

12. september 2023

Ref. CNE

J.nr. 23-007824

## Hørings svar om god praksis for dataetik ved brug af AI i den finansielle sektor

**Følgende myndigheder og organisationer har afgivet høringssvar uden bemærkninger:**

Ingen

**Følgende myndigheder og organisationer har haft bemærkninger til lovforslaget:**

DAFINA

Transparency International Danmark

Finanssektorens Arbejdsgiverforening

Danske Advokater

Dansk Arbejdsgiverforening

Forsikring & Pension

Den Danske Aktuarforening

Den danske Fondsmæglerforening

Dataetisk Råd

Finans Danmark

Digitaliseringsstyrelsen

Nasdaq

FSR – danske revisorer

Dansk Erhverv

Forbrugerrådet Tænk

Finansforbundet

Finansforbundet, Dansk Industri (DI Digital) og Copenhagen Fintech

Datatilsynet

**FINANSTILSYNET**  
Strandgade 29  
1401 København K

Tlf. 33 55 82 82  
CVR-nr. 10 59 81 84  
finansstilsynet@ftnet.dk  
www.finanstilsynet.dk

**ERHVERVSMINISTERIET**

Læs høringssvar med bemærkninger på den næste sider.

**From:** Michael Camphausen <mca@camphausen.dk>  
**Sent:** 01-06-2023 11:31:35 (UTC +02)  
**To:** Camilla Neuenschwander (FT) <CNE@FTNET.DK>  
**Subject:** DAFINA høringsvar vedrørende udkast til god praksis papir om dataetik ved brug af AI i den finansielle sektor

Kære Camilla

På vegne af Dansk Fintech Alliance (DAFINA) fremsendes hermed nedenstående høringsvar vedrørende Finanstilsynets ovennævnte udkast til god praksis papir om dataetik ved brug af AI i den finansielle sektor. DAFINA takker på denne vis for muligheden for at afgive nærværende høringsvar og skal således fremkomme med følgende bemærkninger, input og opfordringer i forhold til tilsynets udkast:

Der er efter vores opfattelse tale om et relevant og gennearbejdet god praksis papir om dataetik, som vi fuldt ud støtter. Vi er således enige i, at der i offentligheden skal være tillid til, at de finansielle aktører og udbydere, herunder eksempelvis fintech virksomheder med nye innovative, digitale løsninger på betalingsområdet og andre finansielle områder, behandler data på en ordentlig og etisk forsvarlig måde. Vi mener dog umiddelbart, at dette ikke kun er relevant i forhold til anvendelsen af nye teknologier (fx AI) til databehandlingen, men at det således har relevans mere generelt i forhold til udviklingen og udbuddet af nye digitale, finansielle løsninger og produkter, dvs. i grunden uafhængigt af, om der bruges nye underliggende teknologier hertil.

Vi bemærker således, at tilsynets sammenfatning, eksempler, overvejelser og de enkelte principper i god praksis papiret både formuleringsmæssigt og indholdsmæssigt sigter til dataetik i mere bred forstand og på denne vis ikke kun er begrænset til anvendelsen af AI i den finansielle sektor. Derfor bør det overvejes at præcisere i god praksis papiret, at de pågældende overvejelser og principper om dataetik rækker videre end til brugen af AI på det finansielle område, dvs. at de enkelte principper har relevans mere generelt i forhold til nye digitale løsninger på det finansielle område (AI er med andre ord blot et eksempel herpå, om end naturligvis et vigtigt eksempel).

Netop fordi tilsynets eksempler, overvejelser og principper om dataetik er formuleret mere generelt og således ikke er begrænset til AI, skal vi opfordre tilsynet til at overveje, om det allerede i god praksis papirets titel og under alle omstændigheder i den indledende afgrænsning i papiret skal tydeliggøres/ændres, at papiret rækker videre end AI og i grunden omfatter både nye teknologier og nye digitale løsninger i den finansielle sektor baseret på brugen og anvendelsen af data mv. (en mere retvisende titel indholdsmæssigt overensstemmende med selve indholdet undervejs kunne således være "Dataetik ved nye digitale løsninger og brugen af AI i den finansielle sektor").

