

# Bilag: Datagrundlag, metode, prompt og valideringsresultater

## Indholdsfortegnelse

<b>1. Datagrundlag.....</b>	<b>2</b>
1.1 De store virksomheders årsrapport og redegørelse for dataetik .....	2
1.2 Branchekoder .....	2
<b>2. Metode.....</b>	<b>3</b>
2.1 Indhentning af årsrapporter.....	3
2.2 Procedure for identifikation af dataetik i årsrapporterne .....	3
2.3 Brug af GPT-4 til at identificere og klassificere afsnit om dataetik.....	4
2.4 Klassificering af om virksomhederne har en dataetisk politik eller ej.....	5
2.5 Udvælgelse af kategorier .....	6
2.6 Validering af GPT-4 og prompts .....	6
2.7 Statistisk usikkerhed.....	9
<b>3. Promptet.....</b>	<b>9</b>
<b>4. Valideringsresultater .....</b>	<b>14</b>
<b>5. Referenceliste .....</b>	<b>16</b>

# 1. Datagrundlag

Rapporten er baseret på virksomhedernes afsnit om dataetik i deres årsrapport og generel firmastatistik.

---

## 1.1 De store virksomheders årsrapport og redegørelse for dataetik

Virksomheder i regnskabsklasse C (stor) og D skal i deres årsrapport rapportere på, hvordan de forholder sig til dataetik<sup>1</sup>. Dette kan gøres på flere måder. Enten kan årsrapporteringen fremgå som et separat afsnit, eller det kan redegøres for som en del af årsberetningen. Derudover kan virksomhederne både vælge at nedskrive deres dataetiske politik, henvise til den andetsteds (fx ESG-rapport eller hovedkoncern), eller de kan forholde sig til dataetik ved at begrunde, hvorfor de ikke har en dataetisk politik.

Hvordan virksomhederne vælger at formulere deres dataetiske politikker, samt hvilke overvejelser eller argumenter, de gør brug af, er ikke defineret i lovgivningen. Der er heller ikke krav til, hvad en begrundelse, for ikke at have en dataetisk politik, skal indeholde.

Denne rapport medtager alene afsnit om dataetik i årsrapporterne eller virksomhedernes links til deres dataetiske politikker andetsteds. Den indeholder dermed ikke andre afsnit om fx ligestilling, regnskaber, bæredygtighed osv. Det undersøges ikke om virksomhederne evt. har interne retningslinjer eller politikker vedrørende dataetik, som der ikke er henvist til i årsrapporten.

## 1.2 Branchekoder

For at kunne undersøge evt. brancheforskelle i hvordan virksomhederne forholder sig til dataetik, krydses virksomhedernes CVR med hvilken branchekode, der passer til den pågældende virksomhed. Her er anvendt DB07 branchekoden. Brancher med færre end 20 virksomheder indgår ikke i analysen for at sikre, at enkeltvirksomheder ikke kan identificeres.

---

<sup>1</sup> Finansielle virksomheder er ikke indbefattet af årsregnskabsloven og indgår dermed ikke i denne rapport. Lovgivningen for rapportering af dataetik i finansielle rapporter trådte i kraft i 2023 og første rapporteringsår vil dermed være 2024. Af denne årsag indgår der ikke finansielle virksomheder i rapporten, som ikke er en del af regnskabsklasserne C (stor) og D.

## 2. Metode

Rapporten viser en række resultater på baggrund af datamateriale udarbejdet med brug af OpenAI's GPT-4. I nedenstående uddybes en række forhold vedrørende databehandling og metode, der har betydning for tolkningen af resultaterne.

---

### 2.1 Indhentning af årsrapporter

Det er kun store virksomheder i regnskabsklasse C og virksomheder i regnskabsklasse D, som er underlagt kravet om at afrapportere, hvorvidt de har en dataetisk politik i deres årsrapporter. Derfor indhentes kun data om disse virksomheders årsrapporter i analysen<sup>2</sup>.

Virksomheder i regnskabsklasse C og regnskabsklasse D skal ifølge årsregnskabslovens §138 afrapportere deres årsrapport senest 6 måneder efter regnskabsårets slutning. Alle årsrapporter registreres og offentliggøres af Erhvervsstyrelsen. I analysen er der blevet indsamlet to årgange af årsrapporter: Regnskabsåret 2021 (indberettet i 2022 og henvises til i denne rapport som år 2022) og regnskabsåret 2022 (indberettet i 2023 og henvises til i denne rapport som år 2023).

Årsrapporter er i denne analyse defineret som at høre til det årstal, hvor perioden for rapporteringen har foregået. Det vil sige, at alle årsrapporter, hvis regnskabsperiode starter mellem d. 1. januar og 31. december 2022, men hvor fristen for indsendelse af årsrapporten til Erhvervsstyrelsen er juni 2023, kategoriseres som værende en årsrapport fra 2023.

