Sluttrapport

Prosjektinformasjon og rapporteringsmål

Prosjektinformasjon

Prosjektnummer:	328745
Prosjekttittel:	AI enriched journalism to investigate illicit financial flows
Program / Aktivitet:	FINANSMARKED
Prosjektleiar:	Krøvel, Roy
Prosjektansvarleg:	OSLOMET - STORBYUNIVERSITETET
Prosjektperiode:	01.08.2022 - 31.12.2025

Rapporteringsmål

1. **Hovudside framdriftsrapport:** Ajourfør framdriftsrapport fram til prosjektslutt. — **Utført**

2. **Sluttrekneskap:** Gi eit samandrag av økonomien i prosjektet. — **Utført**

3. **Verknader og effektar:** Eg er inneforstått med at teksta i feltet for Virkninger og effekter framstilling vil verta offentleg tilgjengeleg* — **Utført**

4. **Resultatrapport:** Legg ved resultatrapport. — **Utført**

5. **Særskild rapportering:** Dersom det ligg føre krav om særskild rapportering, skal dette gjerast. — **Ikkje aktuelt**

6. **Endeleg datahandteringsplan:** Er endeleg datahandteringsplan lasta opp? — **Utført**

Sluttrapport - **328745** - AI enriched journalism to investigate illicit financial flows
Rapporteringsfrist: 31.01.2026 - Innsendt: 29.01.2026

Side 2 / 10

Sluttrekneskap

Faktisk kostnadsplan (Beløp i NOK 1000)

Konto	2025	2024	2023	2022	Totalsum
Personal- og indirekte kostnader	684	87	23	42	836
Innkjøp av FoU-tenester	0	0	0	0	0
Utstyr	0	0	0	0	0
Andre driftskostnader	416	593	38	0	1 047
Sum	1 100	680	61	42	1 883

Faktisk kostnadsstad (Beløp i NOK 1000)

Konto	2025	2024	2023	2022	Totalsum
Næringsliv	0	0	0	0	0
Instituttsektor	0	0	0	0	0
UoH-sektor	1 100	680	61	42	1 883
Andre sektorar	0	0	0	0	0
Utlandet	0	0	0	0	0
Sum	1 100	680	61	42	1 883

Faktisk finansieringsplan (Beløp i NOK 1000)

Konto	2025	2024	2023	2022	Totalsum
Forskingsrådet	336	337	406	321	1 400
Eigenfinansiering	256	155	30	42	483
Offentleg finansiering	0	0	0	0	0
Privat finansiering	0	0	0	0	0
Internasjonale midlar	0	0	0	0	0
Avvik	−508	−188	375	321	0
Avvik grunnlag	1 100	680	61	42	1 883
Sum	592	492	436	363	1 883

Kommentar

Verknader og effektar

Forventa verknader og effektar - hentet frå søknaden

-

Sluttrapport - **328745** - AI enriched journalism to investigate illicit financial flows
Rapporteringsfrist: 31.01.2026 - Innsendt: 29.01.2026

Side 3 / 10

Oppnådde verknader og effektar - basert på prosjektets resultater

a. Improve research methodologies using AI/ML to investigate beneficial ownership
Achieved. The project developed an AI-based extraction methodology achieving 99.6 percent accuracy on well-formatted documents. The methodology is fully documented (129+ pages) and reproducible. More significantly, we systematically mapped where AI methods succeed and fail, established that extraction can be automated but verification cannot, and documented structural obstacles limiting what any methodology can achieve. This helps investigators allocate resources realistically.
b. Test methodologies by investigating 10 international companies
Partially achieved. We validated extraction across multiple companies including Equinor (513 subsidiaries), Shell (1,824 subsidiaries), Chevron, ExxonMobil, and TotalEnergies. Testing revealed that success depends more on disclosure regime than company characteristics. This shifted our focus from company-by-company investigation toward understanding regulatory structures that determine what can be known.
c. Make results available to regulators, academics, journalists and NGOs
In progress. One research article documenting methodology and findings has been submitted for peer review. Extraction tools and documentation are available for other researchers and journalists. Findings about structural obstacles are directly relevant for policy discussions about disclosure requirements.
d. Improve capacity of media to investigate illicit financial flows
Achieved in part. Tools and methodology can help journalists extract subsidiary data more efficiently. Collaboration with ICIJ ensured development was informed by practitioner needs. The project's most important contribution may be documenting what cannot currently be achieved and why, helping journalists frame investigations appropriately.
Potential longer-term impacts
The finding that corporate opacity results from regulatory choices rather than technical limitations has policy relevance. With standardised disclosure formats and eliminated materiality thresholds, our methodology would enable routine corporate mapping at scale. The gap between technical possibility and regulatory reality is now clearly documented.