Formålet hermed og vores bevæggrund herfor skal ses i sammenhæng med vores generelle opfattelse, hvad angår reguleringen af anvendelsen og behandlingen af data på det finansielle område, herunder eksempelvis betalingsområdet. Vi har således den grundlæggende opfattelse, som også kommer til udtryk i flere af vores tidligere høringsvar mv. til tilsynet, at der efter indførelsen af GDPR ikke er behov for særskilt dansk overregulering eller supplerende finansiell regulering af behandlingen af data, herunder finansielle data, dvs. ikke er behov for særskilt dansk finansiell regulering på området, som indebærer hårdere, mere indgribende regulering – og i forhold til udenlandske finansielle aktører også mere konkurrenceforvridende regulering – end hvad følger af netop GDPR. Og dette især ikke, hvis GDPR på det finansielle område nyttigt suppleres af et sådant dataetisk god praksis papir på det finansielle område, dvs. GDPR har bred anvendelse, herunder på det finansielle område mere generelt, og god praksis papiret har tilsvarende bred

anvendelse, herunder på det finansielle område mere generelt (generelle digitale finansielle løsninger, ikke kun ved brugen af AI) – med andre ord vil GDPR (regulering) og god praksis papiret (ikke-regulering) på denne vis følges ad, når det handler om databehandling og dataetik på det finansielle område i lyset af digitaliseringen og nye innovative finansielle løsninger.

Dermed burde det ikke være nødvendigt efterfølgende at opretholde eksempelvis den danske særregel i BET § 125, dvs. det delvise forbud mod kommercialisering/behandling af betalingsoplysninger, der ikke i forvejen følger af PSD2, og der som bekendt begrænser danske kontooplysningstjenesteudbydere (AISP) / danske AISP-løsninger sammenholdt med udenlandske af slagsen. Ligeledes som bekendt er EU Kommissionen i denne måned på vej med sit udkast til PSD3 (bl.a. mht. Open Banking) og sit samtidige udkast til et regulatorisk rammeværk for Open Finance, og netop i denne henseende vil det være hensigtsmæssigt, at vi herhjemme kunne arbejde for dansk implementeringsregulering, der ikke udvander eller begrænser muligheden for danske Open Banking og Open Finance løsninger, idet den nødvendige datamæssige forbrugerbeskyttelse i så henseende varetages af GDPR "toppet up" med et dansk god praksis papir om dataetik ved digitale finansielle løsninger, dvs. i stedet for og som alternativ til dansk særregulering på området ovenpå GDPR og på PSD3 mv. Et mere bredt god praksis papir om dataetik på det finansielle område ved nye digitale løsninger (og altså ikke blot ved AI) vil med andre ord være til gavn for en kommende implementering af PSD3 og Open Finance regelsættene herhjemme uden behovet for opretholdelsen eller indførelsen af danske særforbud/særregulering på området ovenpå GDPR.

Såfremt ovenstående giver anledning til bemærkninger eller spørgsmål fra Finanstilsynets side, står vi naturligvis til fortsat rådighed.

Med venlig hilsen

**Dansk Fintech Alliance (DAFINA)**

---

**Fra:** Camilla Neuenschwander (FT) <CNE@FTNET.DK>

**Sendt:** 17. maj 2023 14:01

**Til:** Fintech <Fintech@FTNET.DK>

**Emne:** Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Til rette vedkommende

Hermed sender Finanstilsynet udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor i høring.

Finanstilsynets hensigt med god praksis-papiret er ikke at indføre nye forpligtelser for eller krav til virksomhederne. Baggrunden for udarbejdelsen af papiret fremgår i øvrigt af vedhæftede høringsbrev.

