**BÁO CÁO PHÂN TÍCH VÀ TÌM HIỂU VỀ TẬP DỮ LIỆU “EEG-MOTOR-MOVEMENTIMAGERY”**

Dataset **"EEG Motor Movement/Imagery"** là một tập dữ liệu về **điện não đồ (EEG)**, được thu thập từ các tình nguyện viên thực hiện các nhiệm vụ vận động và tưởng tượng vận động. Đây là một tập dữ liệu phổ biến trong lĩnh vực **trí tuệ nhân tạo cho IoT y tế**, đặc biệt trong nghiên cứu về giao diện não – máy (BCI - Brain-Computer Interface).

**Thông tin chi tiết về dataset:**

* **Nguồn gốc**: Được phát hành bởi **PhysioNet**, một cơ sở dữ liệu mở về tín hiệu sinh lý.
* **Mục đích**: Hỗ trợ nghiên cứu về **phân tích tín hiệu EEG**, nhận dạng hoạt động não liên quan đến vận động.
* **Số lượng người tham gia**: 109 người (52 nam, 57 nữ, độ tuổi từ 18 đến 47).
* **Dữ liệu EEG**: Ghi nhận từ **64 kênh điện cực** trên da đầu theo hệ thống chuẩn quốc tế **10-20**.
* **Tần số lấy mẫu**: 160 Hz.
* **Tình huống thí nghiệm**:
  + Nghỉ ngơi (eyes open/closed).
  + Cử động tay (trái/phải) và chân.
  + Tưởng tượng cử động tay và chân.

**Ứng dụng của dataset:**

* **Nhận dạng hoạt động não** dựa trên tín hiệu EEG.
* **Phát triển giao diện não – máy (BCI)** cho người khuyết tật.
* **Huấn luyện mô hình AI** trong việc phân loại tín hiệu sinh lý.

**1. Cấu trúc thư mục và tệp dữ liệu**

Tập dữ liệu này bao gồm nhiều tệp **EDF** (European Data Format) chứa tín hiệu EEG thô. Cấu trúc thư mục thường như sau:

* **SxxxRyy.edf**: Mỗi tệp là một phiên ghi nhận tín hiệu EEG từ một người tham gia.
  + Sxxx: Số hiệu người tham gia (S001 đến S109).
  + Ryy: Số phiên thử nghiệm (R01 đến R14).

**2. Nội dung của mỗi tệp EDF**

Mỗi tệp .edf chứa các tín hiệu EEG được ghi nhận từ 64 điện cực đặt trên đầu người tham gia. Các thông tin chính bao gồm:

**(a) Header Metadata (Thông tin chung)**

* **Patient ID**: Mã định danh người tham gia.
* **Recording Date & Time**: Thời gian ghi dữ liệu.
* **Number of Channels**: Số lượng kênh EEG (thường là 64).
* **Sampling Frequency**: 160 Hz.
* **Duration**: Thời gian ghi nhận.

**(b) Tín hiệu EEG**

Mỗi tệp chứa dữ liệu từ **64 kênh EEG**, tương ứng với các vị trí điện cực theo hệ thống quốc tế **10-20**. Một số kênh tiêu biểu:

* **Fp1, Fp2** – Trán.
* **C3, C4** – Vùng vận động.
* **O1, O2** – Vùng chẩm (xử lý thị giác).
* **T3, T4** – Vùng thái dương.

Mỗi kênh sẽ có một chuỗi giá trị điện thế theo thời gian.

**3. Cách tổ chức thử nghiệm**

Mỗi người tham gia thực hiện **14 phiên thử nghiệm** với các tác vụ khác nhau:

**4. Tệp sự kiện (.edf.event)**

Ngoài dữ liệu EEG, tập dataset này có thể chứa tệp .edf.event để lưu lại **thời điểm diễn ra sự kiện** trong quá trình ghi nhận, ví dụ:

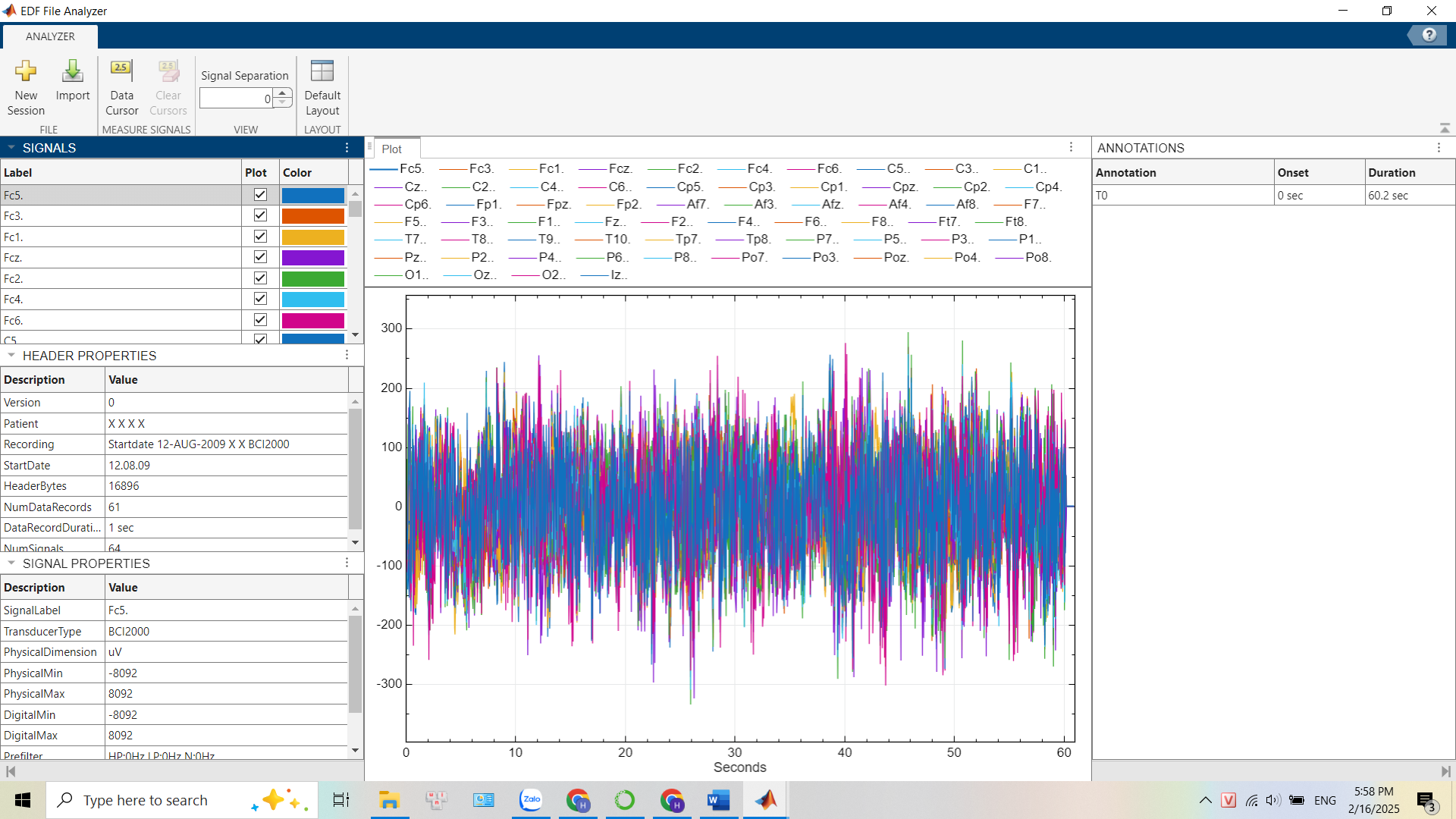
* **Bắt đầu thử nghiệm**
* **Hướng dẫn thay đổi trạng thái (mở mắt, nhắm mắt, vận động, tưởng tượng)**
* **Thời điểm thực hiện một hành động cụ thể**

