

Chceme, aby data získaná z veřejných peněz byla veřejná a dostupná zdarma, popisuje jeden z cílů projektu Lindat/Clariah-CZ

dr. Jan Sládek

Filozofická fakulta UK se ve spolupráci s dalšími českými i mezinárodními institucemi podílí na projektu, jehož cílem je otevřeným způsobem zpřístupnit digitalizované datové zdroje v oblasti jazykových technologií, umění a humanitních věd široké vědecké komunitě i studentům. Projekt představuje dr. Jan Sládek (JS) z Katedry sociologie FF UK a Ondřej Fúsik (OF) z Ústavu anglického jazyka a didaktiky FF UK.

Projekt Lindat/Clariah-CZ (Lindat) klade důraz na digitální a interdisciplinární metody zpracování, včetně moderních metod strojového učení a umělé inteligence. Součástí aktivit skupiny je i analýza právních aspektů použití zdrojů z oblasti digitálních humanitních věd vzhledem k možným autorskoprávním omezením a minimalizace jejich dopadů na vědeckou a výzkumnou práci. Rovněž nabízí „know-how“, softwarové nástroje pro zpracování jazykových a jiných digitálních zdrojů, ale také vývoj zejména jazykových technologií pro potřeby průmyslu a služeb s využitím v nových kulturních a kreativních oborech průmyslu.



Vlevo dr. Jan Sládek, vpravo Ondřej Fúsik

Lindat je digitální výzkumná infrastruktura pro jazykové technologie, umění, sociální a humanitní vědy. Jakým způsobem se tyto oblasti v projektu propojují a co se od tohoto mezioborového spojení očekává?

JS: Mezioborové propojování je dnes standard. My se jej snažíme podpořit tím, že pro taková spojení poskytujeme datovou infrastrukturu. Konkrétní využití je pak na badatelích. Naší vizí je být v tomto propojování leaderem. V tom cítíme velkou podporu vedení FF UK. Také bych rád zmínil kolegu dr. Ondřeje Tichého z Ústavu anglického jazyka a didaktiky FF UK, který se rovněž podepsal na tom, že projekt na FF UK je.

OF: Když se podívám na svůj obor medievistiky, dnešní filolog specializující se na středověk je neoddělitelně spjat s celou řadou databází, digitalizovaných katalogů a rukopisů. Jeho výzkumná činnost se tak stává převážně digitálním podnikem. Je skvělé, že k tomu může projekt výrazně přispět. Díky novým technologiím pomáháme výzkumníkům nejen archivovat své objevy anebo je vytvářet, ale také jim nabízíme nástroje pro atraktivní prezentaci jejich práce, ať už formou digitálních

výstav, interaktivních prezentací, nebo dalších moderních formátů. Nejdůležitějším nástrojem je však možnost uložení digitálních výsledků do repozitáře, díky čemuž je možné je uchovat na dlouhou dobu.

Jak se do projektu zapojují akademici z FF UK? Čím konkrétně se zabýváte vy?

JS: Spolu se začátkem první fáze projektu jsme oslovili celou fakultu a přihlásilo se nám několik týmu, avšak s blížící se novou fází projektu na palubě rádi přivítáme i další akademiky. Pokud jde o mne osobně, pracuji na naplnění toho, že vedle již fungujících archivů sociálněvědních dat bude mít FF UK také svůj – a ten nabídne světu. Přispějeme tak k myšlence „open access“, tedy otevřenému přístupu k vědě. Chceme, aby data získaná z veřejných peněz byla veřejná a dostupná ideálně zdarma. Druhým, lehce osobním cílem je, aby studenti měli k dispozici solidní datové základny a nemuseli už vypracovávat diplomové práce založené na nepříliš dobře řízených anketách na sociálních sítích.

