

Database: una questione così semplice, così complicata...

Vera Gheno

PUBBLICATO: 9 NOVEMBRE 2017

Quesito:

Database è una parola che in italiano si incontra spesso, almeno in contesti legati all'informatica. Il suo uso genera vari interrogativi, ai quali cercheremo di rispondere tramite questa scheda. Hanno posto domande sulla questione Fausto G., Stella C., Antonella D., Angela P. e Christian P.

Database: una questione così semplice, così complicata...

Grafia
I principali vocabolari italiani riportano, generalmente, due grafie: *data base* e *database*. Sia GRADIT che *Treccani* rimandano dalla forma univertata alla forma separata (dando quella separata come prevalente), mentre lo ZINGARELLI 2018 riporta solo la forma *database*, che nell'uso attuale sembra decisamente più diffusa. Il GDLI registra anche la grafia *data-base*, che si riscontra nell'uso in alternanza con le altre.

Significato

La definizione di *Treccani*, confermata, con minime variazioni, dagli altri dizionari, è:

Archivio elettronico di dati correlati, registrati nella memoria di un computer e organizzati in modo da poter essere facilmente, rapidamente e selettivamente rintracciabili uno per uno, oppure per gruppi determinati, mediante appositi programmi di gestione e di ricerca (chiamati anch'essi *data base*, ma più propriamente denominati *data base management system*, in sigla DBMS).

Il database, in altre parole, è un archivio di dati elettronici organizzati in maniera da facilitarne la gestione. Molti dizionari sincronici registrano il termine come tecnicismo dell'informatica; negli ultimi anni appare ormai diffuso anche al di fuori di contesti specialistici, pur rimanendo prevalentemente in tali ambiti. Con la medesima parola si indicano anche i programmi che raccolgono i dati stessi o che gestiscono l'archivio, come nota anche il GDLI (che la registra nel *Supplemento* del 2004). Del termine, come vedremo, esistono vari equivalenti italiani, tra i quali il più diffuso è *banca dati*.

Etimologia e storia

Database arriva a noi dall'inglese, lingua nella quale, stando all'*Online Etymology Dictionary*, è attestata dal 1962. Il termine è composto da due parole:

1) *data*, neutro plurale del latino DATUM, 'cosa data', participio passato del verbo DĀRE, in uso in inglese all'incirca dal 1640, ma dal 1946 nel significato di 'informazione del computer trasmissibile e archiviabile'

Cita come:

Vera Gheno, "Database: una questione così semplice, così complicata", *Italiano digitale*, 2017, 3 (ottobre-dicembre), pp. 32-35.

Copyright 2017 Accademia della Crusca

Pubblicato con licenza creative commons CC BY-NC-ND 4.0

2) *base*, dall'antico francese *bas* 'profondità', dal latino *BASIS* 'fondazione', a sua volta dal greco *βάσις* 'base, piedistallo', derivato del verbo *βαίνειν* 'andare, camminare, fare un passo'; usato in inglese dal XIV secolo con altri significati, e dal 1959 con quello di '*risorse a cui attinge qualcosa per operare*' (traduzioni mie). La data di ingresso in italiano fornita dalla maggior parte dei repertori lessicografici è il 1979. Una ricerca su Google Libri, tuttavia, ha restituito alcuni esempi precedenti a quell'anno; ciò non è insolito, poiché le banche dati in rete di testi digitalizzati danno l'accesso a un numero sempre maggiore di fonti cartacee, prima difficili da reperire. In tal modo, permettendo retrodatazioni come questa, contenuta negli *Atti del Congresso Annuale dell'Associazione Italiana per il Calcolo Automatico* del 1976:

La comprensione del piano-bilancio viene favorita dallo strumento utilizzato per la rappresentazione degli elementi connessi con la formazione e la gestione del piano-bilancio stesso: **la database** ennaria (7).
Uno sviluppo di tale strumento, comprensibile, ma formale, e la visualizzazione dei vari momenti attuativi del piano-bilancio attraverso quadri rappresentativi dei medesimi, più prossimi alla fantasia dell'utente. Un esempio di ciò è mostrato in figura 1-bis, relativo al processo di pianificazione, la cui database corrispondente è mostrata in figura 1.

Le occorrenze dei primi anni sono prevalentemente in testi scientifici, ma non esclusivamente informatici: più o meno in ogni settore, infatti, oggi si compiono studi su banche dati, per cui è normale che il termine ricorra in testi riguardanti l'industria, l'economia, la geopolitica, ecc. Del resto, l'informatica è meno recente di quanto si potrebbe essere portati a credere. Già nel 1966, per esempio, la nota casa produttrice di computer, IBM, **avviava lo sviluppo di un IMS**, *Information Management System*, ossia un sistema per la gestione delle informazioni, su misura per il programma spaziale statunitense APOLLO.

Stando a Google Ngram Viewer, la frequenza del termine *database* in testi italiani subisce un aumento sensibile nel 1995 e continua a crescere fino al 2004, per poi stabilizzarsi e subire anche una piccola flessione (a favore, apparentemente, di *banca dati*, che aumenta di continuo). L'informazione potrebbe indicare che l'espressione inglese è stata ed è preferita nei contesti specialistici, mentre quando il concetto diventa più comune in testi divulgativi, molti ricorrono all'espressione italiana.

