



# DIS18: Green Deal Annotations

## 01: Introduction and Kick-off

Philipp Schaer, Technische Hochschule Köln, Cologne, Germany

Version: 2021-10-07



**in·tro·duc·tion** (in/trə duk/shən), *n.*  
troducting or the state of being introd  
personal presentation of one person to  
**3.** a preliminary part, as of a book, m  
or the like, leading up to the main pa  
tary treatise: *an introduction to botan*  
stance of inserting. **6.** something i  
1400; ME *introduccion* < L *intrōducti*  
*tiō*). See INTRODUCE, -TION]  
—**Syn.** **3.** INTRODUCTION, FOREWORD  
material given at the front of a book  
duce it to the reader. A FOREWORD i  
matter and is usually written by some  
author, often an authority on the sub

# Organization

## Regular appointments + plus additional time slots

- Virtual Tuesday 09:45
- TH: Wednesday 10-13h

## Rooms (virtual and local @TH)

- <https://th-koeln.zoom.us/j/81513492081>

## GitHub

- <https://github.com/irgroup-classrooms/dis18-2021>

## Mattermost

- [https://chat.iim.th-koeln.de/signup\\_user\\_complete/?id=x7hmzjc96ffqmez5pse39w9eur](https://chat.iim.th-koeln.de/signup_user_complete/?id=x7hmzjc96ffqmez5pse39w9eur)
- Use your TH mail account to create an account



# How to pass this project

## Hard requirements

- Do milestone presentations after (1st semester)
- Write a term paper at the end of the course (2nd semester)

## “Soft” requirement:

- Be an active part of the group
- Learn a lot new and interesting stuff about Python, NLP, and other DIS-related topics

# Manifesto for Agile Software Development

We are uncovering better ways of developing software by doing it and helping others do it.  
Through this work we have come to value:

**Individuals and interactions** over processes and tools  
**Working software** over comprehensive documentation  
**Customer collaboration** over contract negotiation  
**Responding to change** over following a plan

That is, while there is value in the items on the right, we value the items on the left more.

# A European Green Deal



“Climate change and environmental degradation are an existential threat to Europe and the world. To overcome these challenges, the European Green Deal will transform the EU into a modern, resource-efficient and competitive economy, ensuring:

- no net emissions of greenhouse gases by 2050
- economic growth decoupled from resource use
- no person and no place left behind

The European Green Deal is also our lifeline out of the COVID-19 pandemic. **One third of the 1.8 trillion euro** investments the NextGenerationEU Recovery Plan, and the EU’s seven-year budget will finance the European Green Deal.”

## A European Green Deal

Striving to be the first climate-neutral continent

≡ MENU

🔍 QUICK SEARCH



📘 Search tips

Need more search options? Use the [Advanced search](#)

EUROPA > EUR-Lex home > EUR-Lex - 52021PC0555 - EN

Document 52021PC0555

? 🖨️ ➦ Share

↕ Expand all ↩ Collapse all

Text

Document information

Procedure

Internal procedure

Back to EUR-Lex homepage

📄 Save to My items

🔖 Permanent link

📄 Download notice

📡 Follow this document

▼ Title and reference

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulation (EU) 2018/842 on binding annual greenhouse gas emission reductions by Member States from 2021 to 2030 contributing to climate action to meet commitments under the Paris Agreement  
COM/2021/555 final

▼ Languages and formats available

	BG	ES	CS	DA	DE	ET	EL	EN	FR	GA	HR	IT	LV	LT	HU	MT	NL	PL	PT	RO	SK	SL	FI	SV
HTML																								
DOC																								
PDF																								

▼ Multilingual display

English (en)



Please choose



Please choose



Display

## ANNEX

### MEMBER STATE GREENHOUSE GAS EMISSION REDUCTIONS PURSUANT TO ARTICLE 4(1)

	Member State greenhouse gas emission reductions in 2030 in relation to their 2005 levels determined in accordance with Article 4(3)	
		Column 2
Belgium	Project DIS18 for the win!	-47%
Bulgaria		-10%
Czechia		-26%
Denmark		-50%
Germany		-50%
Estonia		-24%