OF

:

Já

se

v

pr

oj

ek

tu Logo platformy Medieval Studies in the Digital World

za

mě

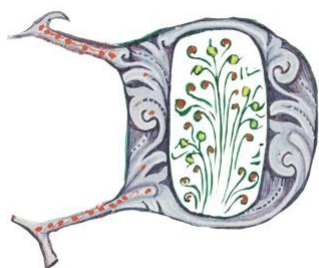
řu

ji

na

li

ng



vi
st
ic
ké
a
me
di
ev
is
ti
ck
é
as
pe
kt
y.
Mý
m
hl
av
ní
m
cí
le
m
je
zp
ří
st
up
ni
t
in
fo
rm
ac
e

o
st
ře
do
vě
ku
co
ne
jš
ir
ší
ve
ře
jn
os
ti
,
od
od
bo
rn
ík
ů
až
po
la
ik
y.
Js
em
rá
d,
že
pr
áv
ě
dí

ky
pr
oj
ek
tu
Li
nd
at
a
ve
sp
ol
up
rá
ci
s
Bi
la
te
rá
ln
í
in
ic
ia
ti
vo
u
Fo
nd
ů
EH
P
a
No
rs
ka

vz
ni
kl
a
pl
at
fo
rm
a
„M
ed
ie
va
l
St
ud
ie
s
in
th
e
Di
gi
ta
l
Wo
rl
d“
,
dí
ky
ní
ž
se
mn
ě
i

os
ta
tn
ím
da
ří
te
nt
o
cí
l
pl
ni
t.
Na
té
to
pl
at
fo
rm
ě
se
se
tk
áv
aj
í
rů
zn
í
vý
zk
um
ní
ci
-

od
pr
of
es
or
ů
a
se
ni
or
ní
ch
vý
zk
um
ní
ků
př
es
po
st
do
ky
až
po
st
ud
en
ty
vš
ec
h
st
up
ňů
. St

ud
en
ti
zd
e
ma
jí
př
íl
ež
it
os
t
uč
it
se
od
zk
uš
en
ěj
ší
ch
ko
le
gů
,
pr
ak
ti
ck
y
se
za
po
ji
t

do
vý
zk
um
ných
ch
ot
áz
ek
a
ro
zv
íj
et
sv
é
do
ve
dn
os
ti
ja
k
v
ob
la
st
i
vý
zk
um
u,
ta
k
př
i
pr

ez
en
ta
ci
sv
ýc
h
po
zn
at
kŭ
.
Li
nd
at
má
te
dy
i
si
ln
ou
di
da
kt
ic
ko
u
sl
ož
ku
a
na
po
má
há
fo

rm
ov
at
no
vo
u
ge
ne
ra
ci
vě
dc
ů.
Js
em
ži
vý
m
dů
ka
ze
m
je
ho
př
ín
os
u
–
dí
ky
Li
nd
at
u
js
em

si
ja
ko
do
kt
or
an
d
os
vo
ji
l
mn
oh
o
do
ve
dn
os
tí
a
tě
ší
m
se
na
da
lš
í
sp
ol
up
rá
ci
v
rá
mc

i
to
ho
to
pr
oj
ek
tu
,
pr
ot
ož
e
si
uv
ěd
om
uj
i,
že
se
má
m
st
ál
e
co
uč
it
.

Filozofická fakulta UK se dlouhodobě hlásí k pronikání myšlenky otevřené vědy, včetně otevřených, sdílených dat, do praxe humanitního a sociovědního výzkumu. Zapojení FF UK do projektu velké výzkumné infrastruktury Lindat/Clariah-CZ je jedním z praktických projevů obecného přesvědčení naší instituce, že věda má být dostupná a sdílená. Věřím, že

vynikající výsledky projektu, ať už z oblasti medievistiky nebo databází soudobých řečových projevů, si další uživatele z řad studujících či badatelek a badatelů najdou.

Děkanka FF UK dr. Eva Lehečková

Lindat vyvinul různorodé lingvistické nástroje a aplikace. Může je využívat i širší veřejnost?