Dataindsamlingen er foretaget d. 2. november 2023. Det betyder, at alle årsrapporter, hvis regnskabsperioder er startet senest d. 2. november 2022 bør være inkluderet. Der er dog risiko for forsinkede afrapporteringer, forsinkelser på hvornår Erhvervsstyrelsen tilføjer årsrapporten til API'en og lignende. Hovedfokus i denne analyse er årsrapporter fra år 2023, hvor rapporter fra år 2022 udelukkende bruges til sammenligning over tid.

### 2.2 Procedure for identifikation af dataetik i årsrapporterne

Efter at have indhentet data fra årsrapporterne blev det klassificeret, hvorvidt virksomheden nævnte dataetik i deres årsrapport. Der blev søgt efter nøgleord som ville fremgå, når virksomheden havde et afsnit om dataetik.<sup>3</sup> Dette blev gjort for at spare serverkraft fra GPT-4 og OpenAI.

---

<sup>2</sup> Dette klassificeres gennem Erhvervsstyrelsens virksomhedsdata.

<sup>3</sup> Ordstrengen: ("dataetisk", "data etik", "dataetik", "data-etik", "dataetisk", "data etisk", "data ethical", "data ethic", "data ethics", "data ethic", "data-ethics", "dataethics", "dataethic", "dataethical", "99d", "§99d"). Det er vigtigt at bemærke, at der i XML-formatet, som årsrapporterne indberettes i, er et afgrænset format, hvor overskriften til afsnittet om virksomhedens dataetiske politik skal være et label, som inkluderer 'data ethics'. Derfor var det ikke nødvendigt at inddrage ord som "dataansvarlighed" eller lignende. En stikprøve på 30 af disse blev gennemgået i PDF-format for at sikre, at metoden var korrekt.

De virksomheder, som slet ikke afrapporterer om dataetik, kategoriseres som, at de ikke forholder sig til dataetik i det pågældende år, og databehandlingen for deres årsrapporter stopper derefter. For 2023-årgangen gjaldt dette 218 årsrapporter.

For de resterende årsrapporter, der nævner dataetik i deres årsrapport, var næste skridt at afgrænse det område i årsrapporten, hvor virksomhedernes dataetiske politik (eller mangel herpå) blev beskrevet. Dette blev gjort på følgende måde:

1. De(t) sted(er), hvor nøgleord, som fx 'dataetik', nævnes i årsrapporten, identificeres, og der laves et udtræk på 2500 tegn, som kommer før ordet og 2500 tegn, som kommer efter ordet.<sup>4</sup>
2. Der laves udtræk for alle fremkomster af nøgleord (der kan fx godt stå 'dataetik' flere steder undervejs i rapporten). Det sikrer, at der ikke frafalder relevant information. Alle årsrapporter har et udtræk på alt mellem 5.000-50.000 tegn, hvor vi ved, at deres afsnit om dataetik indgår, og samtidigt sikres det med høj sandsynlighed, at hele deres dataetiske redegørelse er med.

### 2.3 Brug af GPT-4 til at identificere og klassificere afsnit om dataetik

For at afgrænse *præcis* det afsnit i årsrapporterne, som omhandler virksomhedens dataetiske politik (eller mangel herpå) blev der anvendt en API-adgang til OpenAI's sprogmodel gpt-4-1106-preview (oftest refereret til som GPT-4-Turbo og i denne tekst refereret til som GPT-4). Udover at benytte dette værktøj til at afgrænse det dataetiske afsnit, så blev det også brugt som et analyseværktøj til at klassificere afsnittene ud fra en række parametre.

Databearbejdningen blev opsat på følgende måde:

1. En *pre-prompt* blev defineret. Denne prompt blev genbrugt til hver eneste årsrapport, udelukkende med én variation. Variationen er selve det udtræk, hvor virksomheden redegør for dataetik, der blev identificeret i 2.2.
2. Denne *pre-prompt* blev stillet til hver eneste årsrapport. I prompten er det beskrevet, at modellen skal svare i samme format hver gang for at gøre efterfølgende databehandling nem.
3. Derefter blev et prompt udviklet ud fra en række kategoriseringer og temaer. Udformningen af selve kategoriseringerne uddybes i afsnit 2.5.
4. Modellens output blev til sidst transformeret om til et samlet datasæt.

GPT-4 blev dermed brugt til at klassificere på følgende områder:

1. At afgrænse det præcise dataetiske afsnit.

---

<sup>4</sup> 2500 tegn er udvalgt efter trial-and-error metoden og ved at løbende at validere udtræk af stikprøver i datasættet.

2. At klassificere de virksomheder uden et dataetisk afsnit til ikke at have redegjort for dataetik.
3. At klassificere hvorvidt virksomheden har en dataetisk politik eller ej.
4. At definere hvorvidt virksomhedens dataetiske politik fremgår direkte i årsrapporten, eller om der refereres til et andet sted, hvor politikken kan hentes (fx på et internt link eller i deres moderkoncerns årsrapport).
5. At klassificere alle de indhentede dataetiske politikker ud fra en række prædefinerede kategorier.
6. At klassificere virksomhedens begrundelse for *ikke* at have en dataetisk politik ud fra en række prædefinerede kategorier.

