

En sosiopolitisk maskinlesning av norsk mellomkrigslitteratur, 1920-1945

Sofie Arneberg (Nasjonalbiblioteket), Pål C. Halvorsen (OsloMet), Lars G. B. Johnsen (Nasjonalbiblioteket)

Litteratursosiologi kan handle om å se på litteraturen som et sosialt artefakt, og som en inngang til en bestemt tids tankesett. I vår undersøkelse sammenligner vi tekster skrevet av to grupper forfattere: Forfattere som sympatiserte med okkupasjonsmakten i Norge mellom 1940 og 1945, og forfattere som var nærmere tilknyttet motstandsbevegelsen. Utvalget er basert på en studie av norske forfatters posisjoner innenfor økonomiske, kulturelle og sosiale hierarkier (Hjellbrekke et al., 2025).

Vår tilnærming tar sikte på å hente ut sosiologiske variabler (kjønn, klasse, alder og annet) fra litterære tekster og undersøke hvordan disse reflekterer underliggende samfunnsstrukturer. På denne måten vil vi undersøke hvordan fiktive karakterer blir posisjonert, sosiale roller og allianser kommer til uttrykk, og konflikt er innvevd i språket. I stedet for å basere oss utelukkende på næranalyse, kombinerer vi korpusmetoder som frekvenstelling og navngitt enhetsgjenkjenning (NER) med store språkmodeller (LLM-er).

Korpuset er hentet fra den norske nasjonalbibliografien og tilhørende digitaliserte filer. Det har vært et poeng å ikke bryte koblingen til kilden, den åpne bibliotekskatalogen, med sine internasjonalt standardiserte formater for bibliografisk beskrivelse. Et av målene med arbeidet er derfor å utvikle en metode som ivaretar sporbarhet og som kan anvendes til lignende arbeider med utgangspunkt i bibliografiske beskrivelser.

A socio-political machine reading of Norwegian interwar literature, 1920-1945

Literary sociology can approach literature as a social artifact and as an entry point to the mindset of a given era. In this study, we compare texts written by two groups of Norwegian authors: those who sympathized with the occupying power during 1940–1945, and those who were more closely aligned with the resistance movement. The selection is based on a study of Norwegian authors' positions within economic, cultural, and social hierarchies (Hjellbrekke et al., 2025).

Our approach aims to extract sociological variables (such as gender, class, and age) from literary texts to investigate how these reflect underlying social structures. In doing so, we analyze how fictional characters are positioned, social roles and alliances are expressed, and conflict is embedded in language. Rather than relying solely on close reading, we combine corpus-based methods including frequency counts and named entity recognition (NER) with large language models (LLMs).

The corpus is derived from the Norwegian national bibliography and its associated digitized files. Preserving the link to the source, the open library catalogue with its internationally standardized formats for bibliographic description has been central to our work. One of the aims of the project is therefore to develop a method that safeguards traceability and can be applied to similar studies grounded in bibliographic metadata.