OF: Projekt vytvořil, nebo pomohl vytvořit mnohé aplikace, takže si dovolím vybrat několik, které jsme vyvinuli na FF UK. Nesmíme totiž zapomínat, že projekt Lindat sdružuje mnohá vědecká pracoviště napříč Českou republikou a svým zapojením do dalších projektů je součástí vědecké komunity napříč celou Evropou.

Vyvinuli jsme například zajímavou aplikaci vocabula.lat sdružující různé slovníky latiny, které umožňují hledat nejen jednotlivá slova, ale i celé věty nebo kusy textu. Díky integrované umělé inteligenci umí v pokročilé verzi věty dokonce velmi slušně překládat. Aplikace je volně přístupná každému zájemci o latinu.

Zů
st
an
em
e-
li
u
sl
ov
ní
ků
,
pr
oj



ek
t
Li
nd
at
ta
ké
po
má
há
s
re
al
iz
ac
í
sl
ov
ní
ku
st
ar
oa
ng
li
čt
in
y
Bo
sw
or
th
To
ll
er
's
An

gl
o-
Sa
xo
n
Di
ct
io
na
ry
On
li
ne
,
kt
er
ý
mů
že
vy
už
ív
at
ši
ro
ká
ve
ře
jn
os
t
po
ce
lé
m
sv
ět

ě
se
zá
jm
em
o
ra
no
u
fá
zi
an
gl
ič
ti
ny
. Ta
ké
se
po
dí
lí
me
na
vý
vo
ji
ně
ko
li
ka
ko
rp
us
ů
rů

zn
ýc
h
ja
zy
ků
,
ať
už
hi
st
or
ic
ký
ch
ne
bo
dě
ts
ký
ch
a
so
uč
as
ný
ch
.

Pro širokou veřejnost budou lákavé dvě aplikace fungující na podobném principu – Lipnická bible a Křížův školský sborník A 4. Tyto dvě aplikace umožňují digitálně listovat středověkými rukopisy ve vysokém rozlišení, jejichž jednotlivé části jsou anotovány kolegy medievisty. Lipnická bible vznikla jako součást výstavy o tomto rukopisu a je doprovázena fyzickou knihou *Lipnická Bible – Štít víry v neklidných časech pozdního středověku*, která získala cenu Magnesia Litera za

nakladatelský čin v roce 2022. Aplikace rukopisu A 4 doprovází několik knih a výstavu „Rukopisy Kříže z Telče 1434–1504“.

Jakým způsobem se právě tato výstava, která byla uvedena v červnu v Třeboni, vztahuje k projektu Lindat?

OF: V rámci nejen naší medievistické skupiny se nám díky Lindat daří zprostředkovávat výsledky naší práce široké veřejnosti. Výstava o rukopisech je jedním z takových výstupů, kde jsme představili fascinující postavu latinsky psané kultury pozdního středověku – Kříže z Telče. Tato výstava byla doplňkem již zmíněné aplikace rukopisu A 4, v němž je Kříž většinou jeho písařem, stejně jako databáze pozdně středověkých textů opuscula.ff.cuni.cz. Výstava také sloužila jako představení těchto aplikací široké veřejnosti. Původně neproběhla jen v Třeboni, ale začala v Národní knihovně ČR, což byla skvělá příležitost, jak propojit univerzitu s takto významnými institucemi, jako je Národní knihovna a Třeboňský archiv. Aktuálně je výstava k vidění v Třeboňském archivu.



Vernisáž výstavy



Komentovaná prohlídka výstavy



Vstup na výstavu

Považujete nějakou dílčí část projektu za již dokončenou, nebo lze u jednotlivých výstupů stále něco zlepšovat?

JS: Je to klasická infrastruktura a běh na dlouhou trať. Zlepšovat se může s každým rokem a donekonečna jako každá jiná infrastruktura.