Der blev udelukkende stillet afgrænsede spørgsmål/vejledninger i *promptet*, og langt de fleste var dikotomiske, så det var så simpelt som muligt for modellen at forholde sig til. For GPT-4 er det muligt at indstille en 'temperatur' for, hvor stabilt eller 'frit' modellen skal svare og kategorisere. Her blev modellens 'temperatur' sat lavt – det betyder, at den svarer så stabilt som muligt hver gang, og at den typisk vil svare det samme, hvis man bad den om at gentage samme kategorisering.<sup>5</sup>

I flere årsrapporter (find tal) refererede virksomheden til en dataetisk politik på deres hjemmeside, deres CSR-rapport eller hos deres moderselskab. I disse tilfælde identificerede sprogmodellen, at virksomheden *har* en dataetisk politik, men den kunne ikke klassificere politikken ud fra prædefinerede parametre, da dette var udpenslet i selve rapporten. Derfor blev disse politikker indhentet manuelt af en medarbejder ved at følge virksomhedens reference og ved at oprette simple søgninger på virksomhedens hjemmeside.

I tilfælde hvor politikken ikke kunne opspores, kategoriseres virksomheden til ikke at have forholdt sig til dataetik. I de resterende tilfælde blev årsrapporterne, med det nu indhentede data, klassificeret endnu engang af GPT-4, og svarene blev indsat i datasættet og analyseret på lige fod med resten.

Ovenstående procedure gælder for årsrapporter fra 2023. Årsrapporterne fra 2022 blev udelukkende brugt til at sammenligne over tid på antallet af dataetiske politikker. Her blev der kun anvendt klassificeringstrin 1-3.

## **2.4 Klassificering af om virksomhederne har en dataetisk politik eller ej**

For at kunne undersøge hvor mange virksomheder der har forholdt sig til dataetik, er GPT-4 blevet anvendt. GPT-4 er blevet bedt om at kategorisere, hvorvidt virksomhedernes redegørelser for dataetik nedskriver eller linker til en dataetisk politik eller dataetiske retningslinjer. Hvis en virksomhed på ingen måde nævner dataetik i deres årsrapportering, så kategoriseres de som, at de ikke har redegjort for dataetik.

---

<sup>5</sup> Modellens temperatur var sat til 0,1. For uddybning se Open AI API dokumentation: <https://platform.openai.com/docs/api-reference/audio>.

For at klassificere hvorvidt virksomheder har redegjort for dataetik ved ikke at have en dataetisk politik og i stedet begrunde dette, har GPT-4 analyseret, hvorvidt virksomhederne eksplicit har nedskrevet, at de ikke har en dataetisk politik.

Det er ikke specificeret i årsregnskabsloven, hvordan virksomhederne skal begrunde, hvorfor de ikke har en dataetisk politik. På denne baggrund har Digitaliseringsstyrelsen heller ikke valgt at vurdere, hvorvidt en begrundelse var 'fyldestgørende'.

## **2.5 Udvælgelse af kategorier**

For at kunne udforme et prompt som sprogmodellen kan kategorisere ud fra, har vi udformet en række temaer og kategorier, som vi fandt relevante i denne undersøgelse. Temaerne og kategorierne er valgt ud fra en kombination af tidligere undersøgelser (Erhvervsstyrelsen, 2022; Erhvervsstyrelsen & Rambøll Management Consulting, 2020), vejledning og litteratur om dataetik<sup>6</sup>, samt en fælles manuel læsning af de dataetiske redegørelser i årsrapporterne.

De endelige kategorier anvendt i denne undersøgelse er valgt ud fra en kombination af formodninger og analyser. Det er blevet vurderet, om virksomhedernes dataetiske redegørelser reelt indeholdte samme temaer og kategoriseringer som først antaget.

Derudover har vi løbende valideret kategoriseringer ved hjælp af datasættet og promptet, da det var muligt at identificere hvilke kategorier, der var for upræcise til at sprogmodellen i kombination med promptet kunne anvende dem til at analysere datasættet. Disse kategorier blev sorteret fra.

Dermed er det heller ikke alle de kategoriseringer, som i første omgang blev inddraget, som er endt med at være en del af analysen. De valgte kategorier skal heller ikke ses som udtømmende for datasættets muligheder, men de skal ses som værende givende og relevante inden for det valgte scope.

## **2.6 Validering af GPT-4 og prompts**

Ved anvendelsen af generativ AI til analyse af data er det vigtigt at validere, hvorvidt den performer acceptabelt. Det skal ligeledes evalueres, hvorvidt de fejl, der opstår, kan løses eller er acceptable inden for en given margin. I valideringen af GPT-4's performance er det valgt at manuelt kode en stikprøve, for så at sammenligne det med GPT-4's kodning af det samme stikprøve.

