

令和6年度 シラバス兼授業計画書

学科名	未来創造IT学科	科目名	Webスクレイピング		
授業種類	理論	履修区分	必修	履修時期	通年
授業時間	60単位時間	担当教員	伊藤 治		
授業内容	《授業概要》 本授業では、クローリング、スクレイピングに関する基礎知識と実践的知識の取得を目的とします。 具体的には、基本的なクローリングやAPIを活用したデータ収集、HTMLやXMLの解析からデータ取得後の分析などの処理までを学び、データの収集・解析、活用についての基礎知識の習得を目指します。				
	《学習の到達目標》 1 スクレイピングに関する基礎知識を理解している。 2 クローリングに関する基礎知識を理解している。				
授業計画					
前期			後期		
1	ガイダンス	16	実践とデータの活用①(データセットの取得と活用)		
2	Wgetによるクローリング	17	実践とデータの活用②(データの収集と活用)		
3	gihyo.jpのスクレイピング	18	実践とデータの活用③(Webページの自動操作)		
4	Pythonの基礎知識	19	実践とデータの活用④(ページのスクレイピング)		
5	Webページの取得	20	実践とデータの活用⑤(取得したデータの活用)		
6	Webページからのデータ取り出し	21	フレームワーク Scrapy①(概要、Spider作成・実行)		
7	Pythonによるスクレイピングの流れ	22	フレームワーク Scrapy②(データの処理)		
8	HTMLのスクレイピング	23	フレームワーク Scrapy③(Scrapyの設定・拡張)		
9	XMLのスクレイピング	24	フレームワーク Scrapy④(データの収集と活用)		
10	クローラーとURL	25	フレームワーク Scrapy⑤(画像の収集と活用)		
11	Pythonによるクローラーの作成	26	クローラー運用①(サーバでの実行)		
12	実用のためのメソッド①(クローラーの特性)	27	クローラー運用②(クローリングとスクレイピングの分離)		
13	実用のためのメソッド②(注意ポイント)	28	クローラー運用③(クローリングの高速化・非同期化)		
14	実用のためのメソッド③(設計ポイント)	29	クローラー運用④(クラウド活用)		
15	前期まとめ	30	後期まとめ		
前期試験			後期試験		
教科書参考書	教科書 Pythonクローリング&スクレイピング[増補改訂版] -データ収集・解析のための実践開発ガイド 加藤耕太 著 技術評論社				
成績評価	スクレイピング・クローリングに関する基礎知識について8割程度理解できることを到達目標とする。 【評価方法】 1.授業態度 20% 2.前期試験と後期試験 80%				
履修上の留意点	PCを使用した実習を含むため、以下の推奨スペックを満たすPCが必要です。お使いのPCが以下の推奨スペックを満たしていない場合、スムーズに学習を進める事ができない可能性があります。 要件) OS:Windows8 / Windows10 / Windows11 CPU:インテルcore i5以上、メモリ:8GB以上、ハードディスク:256GB以上				