**Cách Hệ Thống "Kiểm Tin Thật" Hoạt Động và Chấm Điểm**

**I. Nguyên tắc chung:**

* **Điểm khởi tạo:** Mỗi tin tức bắt đầu với **100 điểm tin cậy**.
* **Cộng/Trừ điểm:** Hệ thống quét nội dung (hoặc URL) để tìm kiếm các từ khóa, cụm từ, cấu trúc, hoặc các dấu hiệu đặc trưng. Tùy thuộc vào việc phát hiện dấu hiệu tích cực hay tiêu cực, điểm sẽ được cộng hoặc trừ.
* **Phạm vi điểm:** Điểm cuối cùng được giới hạn trong khoảng **0% đến 100%**.
* **Giải thích:** Hệ thống cung cấp danh sách các lý do chi tiết (các yếu tố đã ảnh hưởng đến điểm số) để người dùng hiểu rõ hơn về kết quả.
* **Khuyến nghị:** Dựa trên điểm số cuối cùng, hệ thống đưa ra khuyến nghị về độ tin cậy của tin tức (Đáng tin cậy, Cần kiểm chứng, Độ tin cậy thấp).

**II. Cách chấm điểm cho từng loại đầu vào:**

**A. Đối với Đầu vào là URL (Link bài viết):**

Hệ thống phân tích tên miền, cấu trúc URL và các từ khóa trong URL.

1. **Uy tín tên miền (Domain Authority):**
   * **Cộng điểm (+5 điểm):** Nếu tên miền thuộc danh sách các nguồn tin uy tín đã biết (ví dụ: vnexpress.net, tuoitre.vn, thanhnien.vn, vietnamnet.vn, baochinhphu.vn, vtv.vn, v.v.).
   * **Trừ điểm (-40 điểm):** Nếu tên miền là các nền tảng blog cá nhân hoặc tên miền cấp cao nhất (TLD) ít phổ biến, thường liên quan đến spam (.blogspot.com, .wordpress.com, .xyz, .club, .info, .biz, .icu, .top).
   * **Trừ điểm (-30 điểm):** Nếu tên miền không nằm trong danh sách uy tín hoặc ít phổ biến, cần đánh giá cẩn trọng hơn.
2. **Từ khóa gây hiểu lầm/giật gân trong URL:**
   * **Trừ điểm (-35 điểm):** Nếu URL chứa các từ khóa như shock, sensational, unbelievable, secret, exposed, top10, cucnong, kinhhoang, khongthetin, batngo.
3. **Dấu hiệu Clickbait trong URL:**
   * **Trừ điểm (-45 điểm):** Nếu URL khớp với các mẫu clickbait như /tin-giat-gan/, /cuc-nong/, /khong-the-tin-noi/, /bi-mat/, /su-that/, /chan-dong/, /lan-truyen-chong-mat/, /gay-bao-mang-xa-hoi/.
4. **Độ dài URL:**
   * **Trừ điểm (-20 điểm):** Nếu URL quá dài (>150 ký tự) hoặc quá ngắn (<25 ký tự).
5. **URL không hợp lệ:**
   * **Giảm mạnh điểm (-95 điểm, về 5 điểm):** Nếu URL có định dạng không hợp lệ, không thể phân tích được.

**B. Đối với Đầu vào là Văn bản (Nội dung bài viết):**

Hệ thống phân tích từ khóa, cụm từ, cấu trúc câu, và các dấu hiệu ngữ nghĩa trong toàn bộ văn bản và đặc biệt là 200 ký tự đầu.