Pronuncia

Come indica la trafilata etimologica, entrambi i termini derivano o passano dal latino prima di arrivare in inglese (il primo elemento in forma più diretta, il secondo seguendo un percorso più complesso); nel valutare la corretta pronuncia della parola in italiano, occorre tenere conto della lingua in cui l'espressione si forma, che è l'inglese. In altre parole, in latino la parola *database*, pur composta di elementi latini o derivati dal latino (ma transitati da altre lingue), non esisteva. Questo fa pendere l'ago della bilancia a favore della pronuncia all'inglese, o perlomeno anglicizzante, come nel caso di altri termini che derivano alla lontana dal latino, ma arrivano in italiano tramite la mediazione di altre lingue, quali *mass media* e *summit*.

La pronuncia inglese è all'incirca /dɛɪtəbeɪz/ (per ascoltarla: <https://www.merriam-webster.com/dictionary/database>), anche se in italiano la parola viene più spesso pronunciata in una versione "spuria" che molti dizionari registrano: /databeɪz/. Del resto, accade sovente che parole straniere che entrano in italiano subiscano un parziale processo di adattamento fonetico per renderle più semplici da pronunciare, come nel recentissimo caso di *influencer* (più o meno /ɪnfluenser/ in inglese, che diventa /influènsɛr/ in italiano come riportato nello ZINGARELLI 2018); in ogni caso, tornando al termine che stiamo trattando, la pronuncia /databàzɛ/ non sembra particolarmente giustificata.

Genere e plurale

In italiano, oggi, il termine o la locuzione sono considerati maschili, quindi *il database* e non *la database*. Google, per le pagine in lingua italiana, conferma tale dato: per “*il database*” troviamo 548.000 occorrenze, per “*la database*” circa duemila; praticamente tutte le occorrenze femminili odierne si riferiscono a contesti in cui *database* è in funzione aggettivale: *la database security*, *la database library* eccetera (e dunque, a rigore, non andrebbero neppure conteggiate, perché l'accordo dell'articolo al femminile riguarda il nome che segue).

Abbiamo visto, però, che nell'esempio del 1976 il termine compare al femminile (*la database ennaria*): un certo grado di oscillazione di genere nei primi anni dall'ingresso di un prestito integrale in italiano è normale (come nel caso di *email* o quello, ancora più recente, di *cheesecake* e *altre leccornie di provenienza statunitense*), dato che in italiano uno dei modi di assegnare il genere a un forestierismo non integrale è quello di considerare il genere del traduttore o dei traduttori più vicini al termine straniero, che in questo caso sono sia *archivio* che *banca* (nella fattispecie, può aver anche influito il possibile accostamento all'it. *base*). Questa oscillazione, comunque, oggi appare definitivamente risolta a favore del maschile, come hanno mostrato i dizionari consultati. Per approfondimenti sul genere dei forestierismi cfr. anche *la scheda di Raffaella Setti in proposito*.

Per quanto riguarda il plurale, i dizionari sincronici mostrano qualche discrepanza: molti considerano il termine invariabile, in italiano, mentre alcuni riportano ancora il regolare plurale della parola in inglese, *databases*. Tuttavia, considerando il buon grado di penetrazione di *database* nella nostra lingua, e seguendo *le indicazioni già fornite su questo sito*, ci sentiamo di consigliare l'uso del plurale invariato: quindi, *i database*.

Diffusione e alternative

Le possibili alternative italiane all'espressione *data(-)base* sono *banca dati*, *banca di dati*, *base di dati* e *base dati*. Consideriamo i numeri sulla diffusione della forma univerbata e quella separata accanto alle espressioni italiane. Per quanto riguarda Google, la ricerca è stata svolta limitatamente a contesti italiani, usando la modalità di navigazione in incognito e la ricerca per frase esatta.

	Google	Google Libri	Archivio di Repubblica
<i>Database</i>	41.900.000	144.000	6.552
<i>Banca dati</i>	4.640.000	90.100	8.191
<i>Base di dati</i>	1.340.000	70.900	662
<i>Data base</i>	520.000	29.900	1.066
<i>Base dati</i>	297.000	12.800	108
<i>Banca di dati</i>	249.000	7.380	13

I numeri sono sufficientemente coerenti: le due forme più usate, in italiano, appaiono essere *database* e *banca dati*. La “forbice” è ampia a favore di *database* se facciamo una ricerca generica, in contesti indifferenziati, mentre si riduce in un contesto più formale quale quello dei libri stampati, per capovolgere, anche se con poco scarto, su uno dei maggiori quotidiani italiani, che potrebbe tenere conto della maggiore necessità di chiarezza da parte del pubblico non di esperti.

In conclusione, volendo essere il più possibile comprensibili e trasparenti, le due opzioni da preferire rimangono *database* e *banca dati*. Tutte le altre forme sono minoritarie, e quindi rischiano di non essere altrettanto chiare per il destinatario.



[Un... database di altri tempi: parte del catalogo cartaceo della biblioteca dell'Accademia della Crusca, prima che lo stesso venisse informatizzato. Oggi lo si trova qui: <http://www.accademiadellacrusca.it/it/biblioteca/catalogo-biblioteca>.]