Processen har været iterativ, dvs. at det i udformningen af kategorier og prompts løbende er blevet vurderet, hvorvidt GPT-4 performer efter hensigten. Dette er sket ved, at kategorier er blevet genovervejet, og prompts er blevet udviklet, således at de til sidst performede hensigtsmæssigt, og at prompten i langt størstedelen af tilfældene gjorde, at GPT-4 performede korrekt og i overensstemmelse med den manuelle kodning af samme stikprøve.

Slutteligt blev der udført en endelig validering af en ny stikprøve på 100 tilfældigt udvalgte dataetiske afsnit i årsrapporterne. Disse 100 tilfældigt udvalgte dataetiske afsnit er manuelt kodet på hver

---

<sup>6</sup> Håndbog i dataetik 2023 af Birgitte Kofoed Olsen, Virksomhedsguiden.dk, samt Dataetisk råd.

kategori ud fra en definition af hver kategori. Her er hvert dataetiske afsnit i årsrapporterne manuelt kodet for alle de respektive kategorier.

En stikprøve på 100 er valgt på baggrund af:

- Vi har løbende tjekket, valideret og forbedret prompten, således at vi kunne forvente, at en stikprøve på 100 højst sandsynligt ville give et retmæssigt billede af performancen
- Det høje antal af kategorier, de dataetiske politikkers længde, samt en ikke-udtømmende mængde af menneskelige ressourcer har gjort, at en stikprøve på 100 var realistisk, uden at være mangelfuld.

Når et datasæt indeholder tekst, kan der opstå situationer, hvor vurderingen af en kategoris tilstedeværelse er op til fortolkning. For at imødekomme dette, blev det derfor valgt at splitte den manuelle kodning på 100 stikprøver op, så flere medarbejdere kodede forskellige udsnit af data. Der hvor medarbejderne var i tvivl, blev det i fællesskab diskuteret og besluttet, hvorvidt en kategori var til stede eller ej.

Efter den manuelle kodning er stikprøven blevet sammenlignet med GPT-4's kodning af samme stikprøve. Her er der blevet sammenlignet hvorvidt:

- Den positive kodning er korrekt (True positive)
- Den positive kodning er forkert (False positive)
- Den negative kodning er korrekt (True negative)
- Den negative kodning er forkert (False negative)

De ovenstående kodninger giver et billede af, hvordan GPT-4 performer på tværs af alle kategorier. For bedre at kunne få et overblik over resultatet af stikprøven er der blevet udregnet tre metrikker, der alle giver et billede af en models klassifikationsperformance. Præcision (Precision), genkaldelse (Recall) og F1-score er metrikker, der anvendes til at vurdere ydeevnen af en klassifikationsmodel såsom en algoritme – i vores kontekst GPT 4 kombineret med det skrevne prompt som nævnt i afsnit 2.3 (selve promptet kan findes i afsnit 3).

Der er mange måder at validere en sprogmodel på, og vi haft valgt at tage inspiration i en metode, som lægger sig op ad, hvordan man også kan validere machine learning modeller (Hovy, 2020; Hovy 2022). Det bemærkes dog, at valideringen er tilpasset til netop vores valgte metode.

Der findes ikke et gyldent svar på, hvor høj scores for klassifikationsperformance bør være, da det afhænger af datasættet, målet og inputtets omfang, samt hvilke fejl der kan opstå. Derfor er der, selvom en kategori performede godt, løbende blevet kigget på, om der var områder, hvor den kunne blive bedre.

Ved dårligt-performende kategorier er det blevet evalueret, hvorfor kategorien ikke performer tilstrækkeligt, og her er prompten blevet reevalueret løbende. Til sidst er det derudover blevet besluttet, at visse kategorier skulle udgå på baggrund af dårlig performance i alle stikprøver – det er ikke været muligt inden for den valgte model at analysere de respektive kategorier, og det har ikke været ansvarligt at benytte resultaterne fra disse kategorier til videre analyse.

De følgende metrikker er udregnet per kategori, og hvert enkelt resultat er evalueret og godkendt separat.

*Præcisionsscore (Precision):*

Præcision måler, hvor mange af de positive instanser, der blev identificeret af modellen, der rent faktisk er korrekte. En høj præcision indikerer, at modellen har en lav rate af forkerte positive, og modellen derfor er god til at identificere de positive instanser. Modsat måler præcisionsscoren ikke, hvorvidt modellen performer godt på negative instanser, hvilket er essentielt i forhold til at evaluere, om resultaterne i de respektive kategorier er kunstigt høje eller lave (Hovy, 2022).