**5. Sơ đồ dữ liệu của người thử nghiệm (chỉ người đầu tiên) (đọc bằng MATLAB):**

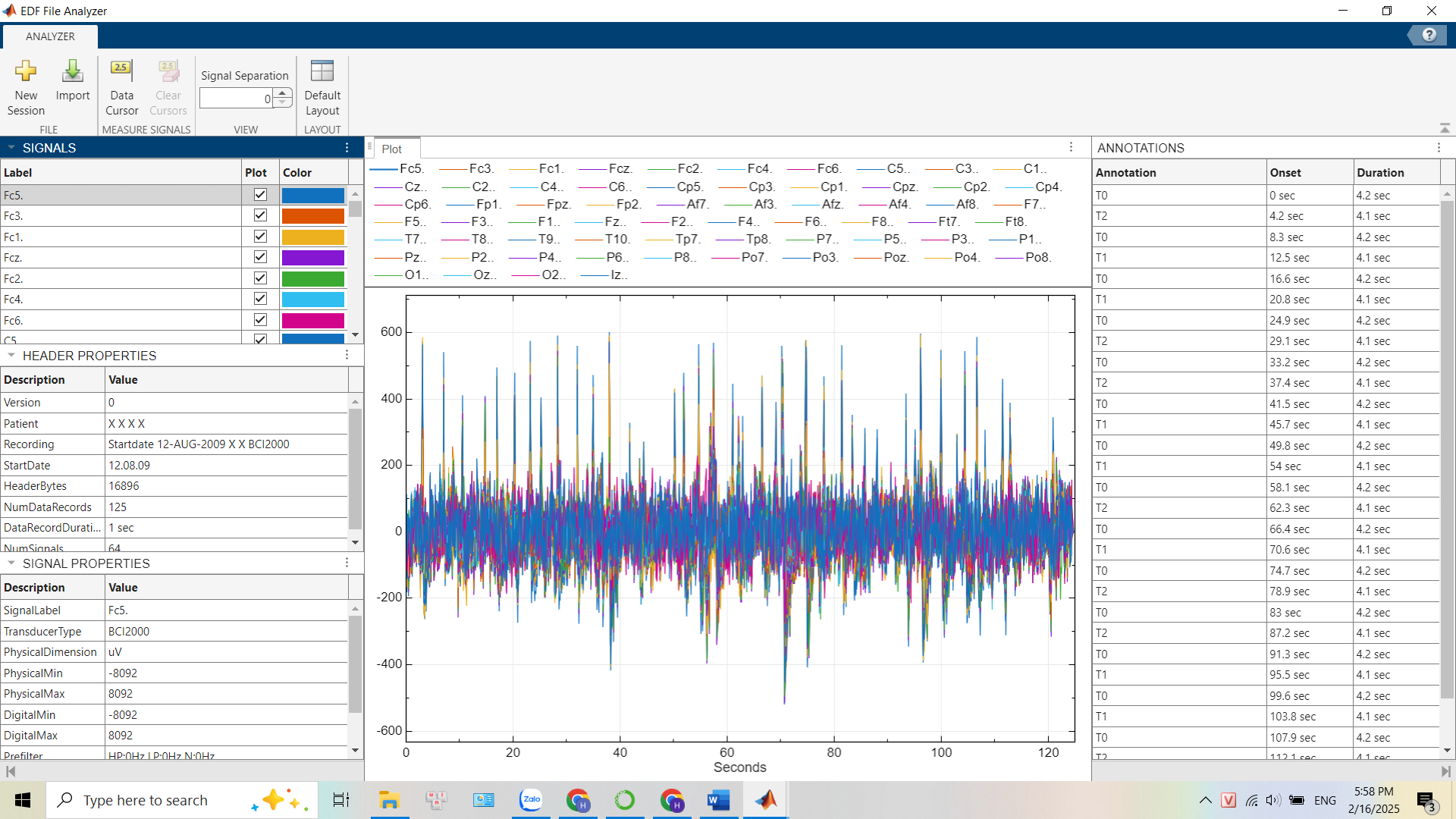
