1067.31	-10.99	-20.45
4.	4.	4.	PostgreSQL +	Relational, Multi-model ⓘ	522.99	+8.19	+46.36
5.	5.	5.	MongoDB +	Document, Multi-model ⓘ	437.08	-1.92	+33.17
6.	6.	6.	IBM Db2 +	Relational, Multi-model ⓘ	161.81	-0.83	-10.39
7.	7.	7.	Elasticsearch +	Search engine, Multi-model ⓘ	149.69	+0.56	+0.86
8.	8.	8.	Redis +	Key-value, Multi-model ⓘ	145.64	+2.17	-0.48
9.	9.	↑ 11.	SQLite +	Relational	124.82	+1.78	-0.07
10.	↑ 11.	10.	Cassandra +	Wide column	119.01	-0.15	-6.17
11.	↓ 10.	↓ 9.	Microsoft Access	Relational	117.18	-2.72	-23.83
12.	12.	12.	MariaDB +	Relational, Multi-model ⓘ	89.79	-0.30	+4.59
13.	13.	13.	Splunk	Search engine	88.08	+0.33	+3.46
14.	14.	14.	Hive	Relational	78.65	-2.89	-0.40
15.	15.	15.	Teradata +	Relational, Multi-model ⓘ	73.28	-0.60	-3.36
16.	16.	↑ 20.	Amazon DynamoDB +	Multi-model ⓘ	64.87	+0.15	+9.61
17.	17.	↑ 21.	SAP Adaptive Server	Relational	53.09	-0.90	-2.03
18.	18.	↓ 16.	Solr	Search engine	51.26	-1.32	-9.22
19.	↑ 20.	19.	SAP HANA +	Relational, Multi-model ⓘ	50.82	+0.29	-5.56
20.	↓ 19.	↓ 18.	FileMaker	Relational	50.16	-0.80	-7.64

從這份排行榜來看，出現的資料庫模型有：Relational、Document、Search Engine、Key-value、Wide Column 等。其中前十名裡有 6 種資料庫為 Relational，甚至前 4 名全都是 Relational 資料庫。這跟一般人(IT 領域)的感覺不太一致，因為 Relational 應該是已經落伍的架構，現在最新架構應該是 NoSQL 架構才對。其實資料庫模型沒有優劣之分，只是適用何種資料結構之分，根據所要儲存資料結構，選擇資料庫模型才是正確之道。

以下針對資料結構，介紹資料庫模型的優勢：




















### 關聯式模型(Relational)

應用範例	結構化資料：ERP、CRM、SCM 等資料庫應用
資料模型	使用關聯式模型儲存結構化資料 (Table – Row – Column)
優點	資料結構嚴謹，查詢效能佳，有統一性查詢語言 SQL。
缺點	資料結構嚴謹，不可動態變更。

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	Oracle +	Relational, Multi-model ⓘ	1343.59	-1.85	+44.37
2.	2.	2.	MySQL +	Relational, Multi-model ⓘ	1277.89	-4.75	+54.26
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model ⓘ	1067.31	-10.99	-20.45
4.	4.	4.	PostgreSQL +	Relational, Multi-model ⓘ	522.99	+8.19	+46.36
5.	5.	5.	IBM Db2 +	Relational, Multi-model ⓘ	161.81	-0.83	-10.39
6.	6.	↑ 7.	SQLite +	Relational	124.82	+1.78	-0.07
7.	7.	↓ 6.	Microsoft Access	Relational	117.18	-2.72	-23.83
8.	8.	8.	MariaDB +	Relational, Multi-model ⓘ	89.79	-0.30	+4.59
9.	9.	9.	Hive	Relational	78.65	-2.89	-0.40
10.	10.	10.	Teradata +	Relational, Multi-model ⓘ	73.28	-0.60	-3.36





















## 文件模式(Document)

應用範例	Web-Based 相關應用
資料模型	使用 Key 當作存取對應鍵值的 Value 但 Value 為結構化資料(Document)
優點	資料結構不嚴謹，可動態變更。
缺點	查詢效能較差，缺乏統一性查詢語言

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	MongoDB 	Document, Multi-model 	437.08	-1.92	+33.17
2.	2.	2.	Amazon DynamoDB 	Multi-model 	64.87	+0.15	+9.61
3.	3.	 4.	Microsoft Azure Cosmos DB 	Multi-model 	30.80	+0.13	+2.56
4.	4.	 3.	Couchbase 	Document, Multi-model 	29.14	+0.56	-4.22
5.	5.	5.	CouchDB	Document	16.07	-0.85	-3.15
6.	6.	 7.	Firebase Realtime Database	Document	13.99	+0.87	+3.14
7.	7.	 6.	MarkLogic 	Multi-model 	11.24	+0.28	-2.26
8.	8.	8.	Realm 	Document	8.36	-0.02	+0.69
9.	9.	 10.	Google Cloud Firestore	Document	7.52	+1.18	+2.64
10.	10.	 11.	ArangoDB 	Multi-model 	5.38	+0.70	+0.81