Resultatrapport

Opplastet resultatrapport

Originalfilnavn: Results Report.pdf

Filreferansenavn: RESULTAT_Sluttrapport11906466.pdf

Melding til Noregs forskingsråd

Særskild rapportering

Kommentar

Opplastet vedlegg

Endeleg datahandteringsplan

Opplastet datahandteringsplan

Originalfilnavn: Datahandteringsplan.pdf

Filreferansenavn: DATAHAND_Sluttrapport11906466.pdf

Framdriftsrapport

Prosjektinformasjon og rapporteringsmål

Prosjektinformasjon

Prosjektnummer:	328745
Prosjekttittel:	AI enriched journalism to investigate illicit financial flows
Program / Aktivitet:	FINANSMARKED
Prosjektleiar:	Krøvel, Roy
Prosjektansvarleg:	OSLOMET - STORBYUNIVERSITETET
Prosjektperiode:	01.08.2022 - 31.12.2025
Rapporteringsperiode:	01.10.2024 - 31.12.2025

Rapporteringsmål

1. **Populærvitenskapeleg framstilling:** Eg er innforstått med at teksten i felta for Populærvitskapleg framstilling vil bli offentleg tilgjengeleg* — **Ja**

2. **Resultatinformasjon:** Er opplysningar om publisering gitt? — **Ja**

3. **Resultatindikatorar:** Alle resultatdata som er komne fram i prosjektet, skal rapporterast. Er rapportering gjort? — **Ja**

4. **Stipend:** Opplysningar om alle stipend må vere fullstendige og rette. Har du oppdatert månadsverk og andre opplysningar for kvar stipendiat? — **Ja**

5. **Internasjonalt samarbeid:** Ein skal gi nemne omfanget av internasjonalt samarbeid. Har det vore slikt samarbeid i perioden? — **Ja**

6. **Særskild rapportering:** Dersom det ligg føre krav om særskilt rapportering i ei eiga melding frå saksbehandlar i Forskingsrådet, skal dette gjerast. Er særskild rapportering gjort? — **Nei**

Populærvitskapleg framstilling

Populærvitskapleg framstilling (Norsk)

Panama Papers, Paradise Papers, Pandora Papers og andre lekkasjer har vist omfanget av problemet med illegitime finansstrømmer. FNs bærekraftsmål understreker behovet for å forstå bedre hvordan illegitime økonomiske strømmer skjer, samt de samfunnsmessige konsekvensene av slike strømmer. Journalister som etterforsker ulovlige strømmer over landegrensene står imidlertid ofte overfor den nesten uoverkommelige oppgaven å finne de reelle eierne bak selskaper som etterforskes. På den ene siden kan selskaper benytte seg av nettverk av datterselskaper registrert i ulike jurisdiksjoner, for eksempel såkalte skatteparadiser eller bedriftsparadiser. I tillegg vil nettverkene ofte bestå av flere lag eller nivåer. Å finne de reelle eierne har vist seg å være en tidkrevende og kostbar del av mange journalistiske undersøkelser. Likevel illustrerer undersøkelser som Panama Papers og Pandora Papers også potensialet til undersøkende journalistikk for å avdekke og formidle kunnskap av betydning for samfunnet.