**Phiên 1:**



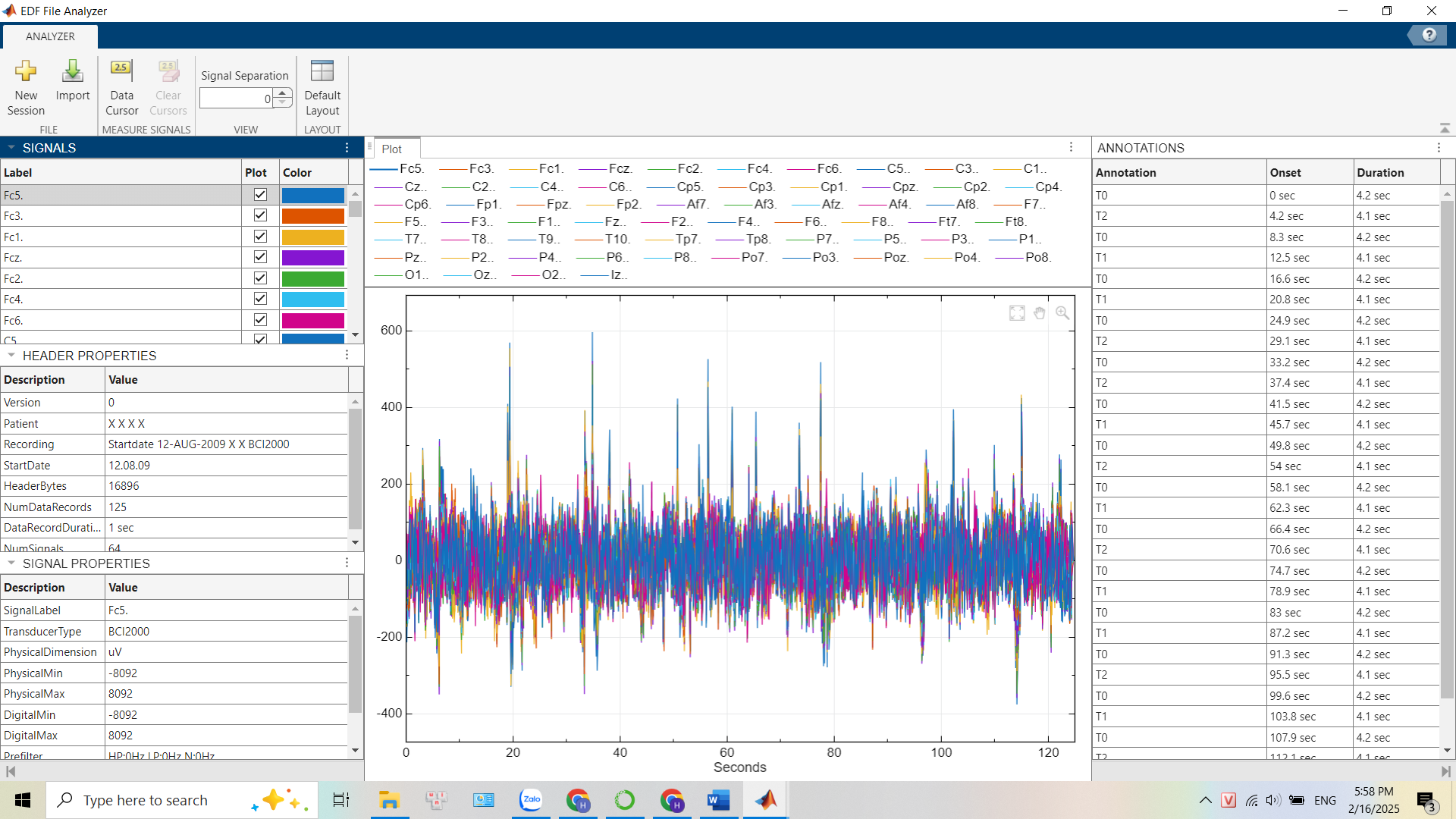
**Phiên 2:**