Finanstilsynet beder om eventuelle bemærkninger til udkastet senest den 27. juni 2023. Bemærkninger kan sendes pr. e-mail på [cne@ftnet.dk](mailto:cne@ftnet.dk) eller pr. post til den i høringsbrevet nævnte adresse.

Med venlig hilsen

**Camilla Neuenschwander**

Chefkonsulent, cand.jur.

Kontor for Fintech, Betalingstjenester og Governance

---



Strandgade 29, 1401 København K  
Tlf.: +45 33 55 82 82  
Direkte tlf.: +45 33 55 82 12  
<mailto:cne@ftnet.dk>  
[www.finanstilsynet.dk](http://www.finanstilsynet.dk)

---

Finanstilsynet er ansvarlig for behandlingen af de personoplysninger, vi modtager om dig. Du kan læse mere om, hvordan vi behandler dine personoplysninger på vores hjemmeside <https://www.finanstilsynet.dk/Kontakt/Privatlivspolitik>

Finanstilsynet gør opmærksom på, at denne e-mail og eventuelle vedhæftede filer er fortrolige. Hvis du har modtaget denne mail ved en fejl, bedes du straks oplyse Finanstilsynet herom ved at besvare denne e-mail og derefter slette e-mailen. Vi gør opmærksom på, at hvis du har modtaget denne e-mail ved en fejl, kan enhver form for kopiering, offentliggørelse eller distribution af denne e-mail være ulovlig.

Finanstilsynet

2. juni 2023

Sendt via [CNE@ftnet.dk](mailto:CNE@ftnet.dk)

**J.nr. 23-007824**

**Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor**

Vi takker for lejligheden til at kommentere udkastet om god praksis.

Artificial Intelligence teknologi påvirker hverdagen mere og mere, også i brancher som den finansielle sektor. Magt flyttes fra mennesker til maskiner. Etisk brug af data er afgørende, når denne magt skal forvaltes. Vi er meget positive overfor jeres tiltag om at sætte standarder og at dele god praksis.

Transparency International ser på AI fra to vinkler: Dels hvordan vi selv kan anvende AI systemer i vores eget analyse arbejde, og dels de udfordringer teknologien fører med sig indenfor åbenhed og korruption. Vore kommentarer er med henblik på disse udfordringer.

Uetisk brug af data kan ske utilsigtet (f.eks. bias ved uheld eller sjusk), helt bevidst (f.eks. prisvandring i eksempel F i udkastet) eller som misbrug for egen vinding skyld (egentlig korruption). Udkastet fokuserer mest på de to første scenarer, vi savner en holdning til et egentlig korrupt og dermed uetisk misbrug.

Tre **illustrative, arketyperiske** eksempler:

- En afdelingsleder bestikker en IT udvikler til at indbygge en bias i Robotic Process Automation algoritmen, der fordeler låneansøgninger, således at netop lederens afdeling får tildelt de ansøgninger, der er lettest at behandle, og så han opnår effektivitetsmål og bonus.
- En algoritme, der udvælger jobansøgninger til interview, har en unfair bias for at udvælge ansøgninger fra kristne, hvide mænd og understøtter dermed IT udviklerens diskriminerende politiske dagsorden.
- En sagsbehandler har adgang til at manipulere tekst og fotos fra skadesanmeldelser og påvirker dermed om AI systemet udtager skadesanmeldelsen til kontrol eller ej. Sagsbehandleren sælger denne 'service' for tjenesteydelser eller kontanter.

Se eventuelt vores generelle kortlægning i 2022.

<https://www.transparency.org/en/blog/bribes-for-bias-can-ai-be-corrupted>

Vi foreslår, at I tilføjer et afsnit om risikovurdering og kontrol med specielt fokus på bevidst misbrug af AI systemerne ved udnyttelse af betroet adgang.

Uigennemskueligheden af AI algoritmer i drift gør det svært at afsløre et misbrug, og skaleringen af automatisering forstærker effekten. De tekniske metoder til at afsløre og modvirke bevidst eller ubevidst uetisk databrug, eller egentligt misbrug, er langt hen ad vejen de samme.

