

WORKSHOP

21/10/2019

9:30 – 13:30

AKC, Husova 4a, Praha 1

VYUŽITÍ DAT WEBOVÝCH ARCHIVŮ: Možnosti a limity

Workshop se bude věnovat problematice českých webových zdrojů, jenž jsou dlouhodobě ukládány a zpracovávány prostřednictvím Webarchivu Národní knihovny ČR. Cílem workshopu je představit problematiku budování, správy, ukládání a zpřístupňování dat webového archivu v širší perspektivě. Bude reflektován mezinárodní kontext budování a správy webových archivů, představen Webarchiv NK ČR, a diskutovány právní limity využití dat webových archivů. Workshop se rovněž zaměří na webové archivy jako na nový zdroj dat pro sociální a humanitní vědy. Na workshopu budou představeny praktické ukázky a zkušenosti s prací s tímto typem dat.

Workshop je pořádán v rámci aktivit výzkumného projektu „Vývoj centralizovaného rozhraní pro vytěžování velkých dat z webových archivů“ (ev. číslo DG18P020VV016) financovaného prostřednictvím dotačního programu NAKI II. Na realizaci projektu se společně podílejí odborníci z Národní knihovny ČR, Sociologického ústavu AV ČR a Katedry kybernetiky Západočeské univerzity v Plzni.

Registrace nutná předem na e-mailu: asistentka.cvvm@soc.cas.cz.

PROGRAM

Využití dat webových archivů: možnosti a limity

9:30 – 9:40

Přivítání účastníků workshopu

9:40 – 10:00

MARIE HAŠKOVCOVÁ (Národní knihovna ČR)

Klíčové aspekty budování a správy webových archivů v celosvětové perspektivě

10:00 – 10:20

JAROSLAV KVASNICA (Národní knihovna ČR)

Webarchiv NK ČR – stávající stav, typologie dat a použité technologie

10:20 – 10:40

TOMÁŠ FOLTÝN (Národní knihovna ČR)

Představení projektu „Vývoj centralizovaného rozhraní pro vytěžování velkých dat z webových archivů“

10:40 – 11:00

LUCIE STRAKOVÁ (Masarykova univerzita v Brně)

Právní aspekty a limity spolupráce v rámci zpracování dat webového archivu NK ČR

11:00 – 11:30

Přestávka na kávu, občerstvení

11:30 – 11:50

PAULÍNA TABERY, MATOUŠ PILNÁČEK (Sociologický ústav AV ČR)

Využití dat webových archivů v sociálních vědách

11:50 – 12:10

PAVEL IRCING, ZDENKO VOZÁR (Katedra kybernetiky, Západočeská univerzita v Plzni; Národní knihovna ČR)

Tvorba intermediary formátu a příklady využití moderních technologií (HADOOP, SPARK, jusTEXT atd.) pro práci s obsahem webového archivu

12:10 – 12:30

MARTA ŠKUBALOVÁ (Ústav pro českou literaturu AV ČR)

Prezentace dosavadních zkušeností ohledně práce s webovým archivem

12:30 – 13:20

ANDREA PROKOPOVÁ (Národní knihovna ČR)

WARC hands on workshop - praktická ukázka práce s WARC soubory

13:20 – 13:30

Ukončení workshopu



Národní knihovna
České republiky
National Library
of the Czech Republic



Sociologický ústav
Akademie věd ČR



ZÁPADOČESKÁ
UNIVERZITA
V PLZNI