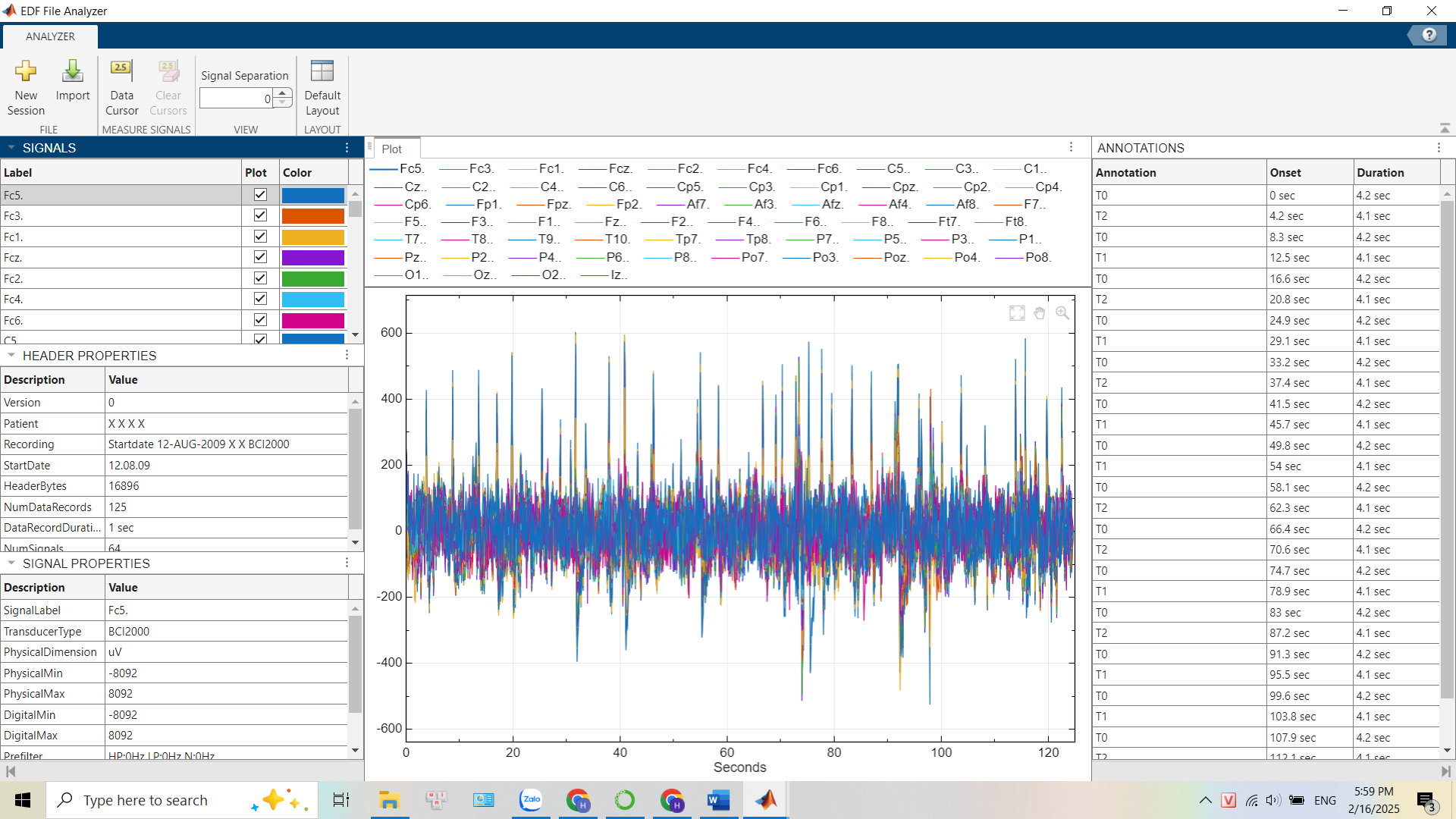
**Phiên 3:**



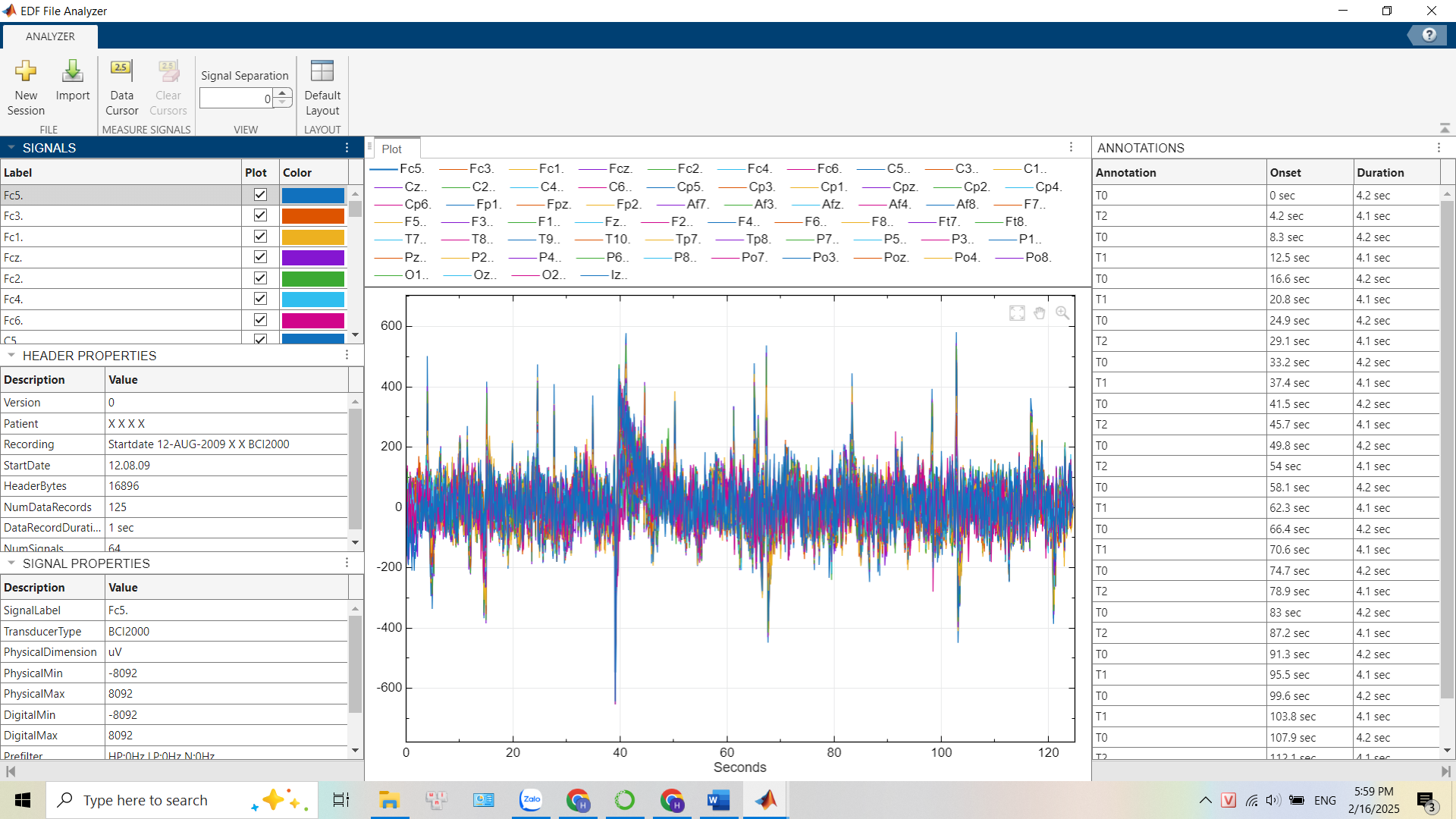
**Phiên 4:**