## 鍵、值儲存(Key Value Store)

應用範例	內容快取，處理大量的資料讀取
資料模型	使用 Key 當作存取對應鍵值的 Value
優點	查詢速度快
缺點	資料無任何結構

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	Redis 	Key-value, Multi-model 	145.64	+2.17	-0.48
2.	2.	2.	Amazon DynamoDB 	Multi-model 	64.87	+0.15	+9.61
3.	3.	3.	Microsoft Azure Cosmos DB 	Multi-model 	30.80	+0.13	+2.56
4.	4.	4.	Memcached	Key-value	24.81	+0.88	-3.20
5.	5.	5.	Hazelcast 	Key-value, Multi-model 	8.41	-0.27	+0.55
6.	6.		etcd	Key-value	8.05	+0.77	
7. 	8.	6. 	Aerospike 	Key-value, Multi-model 	6.66	+0.60	+0.19
8. 	7.	7. 	Ehcache	Key-value	6.27	+0.04	-0.04
9. 	10.	10. 	ArangoDB 	Multi-model 	5.38	+0.70	+0.81
10. 	9.	8. 	Riak KV	Key-value	4.99	+0.23	-1.03

## 寬欄位(Wide Column)

應用範例	分散式檔案系統
資料模型	使用 column-based 架構將不同 row 的相同 column 資料儲存在一起
優點	查詢速度快，可擴充性強
缺點	功能相對受限

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	Cassandra 	Wide column	119.01	-0.15	-6.17
2.	2.	2.	HBase	Wide column	48.73	-0.99	-9.30
3.	3.	3.	Microsoft Azure Cosmos DB 	Multi-model 	30.80	+0.13	+2.56
4.	4.	4.	Datastax Enterprise 	Wide column, Multi-model 	8.36	+0.74	-0.54
5.	5.	5.	Microsoft Azure Table Storage	Wide column	5.27	+0.59	+0.73
6.	6.	6.	Accumulo	Wide column	3.85	+0.18	-0.26
7.	7.	7.	Google Cloud Bigtable	Wide column	2.93	+0.75	+1.06
8.	8.	8.	ScyllaDB 	Multi-model 	2.71	+0.81	+1.29
9.	9.	9.	MapR-DB	Multi-model 	0.63	+0.03	-0.06
10.	10.	 11.	Alibaba Cloud Table Store	Wide column	0.36	+0.04	+0.14

## 搜索引擎(Search Engine)

應用範例	資料搜索
資料模型	Index (Document)
優點	全文檢索、地理資訊搜索、複雜搜索條件，搜索速度快，可擴充性強
缺點	功能相對受限

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	Elasticsearch	Search engine, Multi-model	149.69	+0.56	+0.86
2.	2.	2.	Splunk	Search engine	88.08	+0.33	+3.46
3.	3.	3.	Solr	Search engine	51.26	-1.32	-9.22
4.	4.	4.	MarkLogic	Multi-model	11.24	+0.28	-2.26
5.	6.	5.	Sphinx	Search engine	6.35	+0.33	+0.03
6.	5.	6.	Microsoft Azure Search	Search engine	6.30	+0.18	+0.44
7.	7.	7.	ArangoDB	Multi-model	5.38	+0.70	+0.81
8.	8.	8.	Algolia	Search engine	4.78	+0.32	+0.40
9.	10.	11.	Amazon CloudSearch	Search engine	2.54	-0.05	+0.00
10.	11.	9.	Virtuoso	Multi-model	2.28	-0.07	-0.83

## 圖(Graph)

應用範例	社群網路、推薦系統
資料模型	Graph
優點	使用 Graph 相關演算法：最短路徑或 N 度關係。
缺點	需要對整個 Graph 進行計算，才能得出所需結果，不太適合分散式結構。

Rank			DBMS	Database Model	Score		
Jun 2020	May 2020	Jun 2019			Jun 2020	May 2020	Jun 2019
1.	1.	1.	Neo4j	Graph	48.27	-1.49	-1.28
2.	2.	2.	Microsoft Azure Cosmos DB	Multi-model	30.80	+0.13	+2.56
3.	3.	4.	ArangoDB	Multi-model	5.38	+0.70	+0.81
4.	4.	3.	OrientDB	Multi-model	4.82	+0.68	-0.77
5.	5.	5.	Virtuoso	Multi-model	2.28	-0.07	-0.83
6.	6.	7.	Amazon Neptune	Multi-model	2.17	+0.41	+0.93
7.	7.	6.	JanusGraph	Graph	2.01	+0.36	+0.46
8.	9.	11.	Dgraph	Graph	1.40	+0.31	+0.51
9.	8.	8.	GraphDB	Multi-model	1.25	+0.06	+0.16
10.	12.	18.	FaunaDB	Multi-model	1.19	+0.25	+0.84

以上介紹 Relational、Document、KV Store、Search Engine、Wide Column、Graph 等資料模型的適用情境、優缺點等，希望讀者可以依照所要處理、儲存的資料結構，正確地選擇資料模型，方能達到事半功倍之效。