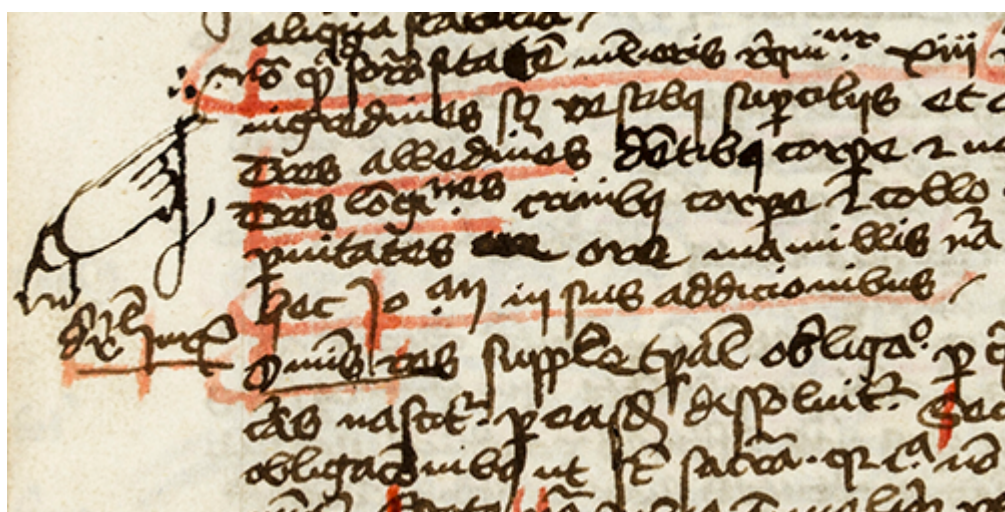
OF: Pokud jde o jednotlivé výstupy, můžeme říci, že například již zmíněné digitální prezentace rukopisů jsou u nás již uzavřené. Nyní pracujeme v rámci dané skupiny na jiných výzvách. Kdykoliv se nicméně můžeme k aplikacím vrátit, například díky novému objevu, a anotace jednotlivých míst obohatit o nové poznatky.

Jedna z větších sekcí výzkumné infrastruktury je vzdělávání. Co konkrétně v této oblasti nabízíte?

JS: Spolu s Centrem pro digital humanities FF UK, které bylo

letos ustanoveno, nabídneme různé workshopy. Konkrétní věci ladíme a něco by se mělo uskutečnit již příští rok.

OF: Myslím si, že je aspekt vzdělávání především dobře vidět na příkladu medievistické skupiny. Třeba právě výstava o Kříži z Telče umožnila studentům podílet se na vytváření panelové výstavy, napsat kapitoly odborné knihy a získat digitální dovednosti v podobě databáze pozdně středověkých textů Opuscula a digitální anotace rukopisu A 4.



Státní oblastní archiv Třeboň, Rukopisy Třeboň, A4, ff. 268v – 269r

Mgr. Jan Sládek, Ph.D., absolvoval magisterské i doktorské studium sociologie na FF UK, kde v současnosti i působí na Katedře sociologie FF UK. Odborně se zabývá především historií sociálního myšlení a také sociologií lidských sídel.

Mgr. Ondřej Fúsik je doktorand programu Anglický jazyk na Ústavu anglofonních literatur a kultur FF UK, kde zároveň spolu s kolegyní Mgr. Světlanou Müllerovou vyučuje předmět Čtení anglických historických textů I: Staroanglické období. Odborně se zaměřuje především na jazykový kontakt mezi latinou a starou angličtinou, paralelní korpus staré angličtiny a latiny, latinské nefinitní slovesné tvary a jejich překladové

ekvivalenty ve staré angličtině, ale také na lexikální pole staroanglických adjektiv a staroanglické překlady bible.

Projekt Lindat/Clariah-CZ je realizován za podpory Ministerstva školství, mládeže a tělovýchovy ČR v rámci velkých výzkumných infrastruktur pod kódem LM2018101.

Michal Otáhal