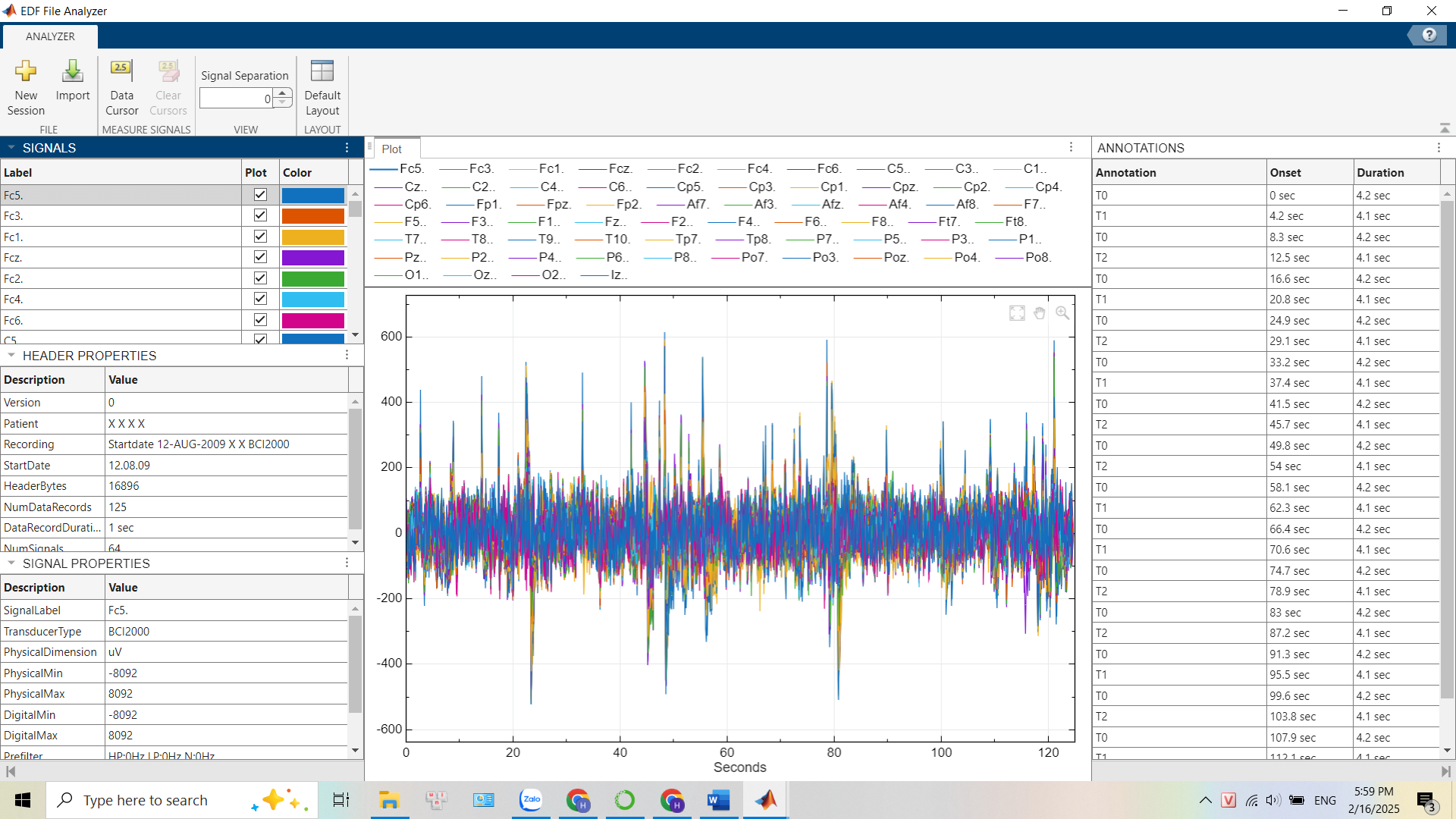
**Phiên 5:**



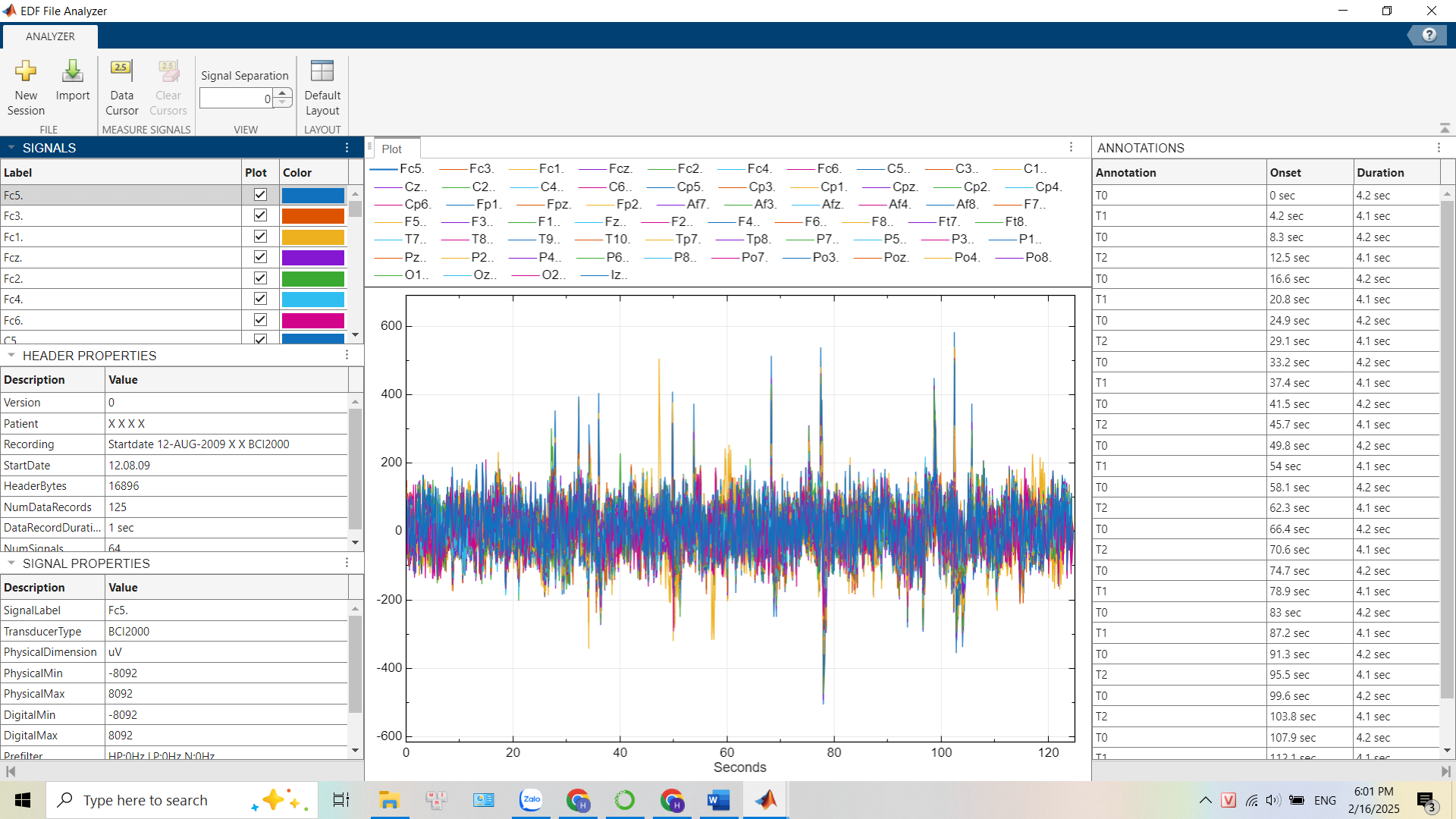
**Phiên 6:**