Formel:

$$\text{Precision Score} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

*Genkaldelsesscore (Recall):*

Genkaldelse måler, hvor mange af de faktiske positive instanser, der blev identificeret af modellen. Det vil sige, hvor mange af de reelt positive datapunkter, der også blev opfanget af modellen som værende positive. En høj genkaldelse indikerer, at modellen har et lavt antal forkerte negative, og dermed er god til at fange alle de reelt positive instanser (Hovy 2022).

Formel:

$$\text{Recall Score} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

*F1-score:*

F1 er et harmonisk gennemsnit af præcision og genkaldelse og giver en balance mellem de to metrikker – de positive og de negative instanser. F1-score er især anvendelig, når der er en ubalance mellem fordelingerne af positive og negative, hvilket er relevant i denne kontekst. Nogle af de analyserede kategorier indeholder en ulige fordeling af positive og negative, hvilket de nævnte valideringsscores ikke i samme grad tager højde for (Hovy, 2022).

Formel:

$$\text{F1 Score} = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Ovenstående valideringsscores vurderes én efter én. Det samlede resultat af hver kategori og hver score kan findes i afsnit 4. Undervejs i processen blev det tydeligt, at visse kategorier, som fremgår meget få gange i datasættet, var udfordrende at validere med den valgte valideringsmetode. Dette er eksempelvis kategorierne: 'Virksomheden begrundet at de ikke har en dataetisk politik fordi de kun anvender medarbejderdata'. I stikprøven fremgik kategorien kun én gang som positiv, men dog heller ingen gange som forkert negativt. Derfor blev det besluttet, at det var vigtigt, at alle kategorier blev udregnet for sig.

Alle kategorier med tre eller færre kodede positive instanser blev valideret yderligere. Her bestod valideringen i at manuelt kode 50 dataetiske politikker, som GPT-4 havde kodet som *positive* i den respektive kategori. Dermed fik vi et overblik over, om kategorierne synes større end de egentlig var. Hver præcisionsscore for de 'små' kategorier kan findes i afsnit 4.

## 2.7 Statistisk usikkerhed

Ved brugen af GPT-4 til analyse af tekst er det nødvendigt at adressere den statistiske usikkerhed, der er forbundet med brugen af generativ AI. GPT-4 kan generere tekstoutput og analysere tekst effektivt, men det er imidlertid afgørende at understrege, at modellen ikke besidder en internaliseret forståelse af 'sandhed' eller nuance på samme niveau som en menneskelig analytiker.

Statistisk usikkerhed kan ses på flere niveauer, herunder potentielt forkert genereret information og begrænsninger i håndteringen af data. Desuden kan modellen reagere varierende på lignende input afhængigt af subtile forskelle i formuleringen af promptet, som er skrevet af mennesker, og som den udelukkende analyserer ud fra. Disse aspekter indikerer nødvendigheden af en nøje fortolkning af resultater samt en bevidsthed om, at genererede svar ikke nødvendigvis reflekterer objektive kendsgerninger.

Derudover baserer undersøgelsen sig udelukkende på nedskrevne dataetiske redegørelser i virksomhedernes årsrapportering. Uddybende forklaringer, interne eller ikke-offentliggjorte dataetiske retningslinjer eller politikker er ikke medtaget.

## 3. Promptet

Bemærk at følgende prompt er udviklet og anvendt i november 2023. Det afspejler dermed den viden, der var tilgængelig, og de funktioner, som GPT-4 havde på daværende tidspunkt.

Prompt:

*You will be presented with a section from a company's annual report in parsed text from an xml file. Your task is to identify and analyze any paragraphs that relate to data ethics based on the following instructions. You can ONLY answer via the provided JSON keys I give you. Not ANYTHING else.*

*Approach this task step-by-step. Each step will end with a JSON-key format with allowable answers for the step. This is the only format you can answer in. Take your time and do not skip steps unless directly instructed to. It is important that your response have the exact number of JSON-keys as there are steps.*

1. Read the provided section of the annual report. Identify the paragraph that is regarding the company's data ethics policy. I only want you to extract the section about data ethics, but you must secure that you get all of that section. It is important, that you follow my following instructions clearly: 1. only answer with the extract from the text. You should not use ANY words yourself. Only use sentences from the passage I give you. 2. Only extract meaningful paragraph(s) or sentence(s) from the report. Make sure you include everything that a company writes about data ethics. 3. Ignore everything that relates to gender equality or social responsibility in general. This has nothing to do with data ethics, so NEVER include this in the extract. 4. Don't look at numbers. Ignore all numbers except for years and the number "99d" which might be a reference to data ethics. 5. be sure that you get the entire context the company writes regarding data ethics. 7. The company might state that their policy for data ethics can be downloaded or found at another source. If they write this, remember to include that reference. Remember all these 7 instructions. JSON-key: data\_ethics\_paragraph: contains the direct extract of the paragraph regarding data ethics policy. If there is no such paragraph, value should be 'NA'. If this step resulted in NA, then also put 'NA' in the following steps. Continue to step 2.