Dette prosjektet søker å bidra til slik innsats ved å utvikle verktøy basert på kunstig intelligens for å hjelpe etterforskning til journalister som undersøker ulovlige finansstrømmer og reell eierskap. Prosjektet er tverrfaglig med deltakere fra journalistikk, journalistikkstudier, informatikk og økonomi. Prosjektet inneholder også aktiviteter for å kommunisere med undersøkende journalister.

Populærvitskapleg framstilling - Oppdatert (Norsk)

Panama Papers, Paradise Papers, Pandora Papers og andre lekkasjer har vist omfanget av problemet med illegitime finansstrømmer. FNs bærekraftsmål understreker behovet for å forstå bedre hvordan slike strømmer skjer. Journalister som etterforsker ulovlige strømmer over landegrensene står imidlertid ofte overfor den nesten uoverkommelige oppgaven å kartlegge selskapsstrukturer og finne de reelle eierne. Selskaper kan benytte nettverk av datterselskaper registrert i ulike jurisdiksjoner, ofte i flere lag eller nivåer. Å kartlegge slike strukturer har vist seg å være en tidkrevende del av mange journalistiske undersøkelser.

Dette prosjektet utviklet og testet metoder basert på kunstig intelligens for å automatisere uttrekk av datterselskapsdata fra årsrapporter og børsmeldinger. Vi oppnådde høy nøyaktighet (99,6 prosent) på godt formaterte dokumenter. Men prosjektets viktigste funn var at hindringene for å kartlegge selskapsstrukturer i hovedsak ikke er tekniske. De er strukturelle og politiske: regulatoriske unntak som tillater selektiv rapportering, inkonsistente formatkrav, og fragmenterte registre på tvers av jurisdiksjoner. Prosjektet dokumenterte systematisk disse hindringene og utviklet metodikk som kan hjelpe journalister å arbeide mer effektivt innenfor dagens begrensninger, samtidig som funnene viser hva som må endres for at åpenhet om selskapsstrukturer skal bli praktisk mulig.

Populærvitskapleg framstilling (Engelsk)

The Panama Papers, Paradise Papers, Pandora Papers and other leaks have shown the extent of the problem of illicit financial flows. The UN's sustainability goals emphasize the need to better understand how illicit economic flows occur, as well as the societal consequences of such flows. However, journalists investigating illicit flows across national borders are often faced with the almost insurmountable task of finding the real owners behind the companies being investigated. On the one hand, companies can make use of networks of subsidiaries registered in various jurisdictions, for example so-called tax havens or corporate havens. In addition, the networks will often consist of several layers or levels. Finding the real owners has proven to be a time-consuming and expensive part of many journalistic investigations. Nevertheless, investigations such as the Panama Papers and Pandora Papers illustrate the potential of investigative journalism to uncover and disseminate knowledge of importance to society.

This project seeks to contribute to such efforts by developing tools based on artificial intelligence to aid investigations for journalists investigating illicit financial flows and beneficial ownership. The project is interdisciplinary with participants from journalism, journalism studies, computer science and economics. The project also includes activities to communicate with investigative journalists.

Sluttrapport - **328745** - AI enriched journalism to investigate illicit financial flows
Rapporteringsfrist: 31.01.2026 - Innsendt: 29.01.2026

Side 6 / 10

Populærvitskapleg framstilling - Oppdatert (Engelsk)

Panama Papers, Paradise Papers, Pandora Papers and other leaks have revealed the scale of the problem of illicit financial flows. The UN Sustainable Development Goals underline the need to better understand how such flows occur. Journalists investigating cross-border financial crime often face the near-insurmountable task of mapping corporate structures and finding the real owners. Companies can use networks of subsidiaries registered in different jurisdictions, often in multiple layers. Mapping such structures has proven to be a time-consuming part of many journalistic investigations.
This project developed and tested methods based on artificial intelligence to automate the extraction of subsidiary data from annual reports and regulatory filings. We achieved high accuracy (99.6 percent) on well-formatted documents. But the project's most important finding was that the obstacles to mapping corporate structures are not primarily technical. They are structural and political: regulatory exemptions that permit selective reporting, inconsistent format requirements, and fragmented registries across jurisdictions. The project systematically documented these obstacles and developed methodology that can help journalists work more efficiently within current constraints, while the findings also show what must change for transparency about corporate structures to become practically achievable.