1. **Dấu hiệu nguồn tin chính thống, uy tín mạnh:**
   * **Cộng điểm (+15 điểm/dấu hiệu):** Nếu tìm thấy các cụm từ như bộ trưởng y tế, chính phủ, ủy ban thường vụ quốc hội, thông qua pháp lệnh, cơ quan chức năng, báo cáo của, theo thống kê, tổ chức y tế thế giới, tổng cục thống kê, thông tin chính thức từ. (Điểm cộng này có trọng số cao để ưu tiên tin chính thống).
2. **Các tuyên bố giật gân, phi lý, không thực tế:**
   * **Trừ điểm (-70 điểm):** Đối với các kết hợp phi lý cao như mỏ vàng tại chợ Bến Thành hoặc Sài Gòn/TP.HCM.
   * **Trừ điểm (-30 điểm/dấu hiệu):** Đối với các tuyên bố về tự chữa khỏi, thuốc thần, biệt dược, chữa bách bệnh, rồng bạch kim, khủng long sống, người ngoài hành tinh, v.v.
3. **Sử dụng con số cường điệu/không căn cứ:**
   * **Trừ điểm (-15 điểm/dấu hiệu):** Nếu phát hiện các tỷ lệ phần trăm lớn số lượng lớn.
4. **Tin tức về tương lai quá xa:**
   * **Trừ điểm (-40 điểm):** Nếu đề cập đến các sự kiện trong năm hiện tại + 1 hoặc năm hiện tại + 2.
5. **Từ khóa khẳng định/chính thống chung:**
   * **Cộng điểm (+2 điểm/dấu hiệu):** Nếu tìm thấy các từ khóa như chính thức, thông báo, nghiên cứu khoa học, chuyên gia hàng đầu, xác nhận từ cơ quan chức năng, đã được kiểm chứng, v.v.
6. **Từ khóa tiêu cực/giật gân (ảnh hưởng tiêu đề/đầu bài):**
   * **Trừ điểm (-20 điểm):** Nếu các từ khóa như chấn động, sốc toàn tập, kinh hoàng, rò rỉ thông tin mật, thuyết âm mưu, khẩn cấp, tin khẩn, cảnh báo khẩn cấp, tin nóng xuất hiện trong **200 ký tự đầu tiên**.
   * **Trừ điểm (-10 điểm):** Nếu các từ khóa này xuất hiện ở **phần còn lại của văn bản**. (Lưu ý: Một số từ thông thường như tin nóng trong ngữ cảnh tiêu đề báo chí sẽ không bị trừ điểm).
7. **Từ khóa thể hiện ý kiến cá nhân/thiếu khách quan:**
   * **Trừ điểm (-15 điểm/dấu hiệu):** Nếu tìm thấy tôi nghĩ rằng, theo quan điểm của tôi, có lẽ là, dường như là, tôi cảm thấy, quan điểm cá nhân, chưa chắc chắn, nghi ngờ, v.v.
8. **Đề cập mạng xã hội (nguồn tin chưa kiểm chứng):**
   * **Trừ điểm (-25 điểm/dấu hiệu):** Nếu đề cập đến tiktok, facebook, zalo, mạng xã hội, video lan truyền, chia sẻ trên mạng, cộng đồng mạng.
9. **Nguồn tin chung chung/không xác minh được:**
   * **Trừ điểm (-30 điểm/dấu hiệu):** Nếu tìm thấy một số người, nguồn tin riêng, nghe nói, tin từ nội bộ, chưa xác định, theo một nguồn tin, ai đó nói.
10. **Từ ngữ cảm xúc mạnh:**
    * **Trừ điểm (-10 điểm/dấu hiệu):** Nếu sử dụng nhiều phẫn nộ, đau lòng, tức giận, thương tâm, căm phẫn, bàng hoàng, xúc động mạnh.
11. **Độ dài văn bản:**
    * **Trừ điểm (-30 điểm):** Nếu văn bản quá ngắn (<150 ký tự).
    * **Trừ điểm (-10 điểm):** Nếu văn bản quá dài (>7000 ký tự).
12. **Dấu câu cường điệu/Viết hoa toàn bộ:**
    * **Trừ điểm (-10 điểm/dấu chấm than/hỏi):** Nếu sử dụng nhiều dấu chấm than ! hoặc dấu hỏi ?.
    * **Trừ điểm (-15 điểm/từ):** Nếu có nhiều từ/cụm từ viết hoa toàn bộ (trừ các từ viết tắt thông thường như TPHCM, COVID, WHO...).
13. **Lỗi chính tả/ngữ pháp phổ biến:**
    * **Trừ điểm (-15 điểm/lỗi):** Nếu phát hiện các lỗi chính tả hoặc ngữ pháp không chuẩn (chứnng minh, nguoif, ko, khong the, v.v.).
14. **Thiếu thông tin ngày tháng hoặc nguồn chính thức cụ thể (khi không có dấu hiệu chính thống mạnh):**
    * **Trừ điểm (-35 điểm):** Nếu không tìm thấy thông tin ngày tháng cụ thể và không có các dấu hiệu mạnh của nguồn chính thống được đề cập.
15. **Kêu gọi hành động/chia sẻ quá mức:**
    * **Trừ điểm (-20 điểm/dấu hiệu):** Nếu có các cụm từ như hãy chia sẻ ngay, hãy lan truyền, comment bên dưới, đừng bỏ lỡ, hãy hành động ngay.
16. **Tuyên bố về lợi ích cá nhân/làm giàu nhanh:**
    * **Trừ điểm (-30 điểm/dấu hiệu):** Nếu có các cụm từ như cách làm giàu nhanh, kiếm tiền dễ dàng, thoát nghèo cấp tốc, tỷ phú trong \d+ ngày, lợi nhuận khủng, đổi đời.
17. **Ngôn ngữ kích động, chia rẽ:**
    * **Trừ điểm (-15 điểm/dấu hiệu):** Nếu sử dụng kẻ thù, âm mưu, lừa dối, bóc trần, vạch trần, thế lực, đáng sợ.
18. **Sự tồn tại của thực thể/tổ chức được nhắc đến (đánh giá cụ thể):**
    * **Trừ điểm (-60 điểm):** Nếu đề cập rồng bạch kim (phi thực tế).
    * **Trừ điểm (-50 điểm):** Nếu đề cập hiệp hội hang động hoàng gia anh (tổ chức không tồn tại, sai tên).
    * **Trừ điểm (-30 điểm):** Nếu đề cập dr. emily roberts (tên chung chung, khó xác minh là chuyên gia nổi tiếng).
    * **Cộng điểm (+10 điểm):** Nếu đề cập sơn đoòng (địa danh có thật, tăng tính xác thực).
19. **Tuyên bố "độc nhất vô nhị", "lần đầu tiên" thiếu bằng chứng:**
    * **Trừ điểm (-25 điểm/dấu hiệu):** Nếu có các cụm từ như lần đầu tiên trên thế giới, duy nhất trên hành tinh, chưa từng thấy bao giờ, phát hiện lịch sử.
20. **Nêu đích danh "chính phủ" hoặc "cơ quan" nhưng thiếu thông tin chi tiết (và không có dấu hiệu chính thống mạnh khác):**
    * **Trừ điểm (-20 điểm/dấu hiệu):** Nếu các cụm từ như chính phủ đã ra quyết định, cơ quan chức năng đã khẳng định, bộ y tế đã công bố xuất hiện mà không kèm số công văn, ngày ban hành hoặc tên quyết định cụ thể, và không có các dấu hiệu chính thống mạnh khác.