2. Now focus on the paragraph about data ethics, which you have identified in step 1. Determine whether the company or their mother company / their main enterprise (koncern) DO have a data ethics policy or guidelines or if they state that, they explicitly do NOT have a data ethics policy. Return 'Yes' in all instances, where either the company, or a mother company or main enterprise HAS a data ethics policy, even if the company e.g. writes "we do not have a policy, we refer to our mother company" it should be a 'Yes'. JSON-key: do\_they\_have\_a\_policy: allowable values are 'Yes' or 'No' or 'NA'. 'NA' is only possible if they have 'NA' in step 1.

3. Now you should analyze the instances, where you returned 'No' in step 2. You should determine the reason(s) that the company does not have a data ethics policy based on their own statement. A company can have several reasons for not having a data ethics policy. I will now provide you with these reasons, and you should provide a JSON key for each reason. They are not mutually exclusive, so you can put 'yes' in several of the json-keys.

3a. Analyze whether the company mentions data analysis, ai, algorithms, or advanced it-technologies or something similar to this. Always return "Yes" if this is the case. Remember, you can put 'Yes' even if other things are also mentioned. JSON-key: reason\_for\_not\_having\_data\_ethics\_policy\_algorithms: allowable values are: "Yes" or "No" or "NA". 'NA' will be true for ALL the cases where there was either an 'NA' or a 'Yes' in step 2.

3b. Analyze whether the company mentions mentions gdpr or something similar to what is contained in the gdpr rules e.g. personal data / databeskyttelse / data protection / data fortrolighed or anything similar to this. Always return "Yes" if this is the case. Remember, you can put 'Yes' even if other things are also mentioned. JSON-key: reason\_for\_not\_having\_data\_ethics\_policy\_gdpr: allowable values are: "Yes" or "No" or "NA". 'NA' will be true for ALL the cases where there was either an 'NA' or a 'Yes' in step 2.

3c. Analyze whether the company mentions only having employee data as a data point in their company as the reason or one of the reasons for not having a data ethics policy. Remember, you can put

'Yes' even if other things are also mentioned. JSON-key: `reason_for_not_having_data_ethics_policy_employee_data_only`: allowable values are: "Yes" or "No" or "NA". 'NA' will be true for ALL the cases where there was either an 'NA' or a 'Yes' in step 2.

3d. Analyze whether the company has well-thought data ethical consideration even though they state that they do not have a data ethics policy. Remember, you can put 'Yes' even if other things are also mentioned. JSON-key: `reason_for_not_having_data_ethics_policy_dataethics_consideration_no_policy`: allowable values are: "Yes" or "No" or "NA". 'NA' will be true for ALL the cases where there was either an 'NA' or a 'Yes' in step 2.

3e. Summarize the reason that the company does not have a data ethics policy. You can both type reason(s) that have already mentioned in steps 3a-3d, or you can type reason(s) that have not been mentioned in these steps, you decide freely yourself. The option in [] indicates that you should write some headlines yourself, but only within the limit of what is defined in the []. You can only return 'No' if there is a 'No' in all of steps 3a-3d. JSON-key: `reason_for_not_having_data_ethics_policy_any_reason`: allowable values are: [Give headlines on the type of reason, the company gives. It can be any kind, you are very free to judge yourself, no answer is wrong. Separate headlines by |. Headlines should be maximum 3 words. Return your answer in the same format as all other JSON-keys] or 'No' or 'NA'. 'NA' will be true for ALL the cases where there was either an 'NA' or a 'Yes' in step 2.

4. Now you should analyze the instances where you returned 'Yes' in step 2. Look at the paragraph and determine whether the paragraph explicitly describes the company's data ethics policy or if it refers to another place, where their data ethical policy can be found. This can e.g. be if they have a mother company and they refer to their mother company instead, or it can be if they have a data ethical policy at another link or another report. If they do write some of their data ethics policy, and then in the end refers to a report, where you can read MORE, then it counts as though they have their own policy, but that more can be read. JSON-key: `do_they_explicitly_describe_data_ethics_policy`: allowable values are 'Yes, they describe their data ethics policy in full' or 'Yes, they describe most of it, but then refer to another place where you can read the rest' or 'No they do not describe it, but they refer to their mother companies' policy' or 'No they do not describe it, but they refer to another report or link, where their data ethics policy can be found' or 'NA'. 'NA' is for all the cases where there is either an 'NA' or a 'No' in step 2.

4. Now you should analyze the instances where you returned 'Yes' in step 2. You should analyze whether the company's data ethics policy can be found directly in the passage or if they refer to another place, where it can be read or both. There are several possible scenarios, I will now provide you with each of these scenarios, and you should provide a JSON-key for each scenario. They are not mutually exclusive, so you can return yes in several of them.

