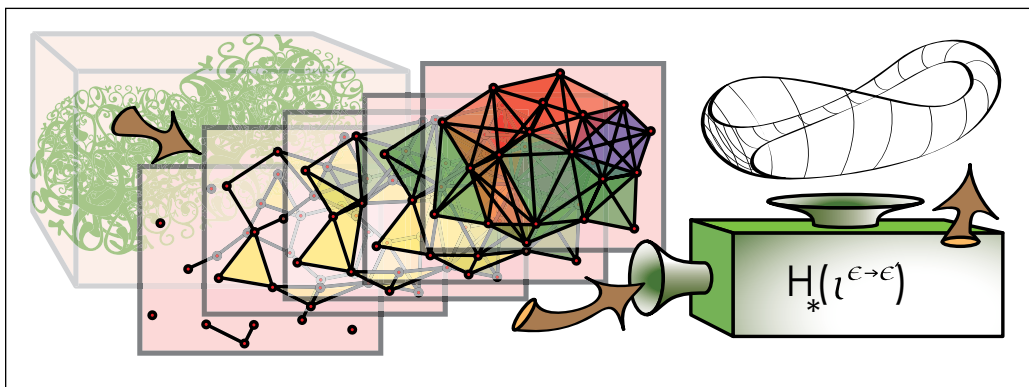




# Phân Tích Dữ Liệu

Rất nhiều nghiên cứu hiện đại, từ chuỗi gen đến khảo sát kỹ thuật số về không gian vũ trụ, tạo ra lượng dữ liệu đa chiều khổng lồ. Thật không may, việc trực quan hóa không gian với số chiều nhiều hơn ba là điều không dễ dàng. Việc phân tích và hiểu dữ liệu cũng vì thế mà trở nên khó khăn hơn. Tô-pô, một nhánh của toán học liên quan đến các tính chất của các cấu trúc hình học, giúp làm sáng tỏ các bộ dữ liệu lớn bằng cách cung cấp một cách phân loại hình dạng của các tập hợp này. Nó đặc biệt hiệu quả để xác định các nhóm điểm tương tự được gọi là “điểm tụ”, các điểm tụ đó rất hữu ích, ví dụ chúng có thể giúp phân biệt các loại bệnh khác nhau, mỗi loại cần một liệu pháp điều trị riêng.

Tô-pô (cụ thể là tô-pô đại số) cũng quan trọng trong hoạt động của các mạng cảm biến không dây, được sử dụng trong các ứng dụng đa dạng như giám sát lưu lượng giao thông ô tô và điều khiển tưới tiêu. Kết hợp với giải tích số, các kết quả từ tô-pô đại số cung cấp hình ảnh toàn cảnh dựa trên dữ liệu cục bộ. Lợi thế là các mạng cảm biến như vậy, được duy trì mà không cần GPS hoặc các phương tiện điều khiển từ xa, thường có chi phí vận hành rẻ hơn. Do đó, trong trường hợp tưới tiêu, các khám phá toán học, có từ trước thời đại công nghệ ngày nay gần một thế kỷ, đã giúp tiết kiệm chi phí, đồng thời giúp chúng ta sử dụng nguồn nước quý giá một cách khôn ngoan. Điều này cho thấy những lý thuyết toán học có thể hữu ích đến mức nào. Hãy nghĩ về dải Möbius mà xem: Hễ có đi thì có đến.



Đồng điều bền bỉ (persistent) của một xấp xỉ ngăn giúp tìm ra các cấu trúc ẩn trong những tập dữ liệu lớn. Ảnh: Robert Ghrist.

**Tài liệu tham khảo:** “Topology and Data” (Tô-pô và Dữ liệu), Gunnar Carlsson, *Bulletin of the American Mathematical Society* (Tập 46, số 2), tháng tư 2009.