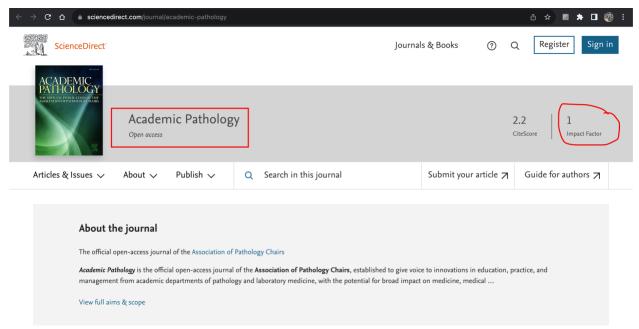
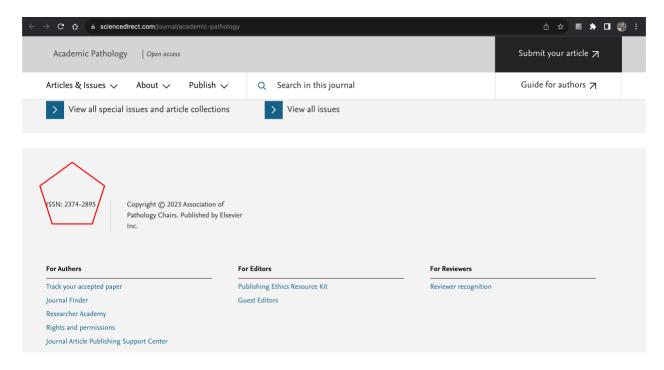
Nhiệm vụ: crawl dữ liệu impact factor (2 years) của các tạp chí của nhà xuất bản Elsevier

- 1. Craw toàn bộ danh mục tạp chí của SAGE từ (lưu ý: đang lọc theo tạp chí) https://www.sciencedirect.com/browse/journals-and-books?contentType=JL
- Crawl tên của mỗi tạp chí, ví dụ, tạp chí: https://www.sciencedirect.com/journal/academic-pathology có tên viết tắt là Academic Pathology, impact factor (2 years) và ISSN/e-ISSN của tạp chí từ link tạp chí, ví dụ, tạp chí ở trên



Ví dụ, như ở trên, ta có tên tạp chí là Academic Pathology, impact factor của tạp chí là: 1, và ở cuối trang, sẽ có mã ISSN hoặc/và e-ISSN



3. Lập thành một bảng dữ liệu (excel file) như bên dưới:

STT	Tên tạp chí	ISSN	eISSN	Impact Factor
1	Academic Pathology	2374-2895		1

Lưu ý: nếu trường dữ liệu nào không có, thì để trống. Sử dụng thư viện Python để crawl dữ liệu html sau đó xử lý (xem thêm bài tập đã làm trong môn Xử lý ngôn ngữ tự nhiên).