Bộ Quy Tắc Gán Nhãn Chi Tiết (V5-Final-Static+URL+ExternalContext)

Mục Tiêu: Phân loại loại hình tấn công phishing một cách **nhất quán và có thể biện luận được** dựa trên phân tích **kết hợp URL gốc và mã nguồn HTML tĩnh**, có xem xét đến thông tin từ các nguồn bên ngoài (nếu có).

Tuyên Bố Miễn Trừ Trách Nhiệm & Hạn Chế Cốt Lõi (BẮT BUỘC PHẢI HIỂU VÀ GHI NHẬN):

1. **KHÔNG THỰC THI JAVASCRIPT:** Quy tắc này **không** mô phỏng trình duyệt hoặc thực thi bất kỳ mã JavaScript nào. Mọi phân tích chỉ dựa trên URL và mã nguồn HTML tĩnh được trả về ban đầu.
2. **BỎ SÓT PHISHING TINH VI:** Do không thực thi JS, phương pháp này **sẽ bỏ lỡ đáng kể** các kỹ thuật phishing hiện đại dựa vào JS để tạo nội dung động, che giấu yếu tố độc hại, và lẩn tránh phát hiện.
3. **ĐỘ NHẠY (RECALL) THẤP:** Hệ thống/người gán nhãn tuân theo quy tắc này sẽ có độ nhạy thấp, đặc biệt trong việc phát hiện Malware Distribution (MW) và các loại hình khác phụ thuộc nhiều vào tương tác động.
4. **ƯU TIÊN ĐỘ CHÍNH XÁC TĨNH (PRECISION):** Quy tắc này ưu tiên gán nhãn (CH, MW, FIN, PII) chỉ khi có **bằng chứng tĩnh mạnh và tương đối rõ ràng** để giảm thiểu gán nhãn sai dựa trên suy đoán yếu.
5. **CẦN KẾT HỢP:** Để đạt hiệu quả thực tế, kết quả từ quy tắc này **nên** được kết hợp với dữ liệu từ các nguồn bên ngoài (Blacklists, Threat Intelligence APIs).

I. Các Lớp Mục Tiêu (Target Classes):

1. **Credential Harvesting (CH)**
2. **Malware Distribution (MW)**
3. **Financial Scam (FIN)**
4. **PII Gathering (PII)**
5. **Other Phishing (OTH)**
6. **(Tùy chọn) Legitimate (LEG)**

(Loại bỏ hậu tố "-Static+URL" để đơn giản hóa, nhưng bản chất vẫn dựa trên phân tích đó)

II. Bước Kiểm Tra Sơ Bộ (Preliminary Checks - Nếu có khả năng tích hợp):

- **Hành động (Khuyến nghị mạnh mẽ nếu có thể):** Trước khi phân tích chi tiết HTML/URL, kiểm tra URL/Domain với các nguồn bên ngoài:

    - **Blacklists:** PhishTank, OpenPhish, Spamhaus DBL, etc.
    - **Threat Intelligence APIs:** VirusTotal (URL/Domain report), Google Safe Browsing API.
- **Xử lý Kết quả Sơ bộ:**
    - Nếu nguồn uy tín đánh dấu rõ ràng là **Malware Distribution** => Có thể ưu tiên gán nhãn **MW** (và vẫn nên kiểm tra tĩnh để xác nhận nếu có thể).
    - Nếu nguồn uy tín đánh dấu rõ ràng là **Phishing** (chung chung) => Thông tin này rất hữu ích, đặc biệt nếu phân tích tĩnh không tìm thấy dấu hiệu mạnh của CH/MW/FIN/PII (có thể dẫn đến OTH).
    - Nếu nguồn uy tín đánh dấu là **Clean/Legitimate** => Có thể ưu tiên gán nhãn **LEG**.
- **Lưu ý:** Bước này là *bổ sung*, quy tắc chính bên dưới tập trung vào phân tích URL/HTML tĩnh.

III. Định Nghĩa Chi Tiết và Dấu Hiệu Then Chốt (URL & HTML Tĩnh):

(Nhấn mạnh bằng chứng tĩnh cần thiết và sự kết hợp với URL)

1. **Credential Harvesting (CH):**
    - **Mục tiêu (Suy đoán):** Đánh cắp thông tin đăng nhập (thường là username/password).
    - **Dấu hiệu Chính yếu Tĩnh (HTML): Bắt buộc phải có** thẻ <input type="password"> trong mã nguồn HTML tĩnh.
    - **Dấu hiệu Mạnh Kết hợp (URL & HTML):**
        - Có <input type="password"> **VÀ** URL chứa từ khóa đăng nhập ("login", "signin", "verify", "account"...) **HOẶC** URL có cấu trúc rõ ràng giả mạo trang đăng nhập (thương hiệu + login/secure...).
    - **Dấu hiệu Hỗ trợ:** Form tĩnh chứa cả trường password và định danh; nút submit login; text/hình ảnh tĩnh mô phỏng trang login; URL đáng ngờ chung (TLD lạ, mới, nhiều subdomain...).
    - **Loại trừ:** Không có <input type="password"> tĩnh.
2. **Malware Distribution (MW):**
    - **Mục tiêu (Suy đoán):** Phân phối mã độc qua link/mã tĩnh. **(Độ tin cậy thấp nhất khi chỉ phân tích tĩnh).**