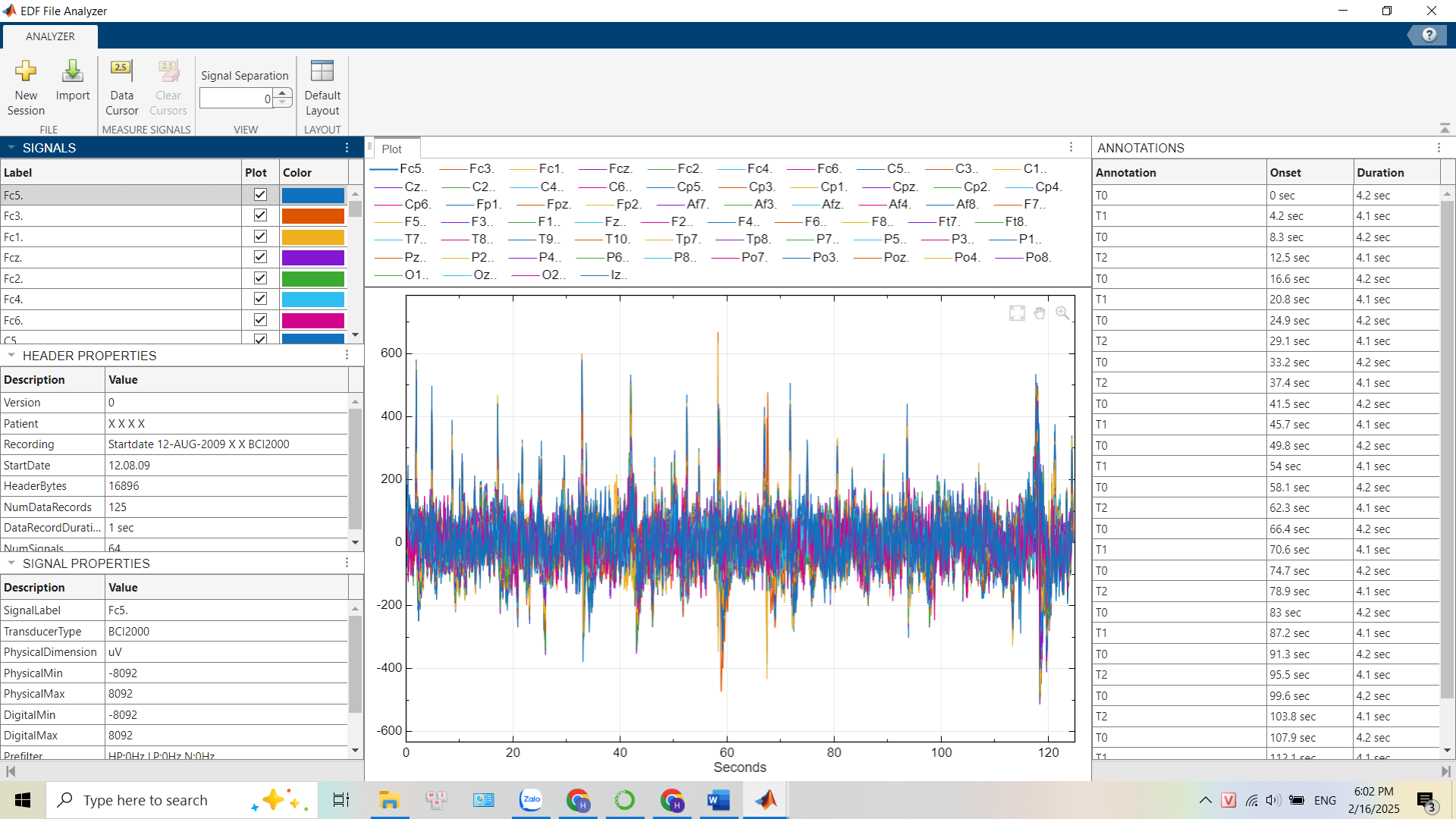
**Phiên 7:**



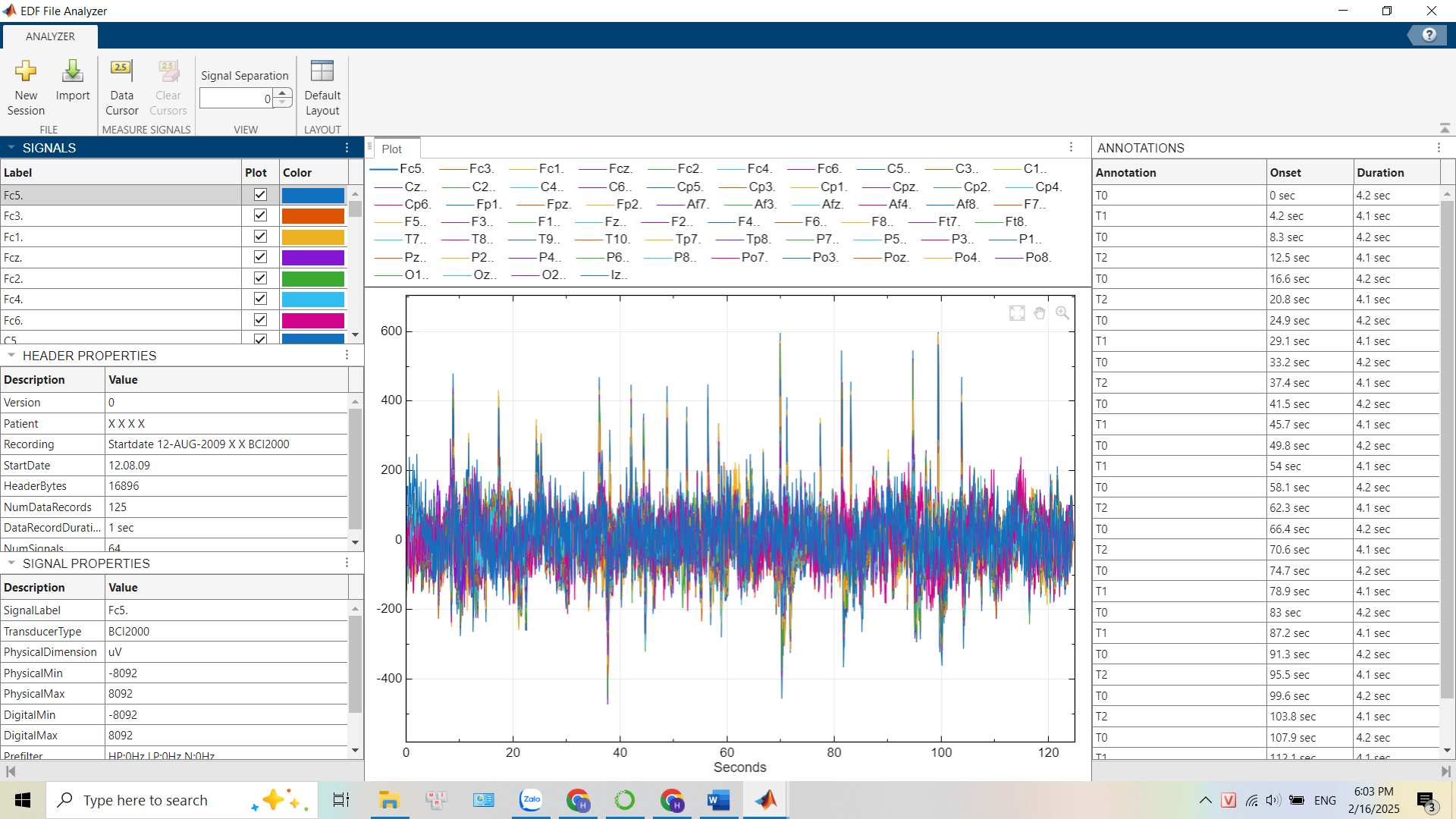
**Phiên 8:**



