

Przejście na dane FAIR okazało się wyjątkowo wymagające w obszarze “długiego ogona” danych badawczych. Dzieje się tak głównie dlatego, że długi ogon składa się z wielu niezależnie tworzonych zbiorów danych, różniących się strukturą i zarządzaniem, choć gromadzonych w podobnym celu. Niniejsza projekt opiera się na niedawno utworzonej sieci i bazie danych SPI-Birds w zakresie długoterminowych badań populacji ptaków, przechowującej i standaryzującej dane na poziomie indywidualnym.

Celem jest rozwój tej sieci jako modelu dla innych (ekologicznych) dziedzin badawczych, ze szczególnym uwzględnieniem wciąż niedocenianych etapów przetwarzania danych, analizy danych i ich zabezpieczania.

Po pierwsze, opracujemy repozytorium z cytowanymi zbiorami danych, które można powiązać z publikacjami, kodami analitycznymi i innymi wynikami, zamykając w ten sposób cykl życia badań w długim ogonie nauki.

Po drugie, rozszerzymy bazę użytkowników zarówno w odniesieniu do producentów danych (zwiększając geograficzny i ekologiczny zasięg zbiorów danych), badaczy w społeczności SPI-Birds i spoza niej, jak i użytkowników spoza nauki, takich jak edukatorzy, dziennikarze, specjaliści ds. projektanci.

Centralnym elementem tych dwóch pierwszych celów jest ciągły rozwój i udoskonalanie istniejącego standardowego formatu danych z bogatymi metadanymi (i dostępnymi przewodnikami po metadanych) zgodnie z powszechnymi standardami, takimi jak Darwin Core.

Po trzecie, stworzymy recenzowane archiwum (włączając w to testy przeprowadzone przez wybranych użytkowników) dla kodów przetwarzania danych i analitycznych (które mogą być stosowane do zbiorów danych hostowanych w bazie danych, ale także poza nią), promując wydajność i powtarzalność analiz danych dużej skali.

Naszym ostatecznym celem jest w pełni przejrzysty krajobraz badawczy FAIR obejmujący zbiory danych, narzędzia analityczne i publikacje, które są połączone cyfrowymi identyfikatorami i bogatymi metadanymi.