**Nguồn:**

[stackexchange directory listing (archive.org)](https://archive.org/download/stackexchange)

**Thông tin miêu tả dataset:**

Dữ liệu đến từ mô hình cơ sở dữ liệu quan hệ, được lưu trữ trên nền tảng BigQuery cho trang hỏi đáp StackOverflow.

[Stack Overflow Data | Kaggle](https://www.kaggle.com/datasets/stackoverflow/stackoverflow)

Dữ liệu được tổ chức thành 16 bảng với định dạng csv, có tiêu đề trong bảng, sử dụng các dấu phân cách như [,]. Bộ dữ liệu bao gồm dữ liệu từ năm 2008 đến tháng 5 năm 2021. Có tổng cộng 16 bảng dữ liệu với:

* Badges có kích thước 1.78 GB, số lượng hồ sơ là 39,178,976 và gồm 6 cột
* Comments có kích thước 16.64 GB, số lượng hồ sơ là 79,220,809 và gồm 7 cột
* posts\_answers có kích thước 27.6 GB, số lượng hồ sơ là 31,169,249 và gồm 20 cột
* posts\_moderator\_nomination có kích thước 473.5 KB, số lượng hồ sơ là 334 và gồm 20 cột
* posts\_orphaned\_tag\_wiki có kích thước 53.34 KB, số lượng hồ sơ là 167 và gồm 19 cột
* post\_history có kích thước 109.24 GB, số lượng hồ sơ là 138,808,451 và gồm 8 cột
* post\_links có kích thước 291.1 MB, số lượng hồ sơ là 7,302,589 và gồm 5 cột
* users có kích thước 2.4 GB, số lượng hồ sơ là 14,080,580 và gồm 13 cột
* posts\_privilege\_wiki có kích thước 3.49 KB, số lượng hồ sơ là 2 và gồm 20 cột
* posts\_questions có kích thước 35.58 GB, số lượng hồ sơ là 20,890,054 và gồm 20 cột
* posts\_tag\_wiki có kích thước 36.57 MB, số lượng hồ sơ là 53036 và gồm 20 cột  
  posts\_tag\_wiki\_excerpt có kích thước 11.02 MB, số lượng hồ sơ là 53036 và gồm 20 cột
* posts\_wiki\_placeholder có kích thước 5.63 KB, số lượng hồ sơ là 5 và gồm 20 cột
* stackoverflow\_posts có kích thước 31.52 GB, số lượng hồ sơ là 31,017,889 và gồm 20 cột
* tags có kích thước 2.48 MB, số lượng hồ sơ là 60,534 và gồm 5 cột
* votes có kích thước 6.67 GB, số lượng hồ sơ là 208,577,841.
* Và gồm 4 cột.

Hình ảnh dưới đây là mối quan hệ giữa các bảng:

[GitHub - rodrigoPM/idt-stackoverflow: The main objective of the project lies in the need to build a dimensional model that allows first-hand data analysis to support the strategic decision-making process of the StackOverflow company.](https://github.com/rodrigoPM/idt-stackoverflow?fbclid=IwAR38Oje3xZkhTG3QVsX4WSnPkXvxiYcT1wH410leHyBeMYKjWLEWuS83o28)

Ảnh có chứa văn bản, bản phác thảo, biểu đồ, Kế hoạch

Mô tả được tạo tự động