- **Dấu hiệu Chính yếu Tĩnh (HTML - Phải rất rõ ràng):** Thẻ <a> có href trỏ **trực tiếp, không mơ hồ** đến tệp có đuôi thực thi/script **cực kỳ nguy hiểm** (.exe, .apk, .scr, .bat, .msi...). *Link đến .zip, .rar, .js cần xem xét kỹ ngữ cảnh URL/tên file.*
- **Dấu hiệu Mạnh Kết hợp (URL & HTML):** Link tải trực tiếp tĩnh (như trên) **HOẶC** URL có đường dẫn chứa tên tệp thực thi rõ ràng (/update.exe) **VÀ/HOẶC** URL thuộc domain đã bị blacklist là phân phối mã độc (từ Bước 0).
- **Dấu hiệu Hỗ trợ (Yếu, Cần xác nhận thêm):** Mã <script> tĩnh cực kỳ đáng ngờ (obfuscation nặng); <iframe> tĩnh có src độc hại đã biết; text tĩnh yêu cầu tải/cập nhật; URL chứa từ khóa download/update.
- **Loại trừ:** Thiếu link tải trực tiếp tĩnh rõ ràng; URL không có dấu hiệu mạnh. **Luôn ưu tiên gắn cờ "Review" cho MW nếu chỉ dựa vào tĩnh.**

3. **Financial Scam (FIN):**
   - **Mục tiêu (Suy đoán):** Chiếm đoạt tiền/thông tin thanh toán qua yếu tố tĩnh.
   - **Dấu hiệu Chính yếu Tĩnh (HTML): Bắt buộc phải có ít nhất một:**
     - Trường <input> tĩnh với name/id/placeholder rõ ràng chỉ **CVV/CVC/CID**.
     - Văn bản tĩnh hiển thị **số tài khoản ngân hàng/địa chỉ ví crypto** chi tiết để nhận tiền.
     - Kết hợp tĩnh của **cảnh báo lỗi/virus đáng sợ + số điện thoại liên hệ tĩnh** (Tech Support Scam).
   - **Dấu hiệu Mạnh Kết hợp (URL & HTML):** Có dấu hiệu chính yếu tĩnh ở trên **HOẶC** URL chứa từ khóa tài chính mạnh **VÀ/HOẶC** URL rõ ràng giả mạo tên miền ngân hàng/cổng thanh toán.
   - **Dấu hiệu Hỗ trợ:** Form tĩnh yêu cầu chi tiết thẻ khác (số thẻ, ngày hết hạn); logo tĩnh cổng thanh toán; từ khóa thanh toán tĩnh; URL đáng ngờ chung.
   - **Loại trừ:** Thiếu các dấu hiệu chính yếu tĩnh mạnh về CVV/chuyển khoản/ví/tech support.
4. **PII Gathering (PII):**
   - **Mục tiêu (Suy đoán):** Thu thập đa dạng PII nhạy cảm qua form tĩnh.
   - **Dấu hiệu Chính yếu Tĩnh (HTML): Bắt buộc phải có:** Form tĩnh chứa **nhiều loại (>2-3 loại)** trường <input>/<select>/<textarea> tĩnh rõ ràng yêu cầu PII nhạy

cảm (SSN, DOB, Địa chỉ đầy đủ, Câu hỏi bảo mật...) **VÀ** form đó **không có <input type="password"> làm trọng tâm VÀ không có trường CVV tĩnh**.

- **Dấu hiệu Mạnh Kết hợp (URL & HTML):** Có form PII đa dạng tĩnh như trên **HOẶC** URL chứa từ khóa liên quan đến hồ sơ/xác minh/khảo sát **VÀ/HOẶC** URL gợi ý về các hoạt động này.
- **Dấu hiệu Hỗ trợ:** Văn bản tĩnh yêu cầu cập nhật/xác minh/đăng ký; URL đáng ngờ chung.
- **Loại trừ:** Form tĩnh chỉ yêu cầu 1-2 PII cơ bản; Thiếu sự đa dạng và nhạy cảm của PII yêu cầu tĩnh.

5. **Other Phishing (OTH):**
    - **Mục tiêu (Suy đoán):** Là phishing nhưng không thể phân loại vào 4 lớp trên dựa trên bằng chứng tĩnh + URL.
    - **Dấu hiệu Chính yếu (Lớp còn lại):** Có dấu hiệu phishing chung từ URL (rất đáng ngờ, blacklist chung) hoặc HTML tĩnh (giả mạo cơ bản) **NHƯNG thiếu bằng chứng tĩnh chính yếu mạnh** của CH, MW, FIN, PII sau khi áp dụng luồng ưu tiên. **Hoặc** được các nguồn bên ngoài xác định là phishing chung chung nhưng không rõ loại.
    - **Trường hợp điển hình:** URL đáng ngờ + HTML tĩnh trống/đơn giản/lỗi; Social engineering tĩnh; Form tĩnh chỉ lấy Email/SĐT; Chuyển hướng meta tĩnh; **Nhiều trang động phức tạp rơi vào đây**.
    - **Loại trừ:** Nếu phù hợp với CH, MW, FIN, PII theo quy tắc.

