

Generování textu v přirozeném jazyce

Jiří Hana¹

Ústav formální a aplikované lingvistiky MFF UK, Malostranské nám. 2/25, 118 00 Praha 1

Jiri.Hana@mff.cuni.cz,

WWW home page: <http://ufal.mff.cuni.cz/hana/>

Abstract: Přednáška představí problematiku generování textu na příkladu sportovních zpráv. Jde o společný projekt nakladatelství Economia a firmy Geneea. Vyvíjený systém vytváří novinové články popisující jednotlivá sportovní utkání. Články jsou generovány na základě strukturovaných dat o událostech na hřišti a dat v databázích, jakou je například znalostní báze Wikidata. V této fázi projektu jde pouze o fotbalové zápasy, ale v blízké budoucnosti bude systém rozšířen na další sporty a výhledově také na zprávy o počasí a zprávy z burzy.

Jiří Hana má zkušenosti jak z firemního, tak z univerzitního prostředí. Vystudoval informatiku na Matematicko-fyzikální fakultě UK a lingvistiku na Ohio State University. V USA strávil téměř 10 let. Mezi jeho zájmy patří morfologie, příklonky a jazyk cizinců. Je vý-



zkumným pracovníkem na Ústavu formální a aplikované lingvistiky MFF UK a spoluzakladatelem společnosti Geneea Analytics, startupu vyvíjejícího systémy pro analýzu a generování textu.