4a. Analyze whether the company's data ethics policy – or at least some of it - can be found directly in text. This also includes if they write some of their policy and in the end refers to a website or another report where you can read more about it. JSON-key: `policy_in_report`: allowable values are 'Yes' or 'No' or 'NA'. 'NA' will be true in all the cases where there is either an 'NA' or a 'No' in step 2.

4b. Analyze either whether the company mentions that their policy can be found, at their website, at their csr report (or other social report) or if they put a direct link to where it can be found. This can

also be 'Yes', if they mention some of the policy in this passage and still mention another place where it can be found. JSON-key: `policy_refer_to_own_policy_elsewhere`: allowable values are 'Yes' or 'No' or 'NA'. 'NA' will be true in all the cases where there is either an 'NA' or a 'No' in step 2.

4c. analyze whether the company mentions words that could indicate that the company is part of a company structure, where different companies share owners. For example, this is the case for words such as 'mother company', 'parent company', 'main enterprise', 'koncern' and similar. Look at the text and decide if there is some indication of the company being in this kind of structure. JSON-key: `policy_refer_to_company_structure`: allowable values are 'Yes' or 'No' or 'NA'. 'NA' will be true in all the cases where there is either an 'NA' or a 'No' in step 2.

5. Now you should analyze the instances where you returned 'Yes' in step 2 AND one of the two possible 'yes' answers in step 4. Alright, so now you are looking at an annual report that has described the company's data ethics policy. Now you should analyze what types of data the company describes in their policy. JSON-key: `type_of_data_described`: allowable values are 'Personal data / persondata' or 'Health data / sundhedsdata' or 'customer data / kundedata' or 'Employee data / medarbejderdata' or 'Economical data / økonomisk data' or 'None of the above' or 'NA'. 'NA' is for all the cases where there is either an 'NA' or one of the 'No'-answers in step 4.

5. Now I want to focus on reports, where the company has put at least SOME of their data ethics policy in the paragraph. Even if it just a little bit. This means, you should take a deep breath now, and go back to step 4a, where you decided whether at least some of the data ethics policy of the company can be found in the text. If you answered 'Yes' to that, then you MUST answer either 'Yes' or 'No' to the rest of the steps I provide. First, I ask about the data types.

5a: Is personal data (persondata) mentioned? JSON-key: `personal_data`: allowable values are 'Yes' or 'No'.

5b: Is employee data (medarbejderdata) or something similar mentioned? JSON-key: `employee_data`: allowable values are 'Yes' or 'No'.

5c: Is customer data (kundedata) or something similar mentioned? JSON-key: `customer_data`: allowable values are 'Yes' or 'No'.

5d: Is health data (sundhedsdata) or something similar mentioned? JSON-key: `health_data`: allowable values are 'Yes' or 'No'.

5e: Is company data (virksomhedsdata) or something similar mentioned? JSON-key: `company_data`: allowable values are 'Yes' or 'No'.

5f: Is economic data (økonomisk data) or something similar mentioned? JSON-key: `economic_data`: allowable values are 'Yes' or 'No'.

5g: Is any other type of data mentioned? The option in [] indicates that you should write the types of data yourself, but only within the limit of what is defined in the []. You can only return 'No' if there are absolutely NO type of data other than already mentioned. It is better that you come with too

many suggestions rather than too few. JSON-key: `any_other_data`: allowable values are: [Give the type of data in a headline. It can be any kind, you are totally free to judge yourself, no answer is wrong. Separate headlines by |. Headlines should be maximum 3 words. Return your answer in the same format as all other JSON-keys] or 'No'.

6. Now you should analyze the instances where you returned 'Yes' in step 4a. Now you should analyze what type of data ethics understanding the company has in its' described data ethics policy.

6a. Does the company understand data ethics in their data ethics policy as the same or similar to GDPR-rules. Look at the company's data ethics policy, and then think about what you know about the gdpr-rules. Return a 'Yes' if the company in their data ethics policy writes about stuff that is similar to gdpr compliance, such as data protection or stuff like that. JSON-key: `same_as_gdpr`: allowable values are 'Yes' or 'No'.

6b. Now you should focus on whether the company understand data ethics as something they need because they apply data analysis, AI, algorithms or advanced IT technologies. JSON-key: `need_bc_AI_algorithms`: allowable values are 'Yes' or 'No'.

6c. Now you should focus on whether the company understand data ethics as something they need because of social responsibility. Examples could be if they mention any of these as their understanding of needing a data ethics policy: 'Digital ansvarlighed', 'socialt ansvar', 'samfundsansvar', 'ansvar for befolkningen'. JSON-key: `need_bc_society_responsibility`: allowable values are 'Yes' or 'No'.

6d. Now you should only focus on whether the company mentions that they understand data ethics as something they need because of cloud computing in any way. That could be that they have all different kinds of data in the cloud ('skyen') or that they have taken some kind of decision regarding using cloud computing because of data ethics. JSON-key: `need_bc_cloud`: allowable values are 'Yes' or 'No'.