Vi foreslår, at finansielle virksomheder, der ønsker at anvender AI på et etisk forsvarligt niveau, også risikovurderer udviklings- og driftsprocesserne for et egentlig korrump misbrug af folk med betroet adgang til systemerne for deres egen vindings skyld. Risiko vurderingen bør omfatte sårbarheder som f.eks.

- Udvælgelse, integration og mark-up af data til træning af AI systemerne har risici for data manipulation, der indbygger en skjult bias som f.eks. eksemplet ovenfor om diskriminering.
- Transaktionsdata der forberedes, før de behandles af AI algoritmerne. F.eks. eksemplet ovenfor om manipulering af fotos, eller rendering af f-eks- jobansøgninger.
- IT systemer, der eksekverer beslutninger baseret på AI algoritmernes sandsynlighedsberegninger, som f.eks. RPA eksemplet ovenfor.

Anbefalingen om risiko vurdering kan eventuelt udbygges med god praksis for opklaring og forebyggelse af bevidst misbrug som f.eks. funktions adskillelse mellem udvikling og kvalitets sikring, sikring af revisionsspor af input og output fra AI systemer i drift, og styrkelse af kontroller i forretningsprocesserne.

Transparency International's holdning til ansvarlig brug af AI teknologi er kort:

- Algoritmisk gennemsigtighed - Faktorer, der påvirker en beslutning, skal defineres og være synlige for de berørte mennesker.
- Algoritmisk ansvarlighed - Ejeren af AI algoritmen, skal stå til ansvar for resultatet.
- Algoritmisk reviderbarhed - AI systemet og data skal være tilgængelige for uafhængige verifikation.
- Algoritmisk regulering - Juridiske regler og retsmæssig håndhævelse

Vi står selvfølgelig til rådighed for yderligere spørgsmål og/eller uddybning, hvis I ønsker det.

Med venlig hilsen

Jesper Olsen

Formand

Lars Wriedt

Bestyrelsesmedlem

**From:** David Luxhøj-Pedersen <dlp@fanet.dk>  
**Sent:** 18-06-2023 17:51:40 (UTC +02)  
**To:** Camilla Neuenschwander (FT) <CNE@FTNET.DK>  
**Subject:** SV: Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Kære Camilla

Finanssektorens Arbejdsgiverforening (FA) takker for høringsmuligheden og har følgende bemærkninger.

FA har ved overenskomstforhandlingerne i 2023 aftalt at nedsætte arbejdsgrupper med både Finansforbundet og Forsikringsforbundet om dataetik generelt.

Arbejdsgruppen afslutter sit arbejde senest ved udgangen af september 2024 med en anbefaling til et protokollat eller kodeks om dataetik.

Arbejdsgruppen skal drøfte dataetik i sammenhæng med gældende lovgivning og den digitale udvikling, og sammen tilvejebringe information om udvikling af dataetik mv. til gavn for virksomheder og medarbejdere. I arbejdet kan bl.a. indgå drøftelser af elementer som:

- Transparens, herunder hvordan medarbejdere efter gældende regler kan få adgang til viden om, hvilke data og med hvilket formål data er indsamlet
- Procedurer for opfølgning, adgang til at gøre indsigelser og sletning af data efter gældende regler
- Saglige begrænsninger i medarbejdernes adgang til indsigt i de data om medarbejderne, virksomhederne behandler
- Tilgængelig information om ovennævnte

Arbejdsgruppen sammensættes af repræsentanter fra FA samt Finansforbundet og Forsikringsforbundet, som kan suppleres af eksperter, hvis parterne er enige.

Arbejdsgruppens resultater kan have betydningen for udkastets sidste halvdel (punkterne 8.1-8.4).

FA vil derfor gerne holdes orienteret og bidrag i det videre arbejde, når arbejdsgruppen har begyndt sit arbejde.