Melding til Noregs forskingsråd

Vi har oppdatert sammendraget for prosjektet for å gjenspeile hva vi faktisk oppnådde og fant ut.

Resulatinformasjon

Kategori: Artikkel (vitensk)

Forfatter(e)	Tittel	Tidskrift tittel	Sidenummer, frå - til	Volum	År	ISSN /ISBN	DOI
Fridman, Marina; Krøvel, Roy; Palumbo, Fabrizio	How (not to) Run an AI Project in Investigative Journalism	Journalism Practice			2023		10.1080/17512786. 2023.2253797

Kategori: Oppslag i massemedia

Forfatter(e)	Tittel	Sted/Type	År	ISSN/ISBN	DOI
Krøvel, Roy	Få hjelp til å utvikle metoder for graving med AI	Skup.no	2024		
Gjengedal , Kjerstin	Prosa #4 23: Kunstig intelligens	Prosa	2023		
NTB	Starter kunnskapsnettverk for AI i mediene — nå ønsker de at flere blir med		2023		

Kategori: Fagleg formidling

Forfatter(e)	Tittel	Tidsskrift/Utgiver/Arrangement	År	ISSN/ISBN	DOI
Palumbo, Fabrizio	AI Journalism Panelist	Global Investigative Journalism Conference Kuala Lumpur	2026		
Cosic, Jelena	AI and Investigative Journalism: Opportunities and Challenges	pulitzercenter.org	2024		
Cosic, Jelena	AI as a Tool for Investigations	Global Investigative Journalism Conference	2023		
Cosic, Jelena	Cross-border Teams: What Could Possibly Go Wrong?	UNCOVERED 23	2023		
Krøvel, Roy	KI i norsk journalistutdanning	Data-SKUP	2023		
Palumbo, Fabrizio	Using AI in investigative journalism	Data SKUP	2022		

Sluttrapport - **328745** - AI enriched journalism to investigate illicit financial flows
Rapporteringsfrist: 31.01.2026 - Innsendt: 29.01.2026

Side 8 / 10

Kategori: Anna publisering

Forfatter(e)	Tittel	Tidsskrift/Type/Utgiver	År	ISSN/ISBN	DOI
SKUP	Få hjelp med AI	SKUP.no	2023		

Sluttrapport - **328745** - AI enriched journalism to investigate illicit financial flows
Rapporteringsfrist: 31.01.2026 - Innsendt: 29.01.2026

Side 9 / 10

Resultatindikatorer

Allmennretta formidlingstiltak
Populærvitenskaplege publikasjonar (artiklar/bøker, debattbøker/-artiklar, høyringar, utstillingar, skjønnlitteratur etc)

2022	2023	2024	2025	Akkumulert hittil
0	0	0	0	0

Oppslag i massemedia (aviser, radio, TV mm)

2022	2023	2024	2025	Akkumulert hittil
0	2	1	0	3

Brukarretta formidlingstiltak
Rapportar, notat, artiklar, foredrag på møte/konferansar retta mot målgruppene i prosjektet

2022	2023	2024	2025	Akkumulert hittil
1	3	1	0	5

Næringsretta FoU resultater
Ferdigstilte nye/forbetra metodar/modellar/prototypar

2022	2023	2024	2025	Akkumulert hittil
0	0	0	0	0

Innføring av nye/forbetra metodar/modellar/teknologi for auka verdiskaping
Bedrifter utanfor prosjektet som har innført nye/forbetra metodar/modellar/teknologi

2022	2023	2024	2025	Akkumulert hittil
0	0	0	0	0

Bedrifter i prosjektet som har innført nye/forbetra arbeidsprosessar/forretningsmodellar