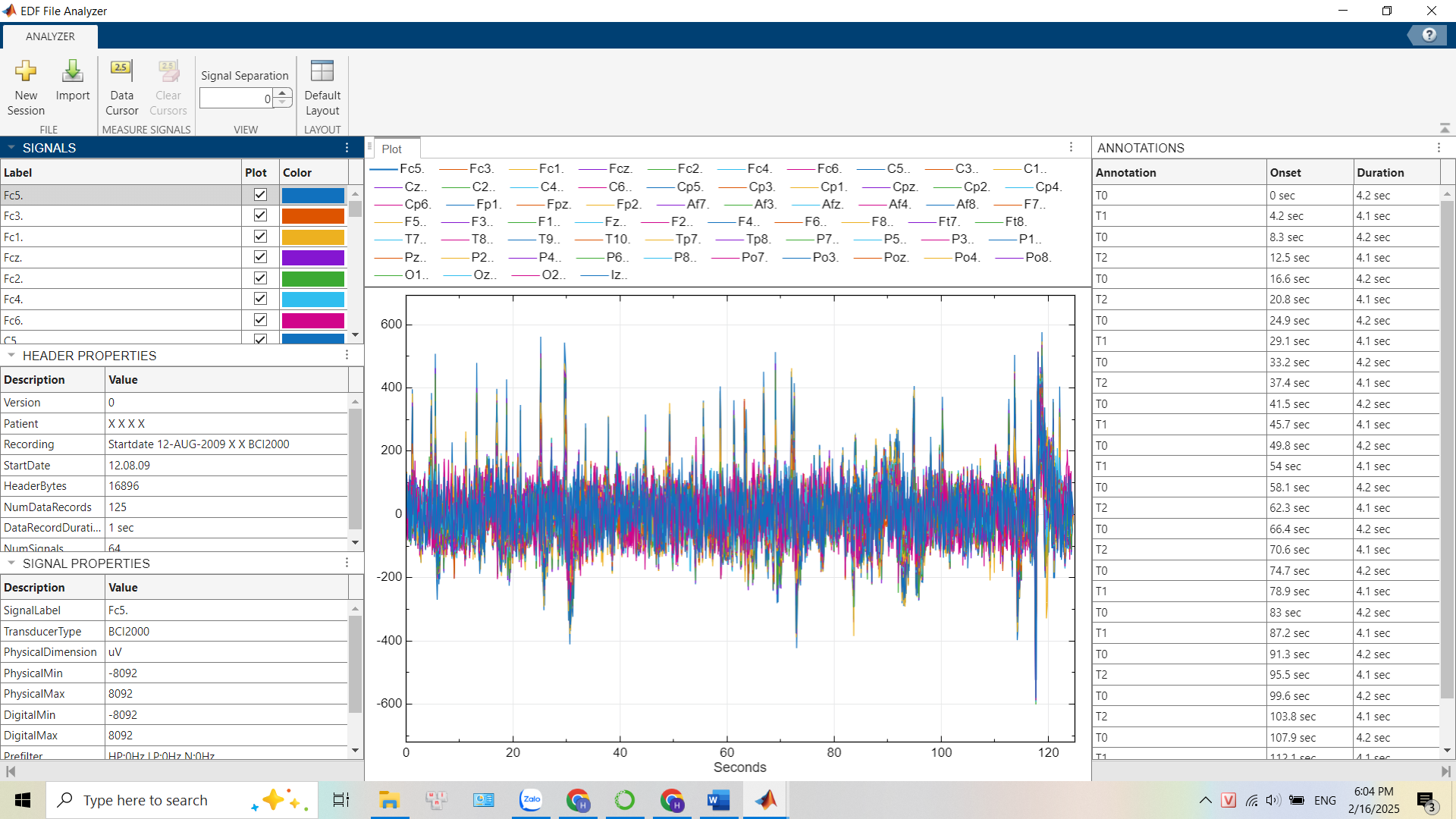
**Phiên 9:**



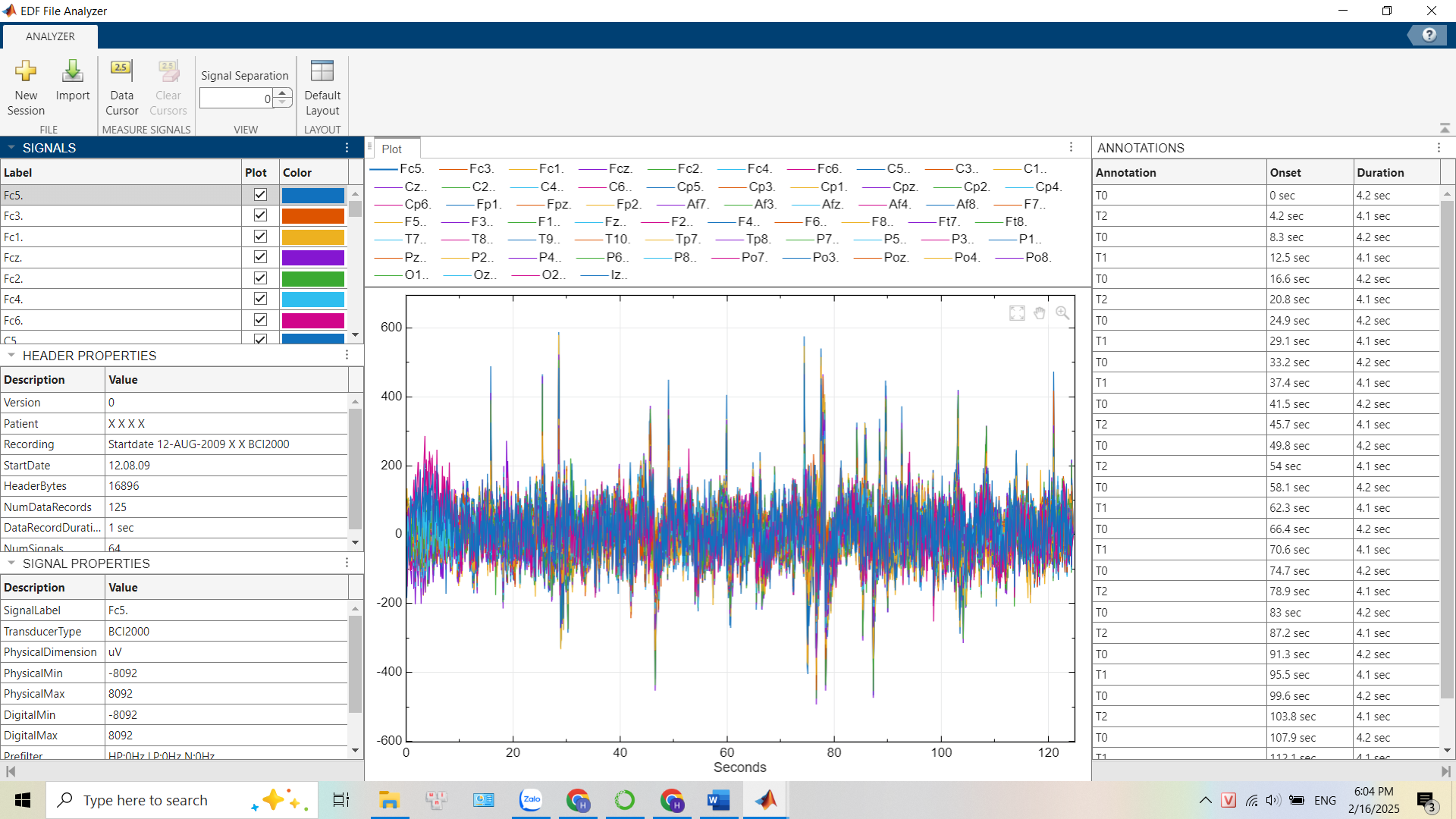