# Project Goals and Work Packages

The following work packages are part of the project

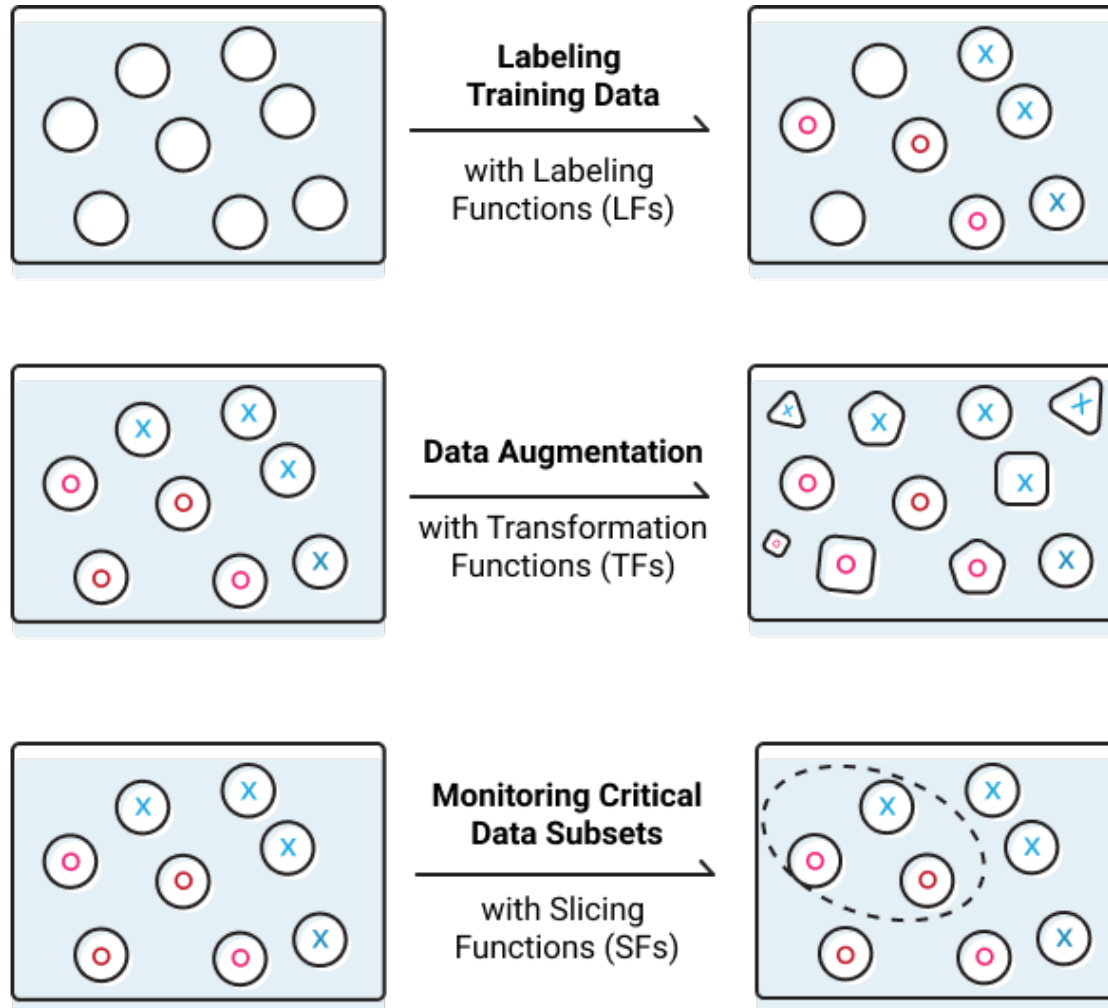
- **Literature** research and **concept** development
- Preprocessing of the documents
- Development and extension of an annotation tool
- Implementation of specific mechanism to support the annotation process
- Evaluation of the systems and processes



# Snorkel



# snorkel



# Label Studio

Person<sup>[1]</sup>
 Organization<sup>[2]</sup>
 Fact<sup>[3]</sup>
 Money<sup>[4]</sup>
 Date<sup>[5]</sup>
 Time<sup>[6]</sup>
 Ordinal<sup>[7]</sup>
 Percent<sup>[8]</sup>
 Product<sup>[9]</sup>
 Language<sup>[10]</sup>
 Location<sup>[11]</sup>

