

Československé sdružení uživatelů \TeX u

$\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$

si vás dovoluje pozvat na

přednášku a valné shromáždění

konané 12. října 2019 od 13.00 hodin (MENDELU, Brno)

s tímto programem:

1. Přednáška **Petra a Ondřeje Sojky**

Překvapivá efektivita generování vzorů

Začátek ve **13.00** hodin

2. Přestávka 15 min. (14.15–14.30, prezenze na valné shromáždění, je možno využít k řešení administrativy $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$ u, placení členských příspěvků apod.).
3. Valné shromáždění $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$ u – v **14.30** hodin.
 - a) Zahájení valného shromáždění, schválení programu.
 - b) Schválení mandátové a volební komise.
 - c) Volba předsedy zasedání.
 - d) Prezenční volba nového výboru a revizorů sdružení. Kandidátní listina bude zveřejněna na webu sdružení a je stále otevřena. Uzávěrka kandidátní listiny je 1. 10. 2019 ve 19 hodin, elektronická volba začne 7. 10. 2019 ve 10 hodin a skončí 11. 10. 2019 v 10 hodin.
 - e) Zpráva o činnosti.
 - f) Zpráva o hospodaření.
 - g) Zprávy revizorů sdružení.
 - h) Informace o Zpravodaji $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$ u.
 - i) Informace o stavu projektů podporovaných finančně $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$ em.
 - j) Rámcový plán činnosti a rozpočet pro rok 2020.
 - k) Stanovení členských příspěvků na rok 2020.
 - l) Vyhlášení výsledků voleb do nového výboru $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$ u a revizorů.
 - m) Různé, diskuse.
 - n) Závěr (cca v 15.30 hodin).

Valné shromáždění i přednáška se bude konat v budově Provozně ekonomické fakulty Mendelovy univerzity, Zemědělská 1, 613 00 Brno v **sobotu 12. 10. 2019**, v posluchárně s označením **Q02** v přízemí budovy Q.

Kromě osobního setkání s \TeX ovými přáteli si budete moci vyslechnout přednášku o užití \TeX ových technologií v praxi či zaplatit členské příspěvky sdružení. Noví členové $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$ u si budou moci osobně vyzvednout DVD \TeX Live 2019.

Valné shromáždění je určeno členům sdružení $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}$, nicméně je to vhodná příležitost, jak se stát členem sdružení. Informace o sdružení, zaměřeném na podporu kvalitní typografie v ČR a SR pomocí sázecího systému \TeX , je možno najít na adrese <http://www.cstug.cz/> nebo si je vyžádat na adrese cstug@cstug.cz.

Abstrakt přednášky

Petr Sojka a Ondřej Sojka: Překvapivá efektivita generování vzorů

Jazyky se neustále vyvíjejí, stejně jako jejich slova a pravidla dělení slov. Účinnost a užitečnost algoritmu dělení slov \TeX u byla prokázána jeho použitím v téměř všech dnes používaných sázečích systémech. Současné vzory pro české dělení slov byly vytvořeny v roce 1995 a žádná databáze slov s vyznačeným dělením slov pro jejich generování dosud nebyla volně dostupná.

Vyvinuli jsme novou databázi českých slov a pomocí programu `patgen` jsme efektivně vytvořili nové české vzory dělení slov a vyhodnotili jejich generalizační vlastnosti. Dosáhli jsme *plného* pokrytí vstupního souboru tří miliónů slov a vyvinuli jsme validační postup nových vzorů pro češtinu na základě testovací databáze 105 000 slov schválených lingvisty Akademie věd ČR.

Naše případová studie generování vzorů ilustruje praktické řešení široce používaného slovníkového problému. Prokázali jsme univerzálnost, efektivitu a rozšiřitelnost Liangova přístupu k segmentaci slov vyvinutého pro \TeX . Překvapivá efektivita technologie vzorů vedla k aplikacím, které jsou a budou používány i téměř 40 let po jejich návrhu.

První autor je předsedou $\mathcal{C}\mathcal{S}\mathcal{T}\mathcal{U}\mathcal{G}\mathcal{u}$ a docentem Fakulty informatiky Masarykovy univerzity, kde podporuje digitální typografii sázečím systémem \TeX od začátku 90. let minulého století. Druhý autor je studentem Gymnázia Matyáše Lercha v Brně.