**Phiên 10:**



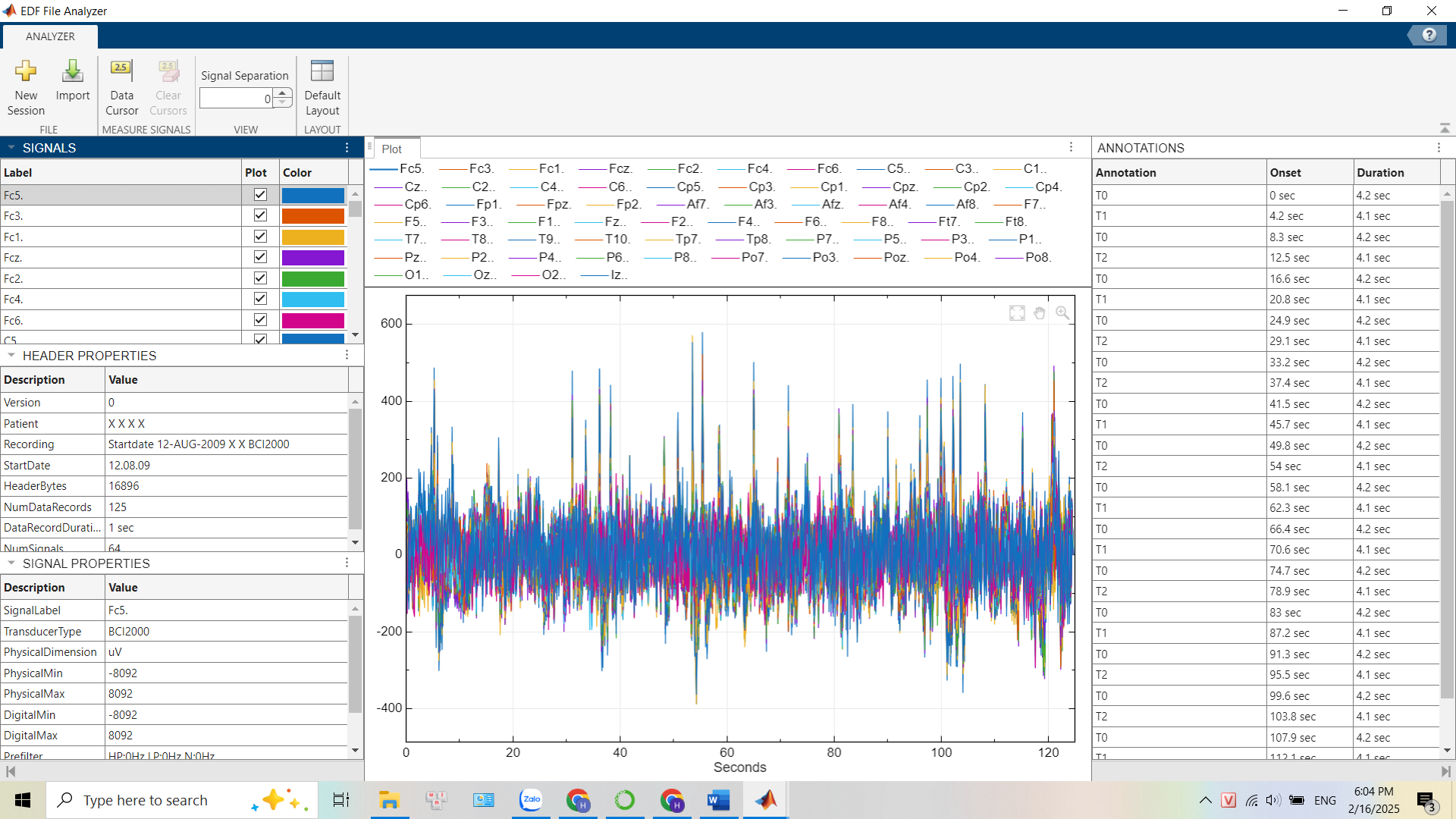
**Phiên 11:**



**Phiên 12:**



**Phiên 13:**



**Phiên 14:**