2022	2023	2024	2025	Akkumulert hittil
0	0	0	0	0

Bedrifter i prosjektet som har innført nye/forbetrae metodar/teknologi

2022	2023	2024	2025	Akkumulert hittil
0	0	0	0	0

Ny verksemd
Nye foretak som følgje av prosjektet

2022	2023	2024	2025	Akkumulert hittil
0	0	0	0	0

Sluttrapport - **328745** - AI enriched journalism to investigate illicit financial flows
Rapporteringsfrist: 31.01.2026 - Innsendt: 29.01.2026

Side 10 / 10

Nye forretningsområde i eksisterande bedrifter som følgje av prosjektet

2022	2023	2024	2025	Akkumulert hittil
0	0	0	0	0

Vitskaplege publikasjonar

Tidsskriftsartikkel/periodika

2022	2023	2024	2025	Akkumulert hittil
0	1	0	0	1

Stipend

Stipend finansierte av prosjektet

Internasjonalt samarbeid

Internasjonalt samarbeid finansiert av prosjektet (i NOK 1000)

Beløp i NOK 1000

Land	2022	2023	2024	2025

Særskild rapportering

Kommentar

Opplastet vedlegg

Results Report: AI-Enriched Journalism for Investigating Illicit Financial Flows

Background and Objectives

The project set out to develop artificial intelligence-based tools to help investigative journalists map corporate ownership structures. Major document leaks such as the Panama Papers and Pandora Papers have demonstrated both the scale of illicit financial flows and the potential of investigative journalism to expose them. However, journalists investigating such flows face the time-consuming task of tracing subsidiaries across multiple jurisdictions and corporate layers. The project aimed to reduce this burden through automated extraction methods.

The project was funded by the Norwegian Research Council in partnership with the International Consortium of Investigative Journalists (ICIJ). It was conducted at OsloMet (Oslo Metropolitan University) with contributions from experienced investigative journalists and data analysts with doctoral-level expertise in AI and machine learning.

The original objectives were twofold. First, to develop AI-based methodologies that could help unveil patterns in how multinational corporations use corporate and tax havens through their subsidiary networks. Second, to test whether web crawling and machine learning could create new research methodologies for investigating beneficial ownership.

Results Compared to Objectives

The project achieved its technical objectives. We developed and validated an extraction system that achieves 99.6 percent accuracy when extracting subsidiary information from well-formatted corporate disclosure documents. The system was validated against independent ground truth data (Equinor ASA's 513 subsidiaries) and tested across multiple companies of different types (Norwegian, US, UK) and sizes (from 30 to over 1,800 subsidiaries).

However, the project's most significant finding was unexpected. We discovered that the obstacles to mapping corporate structures are not primarily technical. The technical problem of extracting subsidiary data from documents is essentially solved. The persistent difficulties are structural and political: regulatory frameworks that permit selective disclosure, inconsistent format requirements that defeat systematic analysis, and fragmented registries that prevent cross-jurisdictional verification.

This finding reframes the contribution. Rather than simply providing better tools, the project documented why comprehensive corporate mapping remains difficult even when technical solutions exist. The barriers include:

The "significant subsidiary" threshold in SEC regulations, which allows US companies to omit subsidiaries below 10 percent of assets

Format variations in corporate disclosures that satisfy legal requirements while remaining practically unusable for systematic analysis

Registry fragmentation across jurisdictions with no common identifiers or data standards

The gap between formal transparency (what regulations require) and practical transparency (what can actually be analysed)

Research Activities and Contributing Partners

The project organised research activities along three complementary paths.

Path A: Verification workflow methodology. Led by an experienced investigative journalist, this path documented the complete workflow for extracting and verifying corporate subsidiary information. The work measured where time goes in corporate ownership investigations and identified the verification bottleneck: while extraction can be largely automated, verification against primary sources remains manual and time-intensive.

Path B: Automated extraction methodology. Led by a journalism researcher at OsloMet working intensively with AI tools, this path developed and validated the extraction pipeline. The system uses large language models (GPT-OSS:120B via Ollama Turbo API) combined with PDF processing tools to extract subsidiary names and jurisdictions from corporate documents. Validation against ground truth data established the accuracy and limits of automated extraction.