possums[Person] are usually solitary[Fact] and nomadic, staying in one area as long as food and water are easily available. Some families will group together in ready-made burrows or even under houses. Though they will temporarily occupy abandoned burrows, they do not dig or put much[Ordinal] effort into building their own. As nocturnal animals, they favor dark, secure areas. These[Date] areas may be below ground or above.[36][37] Juvenile opossum hissing defensively When threatened or harmed, they will 'play possum', mimicking the appearance and smell of a sick or dead animal[Product]. This physiological response is involuntary (like fainting), rather than a conscious act. In the case of baby opossums, however, the brain does not always react this way at the appropriate moment, and therefore they often fail to play dead when threatened[Language]. When an opossum is playing possum, the animal's lips are drawn back, the teeth are bared[Person], saliva foams around the mouth, the eyes close or half-close, and a foul-smelling fluid is secreted from the anal glands. The stiff, curled form can be prodded, turned over, and even carried away without reaction. The animal will typically regain consciousness after a period of a few minutes to four hours, a process[Time] that begins [Location] with a slight twitching of the ears.[38] Some species of opossums have prehensile tails, although dangling by the tail is more common among juveniles. An opossum [Product] may also use its tail[Fact] as a brace and a fifth limb when climbing. The tail is occasionally used as a grip to carry bunches of leaves or bedding materials to the nest.[Date] [39] A mother will sometimes carry her young upon her back, where they will cling tightly even when she is climbing or running. Threatened opossums (especially males) will growl deeply, raising their pitch as the threat becomes more urgent. Males make a clicking 'smack' noise out of the side of their mouths as they wander in search of a mate, and females will sometimes repeat the sound in return. When separated or distressed, baby opossums will make a sneezing noise to signal their mother. The mother in return makes the clicking sound and waits for the baby to find her. If threatened, the baby will open its mouth and quietly hiss until the threat is gone. Diet Opossums eat dead[Time] animals, insects, rodents and birds. They also feed on eggs, frogs, plants, fruits and grain. One source notes [Location] their need for high amounts of calcium. [Fact] [citation needed] Thus, possums eat the skeletal remains of rodents and roadkill animals. Opossums also eat dog food, cat food and human food waste. Many large opossums (Didelphini) are immune to the venom of rattlesnakes and pit vipers (Crotalinae) and [Person] regularly prey upon these snakes[Fact]. [40] This adaptation seems to be unique to the Didelphini, as their closest relative the brown four-eyed opossum is not immune to snake venom. [41] Similar adaptations are seen in other small predatory mammals such as mongooses and hedgehogs. Didelphin opossums and crotaline vipers have been suggested to be in an evolutionary arms race. Some authors have suggested that this adaptation originally arose as a defense mechanism, allowing a rare reversal of an evolutionary arms race where the former prey has become the predator. [42] whereas others have suggested it arose as a predatory adaptation given that it also occurs in other predatory mammals and does not occur [Money] in opossums that do not regularly eat other vertebrates. [9] The fer-de-lance, one of the most venomous snakes in the New World, may have developed its highly potent venom as a means to prey on or a defense mechanism against large opossums. [42] Opo[Fact]ssums are also notable for their ability to clean themselves of ticks, which they then eat. Some estimates suggest they can eliminate up to 5,000 ticks in a season. [43] Opossums are found in North, Central, and South America. The Virginia opossum lives in regions as far north as Canada and as far south as Central America, while other types of opossums only inhabit countries south of the United States. [44] The Virginia opossum can often be found in wooded areas, though its habitat may vary widely. [45] Opossums have been moving north in recent years. [46] Hunting and foodways The Virginia opossum was once widely hunted and consumed in the United States. [47] [48] [49] [50] Opossum farms have been operated in the United States in the past. [51] [52] [53] Sweet potatoes were eaten together with the possum in America's southern area. [54] [55] South Carolina cuisine includes opossum. [56] and President Jimmy Carter hunted opossums [57] [58] in addition to other small game. [59] [60] Raccoon, opossum, partridges, prairie hen, and frogs were among the fare Mark Twain recorded as part of American cuisine. [61] [62] [63] [64] [65] [66] [67] [68] [69] [70] [71] A 'possum joke sign In Dominica, Grenada, Trinidad, Saint Lucia and Saint Vincent and the Grenadines the common opossum or maniocou is popular and can only be hunted during certain times of the year owing to overhunting. [72] The meat is traditionally prepared by smoking, then stewing. It is light and fine-grained, but the musk glands must be removed as part of preparation. The meat can be used in place of rabbit and chicken in recipes. Historically, hunters in the Caribbean would place a barrel with fresh or rotten fruit to attract opossums that would feed on the fruit or insects. In northern/central Mexico, opossums are known. Their tails are eaten as a folk remedy to improve fertility. [73] In the Yucatán peninsula they are known in the Yucatec Mayan language as [74] and they are not considered part of the regular diet by Mayan people, but still considered edible in times of famine. Opossum oil (possum grease) is high in essential fatty acids and has been used as a chest rub and a carrier for arthritis remedies given as topical salves. [75] [76] [77] Opossum pelts have long been part of the fur trade. Classification The range of Didelphis virginiana across North America in green Classification based on Voss and Jansa (2009) [78] Tate's fat-tailed mouse opossum (Thylamys tatei) Dwarf fat-tailed mouse opossum (Thylamys velutinus) Buff-bellied fat-tailed mouse opossum (Thylamys venustus) †Thylamys zettii [90] †Genus Zygolestes †Zygolestes tatei. The tail is occasionally used as a grip to carry bunches of leaves or bedding materials to the nest. [39] A mother will sometimes carry her young upon her back, where they will cling tightly even when she is climbing or running. Threatened opossums (especially males) will