6e. Now you should only focus on whether the company mentions that they understand data ethics as something they need because of the risk of hacking/phishing/ransomware. JSON-key: `need_bc_hacking`: allowable values are 'Yes' or 'No'.

6f. Is any other type of understanding that the company has, on why they need a data ethics policy? Other than the ones mentioned in 6a-6e. The option in [] indicates that you should write a headline for other understandings yourself, but only within the limit of what is defined in the []. You can only return 'No' if there are absolutely NO other type of understanding of why you need data ethics, than already mentioned. It is better that you come with too many suggestions rather than too few. JSON-key: `any_other_understanding`: allowable values are: [Give the type of understanding in a headline. It can be any kind, you are totally free to judge yourself, no answer is wrong. Separate headlines by |. Headlines should be maximum 3 words. Return your answer in the same format as all other JSON-keys] or 'No'.

7. Now you should analyze the instances where you returned 'Yes' in step 4a. Analyze which objectives or measurements that the company has actively taken in other to secure that they adhere to their data ethics policy.

7a. Now you should focus on whether the company mention that they have decided to not sell some – or all – of their data to third parties. JSON-key: `no_data_sales`: allowable values are ‘Yes’ or ‘No’.

7b. Now you should only focus on whether the company has taken active measures in order to not collect or store specific kind of data because of data ethics considerations. JSON-key: `data_decisions`: allowable values are ‘Yes’ or ‘No’.

7c. Now you should only focus on whether the company has introduced voluntary education or training about data ethics or data security to their employees. JSON-key: `education_voluntary`: allowable values are ‘Yes’ or ‘No’.

7d. Now you should only focus on whether the company has introduced obligatory education or training about data ethics or data security to their employees. JSON-key: `education_obligatory`: allowable values are ‘Yes’ or ‘No’.

7e: Now you should only focus on whether the company has introduced education or training about data ethics that is neither specified to be obligatory or voluntary to their employees. Always return ‘No’ if step 7c OR step 7d returned ‘Yes’. JSON-key: `education_general` allowable values are ‘Yes’ or ‘No’.

Only answer in JSON format. Remember you need to answer with one JSON key per step I provided. This means you should answer with exactly 29 JSON keys. It is important that you answer one of the allowable values for each JSON-key. In the extract, you ONLY include words directly from the annual report.

## 4. Valideringsresultater

Nedenstående resultater afspejler udregningerne fra metodeafsnit 2.6. Det stræbes efter at få et resultat så tæt på 1.00 som muligt.

Spørgsmål:	Precision	Accuracy	Recall	F1 Score	With NA: Precision	With NA: Accuracy	With NA: Recall	With NA: F1 Score
<code>do_they_have_a_policy</code>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
<code>reason_for_not_having_data_ethics_policy_algor...</code>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
<code>reason_for_not_having_data_ethics_policy_gdpr</code>	1,00	0,99	0,94	0,97	1,00	0,99	0,94	0,97
<code>reason_for_not_having_data_ethics_policy_emplo...</code>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
<code>reason_for_not_having_data_ethics_policy_datae...</code>	0,95	0,98	0,95	0,95	0,95	0,98	0,95	0,95

Spørgsmål:	Precision	Accuracy	Recall	F1 Score	With NA: Precision	With NA: Accuracy	With NA: Recall	With NA: F1 Score
policy_in_report	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
policy_refer_to_own_policy_elsewhere	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
policy_refer_to_company_structure	1,00	0,96	0,86	0,92	1,00	0,97	0,86	0,92
same_as_gdpr	0,91	0,92	1,00	0,96	0,91	0,96	0,97	0,94
need_bc_AI_algorithms	1,00	1,00	1,00	1,00	1,00	0,99	0,83	0,91
no_data_sales	1,00	1,00	1,00	1,00	1,00	0,98	0,82	0,90
data_decisions	1,00	0,92	0,70	0,82	1,00	0,95	0,58	0,74
education_voluntary	1,00	1,00	1,00	1,00	1,00	0,98	0,33	0,50
education_obligatory	1,00	1,00	1,00	1,00	1,00	0,98	0,50	0,67
education_general	1,00	0,97	0,86	0,92	1,00	0,97	0,67	0,80

## 5. Referenceliste

Erhvervsstyrelsen. (2022). *Store danske virksomheders arbejde med dataetik*. <https://digst.dk/media/28829/store-danske-virksomheders-arbejde-med-dataetik.pdf>

Erhvervsstyrelsen, & Rambøll Management Consulting. (2020). *Danske virksomheders arbejde med dataetik*. <https://digst.dk/media/28828/danske-virksomheders-arbejde-med-dataetik-rapport-fra-ramboell.pdf>

Hovy, D. (2020). *Text Analysis in Python for Social Scientists: Discovery and Exploration* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781108873352>

Hovy, D. (2022). *Text Analysis in Python for Social Scientists: Prediction and Classification* (1st ed.).