Article mis à jour le : 2024/11/25 22:28 / Imprimé le 2025/11/02 10:05

Data Scraping

Récupérer des données du web public pour whatever raison

wget -p -k http://francoisegamma.computersclub.org/

-p : télécharger tous les fichiers nécessaires (css, images, etc.) -k : convertir les liens pour que la page fonctionne hors ligne

Beaucoup d'autres possiblités avec wget

httrack http://www.httrack.com/page/2/fr/index.html existe en paquet debian webhttrack

Article extrait de : http://lesporteslogiques.net/wiki/ - WIKI Les Portes Logiques

Adresse: http://lesporteslogiques.net/wiki/ressource/data_scraping?rev=1732570125

Article mis à jour: 2024/11/25 22:28