# Provisional Schedule for Winter Term

Date	Type	Topics
12.10.21	Zoom	Introduction and kick-off, organization paper distribution
19.10.21	Zoom	Introduction to Snorkel and Label Studio
26.10.21	Self-organized	Paper reading / Work on Snorkel and Label Studio tutorials
03.11.21	Workshop@TH	Student talks on background
09.11.21	Self-organized	Work on Snorkel and Label Studio tutorials
16.11.21	Zoom	Introduction to the data sets
23.11.21		Project week – no meeting
01.12.21	Workshop@TH	Workshop: Usecase definitions and further planning
07.12.21	Zoom	Work on prototype
14.12.21	Zoom	Work on prototype
21.12.21	Zoom	Work on prototype
28.12.21		Christmas holidays
04.01.22	Zoom	Work on prototype
12.01.22	Workshop@TH	Student talks – Milestone presentation

# NLP in a Nutshell

- “**Natural language processing (NLP)** is a subfield of linguistics, computer science, and artificial intelligence concerned with [...] how to program computers to **process and analyze large amounts** of natural language data.”
- “The goal is a computer capable of **"understanding" the contents** of documents [...]. The technology can then accurately extract information and insights contained in the documents as well as categorize and organize the documents themselves.”

## Speech and Language Processing

An Introduction to Natural Language Processing,  
Computational Linguistics, and Speech Recognition

Third Edition draft

Daniel Jurafsky  
*Stanford University*

James H. Martin  
*University of Colorado at Boulder*

*Copyright ©2020. All rights reserved.*

Draft of December 30, 2020. Comments and typos welcome!

# Computational Journalism

## **What do Journalists do with Documents? Field Notes for Natural Language Processing Researchers**

Jonathan Stray  
Columbia Journalism School  
jms2361@columbia.edu

- Robust import
- Robust analysis
- Search, not exploration
- Quantitative summaries
- Interactive methods
- Clarity and Accuracy

Project JoIE: <https://ir.web.th-koeln.de/projects/joie/>



# NLP in a Nutshell – Your next job

Group 1: **Regular Expressions, Text Normalization, Edit Distance** (2.1-2.4 auf S. 2-22)

- Puddu, Blum, Kubon, Bateva, Wolf

Group 2: **Sequence Labeling for Parts of Speech and Named Entities** (8.1 - 8.3 auf S. 153-159)

- Nebatz, Jennert

Group 3: **Question Answering** (23.1-23.3.1 auf S. 470-486)

- Pycha, Hilbrecht, Schmer, Novak

Group 4: **Information Extraction** (17.1-17.4 auf S. 339-357)

- Tanriver; Celik; Alarslan; Romer

# Presentation Guidelines

- ALL of you are expected to read ALL chapters!
  - You and your designated group should present your chapter to the rest of the group, who read the chapter before!
  - Don't start by 0, but dive right in.
  - We are laying a foundation for the rest of the semester here!
- 15-20 minutes per team + 15 minutes discussion and questions
  - Push the slides (in PDF format) to our common GitHub repo a night before the workshop.

# Take-away message and next steps

- Join our GitHub project:  
<https://github.com/irgroup-classrooms/dis18-2021>
- Read, discuss and prepare your group's chapter!
- We see each other again **in one week for an introduction to Snorkel and Label Studio!**