1. Hevo Data: Hỗ trợ tích hợp dữ liệu được xây dựng trước từ hơn 100 nguồn. Là nền tảng tự động hoàn toàn, không yêu cầu việc quản lý và bảo trì. Kiên trúc fault-tolorent đảm bảo dữ liệu được nhất quán và bảo mật.
2. IBM InfoSphere DataStage: Được xây dựng trên công nghệ microservice và container nên dễ dàng phân tích dữ liệu thời gian thực. Các chuyển đổi từ công cụ cung cấp thông tin phong phú và phù hợp tới tổ chức sử dụng. Các tính năng như tìm kiếm tích hợp sẵn, truyền siêu dữ liệu tự đông và đánh dấu đống thời các lỗi biên dịch.
3. Pentaho Data Integration: Là công cụ ETL mã nguồn mở, hỗ trợ nhiều kiến trúc đám mây. Tập các nguồn dữ liệu lớn.
4. Xplenty: Giao diện thân thiện người dùng, không yêu cầu viết mã. Là công cụ dựa trên nền tảng đám mây, giảm tải việc quản lý và bảo trì định kỳ.
5. Tibco Clarity: Giao diện làm sạch dữ liệu trực quan. Có thể trực quan hóa dữ liệu đang được xử lý, cung cấp cái nhìn rõ hơn về tập dữ liệu. Cấu hình việc xác thực dựa trên quy tắc cung cấp thêm 1 lớp quản lý chất lượng dữ liệu.

References:

<https://www.xplenty.com/blog/top-10-data-cleansing-tools/#xplenty>

https://hevodata.com/learn/8-best-data-transformation-tools/#b4