

# Zgłoszenie tematu pracy dyplomowej :: **STUDIA II STOPNIA** ::

na rok akademicki 2020/21

<b>Promotor:</b>	<b>dr hab. Jozef Kapusta, prof. UP</b>
Temat pracy magisterskiej (j. polski, j.angielski):	Application of Morphological Analysis for Speakers Identification <i>Zastosowanie analizy morfologicznej do identyfikacji mówców</i>
Zakres pracy i oczekiwane rezultaty praktyczne:	<p>Natural Language Processing is a sub-field of Artificial Intelligence that is focused on enabling computers to understand and process human languages, to get computers closer to a human-level understanding of language.</p> <p>The aim of the practical part is to create classifier of speakers. There are many full-text speeches datasets (white house speeches, Pope Francis' speeches, etc.). Advanced features of speeches will be obtained from the morphological analysis (part of speech tags). The classifier will work based on a typical morphological characteristic of speakers. The student will be using known libraries (NLTK, UDPipe, Polyglot, TreeTagger etc.) and some selected solution for machine learning.</p>
Aspekt naukowy, problemowy, innowacyjny pracy:	Selection and implementation of natural language processing methods, implementation selected machine learning methods.
Oprogramowanie, język programowania, środowisko systemowe:	Jupyter Notebook environment (Python)
Środowisko uruchomieniowe	Windows or Linux
Dodatkowe wymagania i uwagi:	english language
Literatura:	<ol style="list-style-type: none"> <li>1. Steven Bird, Ewan Klein, and Edward Loper: Natural Language Processing with Python - Analyzing Text with the Natural Language Toolkit. O'Reilly Media, 2009.</li> <li>2. Benjamin Bengfort, Tony Ojeda, Rebecca Bilbro: Applied Text Analysis with Python: Enabling Language - Aware Data Products with Machine Learning, O'Reilly Media, 2018, 332 p.</li> <li>3. UDPipe, online: <a href="http://lindat.mff.cuni.cz/services/udpipe/info.php">http://lindat.mff.cuni.cz/services/udpipe/info.php</a></li> <li>4. Natural Language Toolkit, online: <a href="https://www.nltk.org/">https://www.nltk.org/</a></li> <li>5. scikit-learn: Machine Learning in Python, online: <a href="https://scikit-learn.org/stable/">https://scikit-learn.org/stable/</a></li> </ol>