IV. Luồng Logic Ưu Tiên Xử Lý Chồng Chéo (Adjudication Flow - Tĩnh + URL + External Context):

(Áp dụng tuần tự. Dừng ngay khi gán được nhãn. Xem xét thông tin từ Bước 0 nếu có)

Bước 1: Kiểm tra FIN?

- **Câu hỏi:** Có dấu hiệu tĩnh chính yếu **mạnh và rõ ràng** của FIN (CVV tĩnh, Chuyển khoản/Ví tĩnh, Tech Support tĩnh)? Hoặc URL có dấu hiệu tài chính cực mạnh? Hoặc nguồn ngoài chỉ rõ là lừa đảo tài chính?
- **Quyết định:** Nếu **CÓ**, gán **FIN**. Dừng.

Bước 2: Kiểm tra MW?

- *(Nếu Bước 1 là KHÔNG)*

- **Câu hỏi:** Có dấu hiệu tĩnh chính yếu **cực kỳ mạnh và không mơ hồ** của MW (link .exe trực tiếp rõ ràng)? Hoặc URL/Domain bị nguồn ngoài uy tín đánh dấu là **phân phối mã độc**?
- **Quyết định:** Nếu **CÓ**, gán **MW** (ưu tiên gắn cờ "Review" nếu chỉ dựa vào tĩnh). Dừng.

Bước 3: Kiểm tra CH?

- *(Nếu Bước 1 & 2 là KHÔNG)*
- **Câu hỏi:** Có thẻ <input type="password"> trong HTML tĩnh? Hoặc URL có dấu hiệu giả mạo trang đăng nhập rất mạnh?
- **Quyết định:** Nếu **CÓ**, gán **CH**. Dừng.

Bước 4: Kiểm tra PII?

- *(Nếu Bước 1, 2 & 3 là KHÔNG)*
- **Câu hỏi:** Có form tĩnh chứa **nhiều loại PII nhạy cảm** (không password/CVV làm trọng tâm)? Hoặc URL có dấu hiệu mạnh về yêu cầu hồ sơ/xác minh?
- **Quyết định:** Nếu **CÓ** (chủ yếu dựa vào form tĩnh), gán **PII**. Dừng.

Bước 5: Gán Nhãn OTH?

- *(Nếu Bước 1-4 là KHÔNG)*
- **Câu hỏi:** Có dấu hiệu phishing chung từ URL (đáng ngờ, blacklist chung) hoặc HTML tĩnh, HOẶC nguồn ngoài xác nhận là phishing chung chung, nhưng không đủ bằng chứng tĩnh mạnh cho CH/MW/FIN/PII?
- **Quyết định:** Gán **OTH** (thường nên gắn cờ "Review"). Dừng.

Bước 6: Xem xét lại LEG?

- *(Nếu chưa gán nhãn nào từ CH-OTH)*
- **Câu hỏi:** Sau khi xem xét kỹ URL, HTML tĩnh và thông tin nguồn ngoài (nếu có), liệu có khả năng đây là trang hợp pháp không?
- **Quyết định:** Nếu **CÓ**, gán **LEG**. Nếu không chắc chắn => Gán **OTH** và gắn cờ "Review".

V. Hướng Dẫn Bổ Sung Cho Người Gán Nhãn:

- **Luôn Tham chiếu Quy tắc V5:** Giữ tài liệu này làm kim chỉ nam.
- **Xem xét Đồng thời URL, HTML tĩnh, và Thông tin Ngoài (nếu có):** Đưa ra quyết định dựa trên tổng thể bằng chứng.

- **Ưu tiên Bằng chứng Mạnh:** Chỉ gán CH/MW/FIN/PII nếu có dấu hiệu chính yếu tĩnh mạnh hoặc xác nhận mạnh từ nguồn ngoài.
- **Sử dụng "Review" Flag:** Đừng ngần ngại gắn cờ cho MW tĩnh, OTH, và bất kỳ trường hợp nào bạn không chắc chắn 100%.
- **Ghi Chú Lý Do:** Đặc biệt cho các trường hợp khó hoặc khi ghi đè gợi ý tự động.
- **Nhất Quán:** Áp dụng cùng một logic cho mọi mẫu.

Đây là phiên bản V5, cố gắng cân bằng giữa việc tận dụng tối đa thông tin tĩnh có sẵn (URL + HTML) và thừa nhận sự cần thiết của việc tham khảo ngữ cảnh bên ngoài, đồng thời duy trì một quy trình nhất quán và rõ ràng trong giới hạn không thực thi JavaScript. Hãy sử dụng nó một cách cẩn thận và luôn ghi nhớ những hạn chế đã nêu.