



UNIwersytet
Warszawski



Centrum Kompetencji Cyfrowych UW

ul. Smyczkowa 14

02-678 Warszawa

ckc@uw.edu.pl

Pan Piotr Adamowicz
przewodniczący Komisji Kultury i Środków Przekazu (KSP)

STANOWISKO CKC UW W SPRAWIE NOWELIZACJI USTAWY O PRAWIE AUTORSKIM

Wobec toczących się obecnie prac legislacyjnych nad projektem ustawy o zmianie ustawy o prawie autorskim i prawach pokrewnych oraz niektórych innych ustaw, Centrum Kompetencji Cyfrowych UW chce przedstawić dwa istotne z punktu widzenia praktyki naukowej oraz zabezpieczania dziedzictwa cyfrowego problemy, których dotyczy kształt proponowanej nowelizacji.

Centrum Kompetencji Cyfrowych UW to jednostka UW wspierająca badania naukowe przez projektowanie i opracowywanie ich komponentów cyfrowych. Funkcjonująca w ramach CKC UW Pracownia Archiwistyki Webu to jedyna w Polsce inicjatywa naukowa poświęcona archiwistyce internetu.

Problem I: Wolność badań naukowych i wpływ prawa na metody badań korpusowych

Wprowadzenie sztywnego limitu 20 proc. objętości do regulacji uprawnień przetwarzania treści (art 27 ust. 1 projektowanej ustawy) będzie przepisem blokującym wolność prowadzenia badań. Przepis ten bezpośrednio i negatywnie wpływać będzie na metody badań, których dobór i

kształt zależeć powinien wyłącznie od badacza/badaczki. W badaniach korpusowych (maszynowej analizy i interpretacji dużych zbiorów tekstów) oderwany od konkretnego zadania badawczego i specyfiki badanych treści limit 20 proc. dozwolonej objętości wprost podważy racjonalność metody badawczej.

Fundamentalną cechą korpusów badawczych jest reprezentatywność i zrównoważenie¹. Ta cecha korpusu może być zachowana także w przypadku korzystania wyłącznie z części dostępnych treści (zasobów). To badacz/badaczka jako jedyni są w stanie określić, wybór jak dużej części tekstów do analizy będzie najbardziej racjonalny. Jesteśmy zdania, że decyzja ta nie powinna należeć do legislatora, tym bardziej, że nie doprecyzowano, jak wyliczać ustawiony ogólnie limit 20 proc. (w przypadku książek np. czy ma zawierać okładkę, przypisy, bibliografię, czy tylko tekst merytoryczny). W praktyce badań korpusowych i podobnych takie sztywne ograniczenie będzie powszechnie ignorowane, a przepis stanie się martwy.

Problem II: Zabezpieczanie cyfrowego dziedzictwa i badania internetu

Pandemia COVID czy wojna w Ukrainie to tylko niektóre z wydarzeń, jakie dostępne będą przyszłym historykom w dużej mierze przez zachowane zasoby internetowe (od stron WWW przez materiały multimedialne aż po publikacje w mediach społecznościowych). Zasoby te zdecydowanie przynależą także do polskiego dziedzictwa (cyfrowego), którego ochrona należy do zadań polskich instytucji pamięci (biblioteki, archiwa). Art. 28 proponowanej ustawy mógłby zostać wykorzystany do zabezpieczenia uprawnień sektora nauki oraz instytucji pamięci do gromadzenia, przechowywania, badawczego analizowania i udostępniania (w określony sposób, np. na terminalach w siedzibie instytucji) części ogólnodostępnych zasobów WWW.

¹ V. Kamasa, *Techniki językoznawstwa korpusowego wykorzystywane w krytycznej analizie dyskursu. Przegląd*, "Przegląd Socjologii Jakościowej", 2014, 2, s. 101.

Archiwistyka Webu (*web archiving*) to dziedzina rozwijająca się od połowy lat 90., posiadająca własne metody, zbiory (archiwa Webu takie jak Wayback Machine), instytucje naukowe, czasopisma itp. Rozwój polskiej archiwistyki Webu, ale przede wszystkim możliwość zabezpieczania nietrwałych zasobów WWW dokumentujących ważne wydarzenia społeczne, polityczne czy kulturalne, wymaga odpowiednich regulacji prawnych. Polskie ustawodawstwo jest pod tym względem mocno opóźnione, np. większość krajów Europy zmodyfikowała swoje krajowe przepisy dotyczące egzemplarza obowiązkowego tak, aby umożliwić bibliotekom narodowym gromadzenie, przechowywanie, opracowywanie i udostępnianie w kontrolowanych warunkach archiwalnych zbiorów WWW².

Podkreślamy też, że gromadzenie i przetwarzanie zasobów WWW jest nie tylko sprawą ochrony dziedzictwa cyfrowego, ale też należy do metod badań naukowych. Niekiedy w praktyce badawczej konieczne jest pozyskanie i przechowywanie obszernych zbiorów treści dostępnych online (w tym stron WWW), tak aby dalej je maszynowo przetwarzać i analizować.

Nowelizacja prawa autorskiego nie powinna ograniczać, ale ułatwiać badania naukowe i stworzyć warunki do legalnego pozyskiwania i gromadzenia publicznych zasobów WWW, czy to prowadzonego w ramach projektów badawczych, czy w ramach misji ochrony dziedzictwa cyfrowego, podejmowanych przez instytucje takie jak biblioteki i archiwa.

² E. Vlassenroot et al., *Web archives as a data resource for digital scholars*, "International Journal of Digital Humanities", 2019, 1. s. 89, <https://doi.org/10.1007/s42803-019-00007-7>