Path C: Population-level analysis. Led by data analysts with doctoral training in AI and data science, this path tested automated extraction at scale on SEC-registered companies. The work identified which companies and document types support reliable extraction and which do not, revealing that success rates vary significantly depending on disclosure format and regulatory regime.

The three paths developed in mutual support, with extraction tools from Path B incorporated into verification workflows from Path A, and validation requirements from Path A defining accuracy standards for Path B. Path C identified patterns warranting deeper investigation through the methods developed in Paths A and B.

Assessment of Project Implementation and Resource Use

The project was implemented largely as planned, though the focus shifted as findings emerged. Initial plans emphasised tool development and pattern discovery. As work progressed, the systematic documentation of structural obstacles became increasingly central.

Resource allocation reflected this shift. Significant effort went into testing extraction approaches that ultimately failed or proved limited, including IBM Docling, Qwen-3, and various regex-based methods. This testing was necessary to establish which approaches work reliably and which do not. The negative findings are themselves valuable, as they prevent future investigators from repeating unsuccessful approaches.

The collaboration with ICIJ functioned as intended. ICIJ partners provided insight into how journalists approach corporate ownership investigations, what challenges they face, and what tools would be useful. This input shaped both tool development and the documentation of obstacles.

International Stays

There were no extended international stays during the project period. Collaboration with ICIJ was conducted remotely through interviews, discussions, and feedback on methodological development.

Value and Expected Utility of Results

The results have value for multiple audiences.

For investigative journalism: The extraction tools and documented methodology can help journalists work more efficiently within current constraints. The Skills documentation (over 129 pages) encodes practical knowledge about which data sources serve which investigative purposes, how to navigate format variations, and where automation reaches its limits. This knowledge is immediately usable by other investigators.

For journalism research: The project demonstrates what methodological research in journalism can look like. By building and testing rather than only observing, we produced empirical findings about where AI tools add value and where human judgment remains essential. The extraction-verification asymmetry, the scale effects that limit hybrid workflows, and the source selection findings could not have been discovered through interviews or retrospective analysis alone.

For transparency policy: The systematic documentation of obstacles has implications for how disclosure regimes are designed and evaluated. Our findings show that compliance-oriented disclosure and investigation-oriented disclosure are not the same thing. The gap between formal and practical transparency documented in this project suggests that transparency policies should be evaluated not only by what they require to be disclosed but by whether the disclosures produced are practically usable for systematic analysis.

For SDG 16.4: The UN Sustainable Development Goals include a target to reduce illicit financial flows. Our findings show that measuring and mapping corporate structures at scale is not currently possible, even for companies subject to extensive disclosure requirements. The obstacles are not technical limitations but policy choices that could be changed. Comprehensive disclosure requirements, standardised formats, and integrated registries would transform corporate mapping from an investigative challenge into a routine database query.

Plans for Dissemination and Utilisation

Academic publication: One research article has been completed and submitted for peer review. The article documents the three methodological paths, presents findings on the extraction-verification asymmetry and structural obstacles to transparency, and discusses implications for investigative journalism methodology. The article has not yet been accepted for publication. We can share the submitted manuscript with the Research Council if more detailed information is of interest.

Tool availability: The extraction code and Skills documentation are available for other researchers and journalists to use and build upon. The methodology is documented in sufficient detail for reproduction.

Practitioner communication: Project findings have been shared with ICIJ partners and will be communicated to the broader investigative journalism community through appropriate channels.

Results Expected After Project Completion

The submitted research article is expected to complete peer review and be published following any required revisions. Depending on reviewer feedback, additional analyses or documentation may be developed.

The extraction tools and methodology will remain available for use by journalists and researchers. We anticipate that others will adapt the tools for additional companies, jurisdictions, and document types beyond those tested in the project.

The structural findings about obstacles to corporate transparency may inform policy discussions about disclosure requirements, though such impact depends on factors beyond the project's control.