**Med venlig hilsen**

David Luxhøj-Pedersen  
Advokatfuldmægtig

E: [dlp@fanet.dk](mailto:dlp@fanet.dk)  
M: [+45 33 38 16 23](tel:+4533381623)

Amaliegade 7  
1256 København K

*OBS! Vores juridiske hotline har fået nyt nummer og e-mail*

T: [+45 33 38 16 10](tel:+4533381610)  
E: [jurahotline@fa.dk](mailto:jurahotline@fa.dk)



---

**Fra:** Camilla Neuenschwander (FT) <CNE@FTNET.DK>

**Sendt:** 17. maj 2023 14:01

**Til:** Fintech <Fintech@FTNET.DK>

**Emne:** Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Til rette vedkommende

Hermed sender Finanstilsynet udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor i høring.

Finanstilsynets hensigt med god praksis-papiret er ikke at indføre nye forpligtelser for eller krav til virksomhederne. Baggrunden for udarbejdelsen af papiret fremgår i øvrigt af vedhæftede høringsbrev.

Finanstilsynet beder om eventuelle bemærkninger til udkastet senest den 27. juni 2023. Bemærkninger kan sendes pr. e-mail på [cne@ftnet.dk](mailto:cne@ftnet.dk) eller pr. post til den i høringsbrevet nævnte adresse.

Med venlig hilsen

**Camilla Neuenschwander**

Chefkonsulent, cand.jur.  
Kontor for Fintech, Betalingstjenester og Governance

---

 **FINANSTILSYNET**

Strandgade 29, 1401 København K

Tlf.: +45 33 55 82 82

Direkte tlf.: +45 33 55 82 12

<mailto:cne@ftnet.dk>

[www.finanstilsynet.dk](http://www.finanstilsynet.dk)

---

Finanstilsynet er ansvarlig for behandlingen af de personoplysninger, vi modtager om dig. Du kan læse mere om, hvordan vi behandler dine personoplysninger på vores hjemmeside <https://www.finanstilsynet.dk/Kontakt/Privatlivspolitik>

Finanstilsynet gør opmærksom på, at denne e-mail og eventuelle vedhæftede filer er fortrolige. Hvis du har modtaget denne mail ved en fejl, bedes du straks oplyse Finanstilsynet herom ved at besvare denne e-mail og derefter slette e-mailen. Vi gør opmærksom på, at hvis du har modtaget denne e-mail ved en fejl, kan enhver form for kopiering, offentliggørelse eller distribution af denne e-mail være ulovlig.



Finanstilsynet  
Strandgade 29  
1401 København K

Sendt via mail til: [cne@ftnet.dk](mailto:cne@ftnet.dk)

Vesterbrogade 32  
1620 København V

Telefon 33 43 70 00  
[mail@danskeadvokater.dk](mailto:mail@danskeadvokater.dk)  
[www.danskeadvokater.dk](http://www.danskeadvokater.dk)

21. juni 2023

Dok.nr. D-2023-024500

## **Høringsvar vedrørende udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor**

Finanstilsynet har 17. maj 2023 sendt udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor i høring. Danske Advokater takker for muligheden for at afgive høringssvar. Dette høringssvar er udarbejdet af Danske Advokaters fagudvalg for IP-rettigheder, forbruger- og markedsføringsret.

### **Overordnede bemærkninger**

Danske Advokater støtter overordnet set udkastet til papiret, fordi det vil, hvis det offentliggøres, give virksomheder i den finansielle sektor værktøjer og inspiration til overvejelser, som de kan gøre sig, når og hvis de ønsker at arbejde med dataetik. På denne måde vil virksomhederne bedre kunne klarlægge og beskæftige sig med dataetik samtidig med, at virksomhedernes kunder vil få større indblik i, hvordan deres data bliver behandlet af de finansielle virksomheder. Dette vil øge kundernes retssikkerhed ved virksomhedernes brug af deres data.

### **Den dataetiske politik**

Danske Advokater anser den dataetiske politik for at være et godt og effektivt værktøj for at øge og sikre tilliden til det finansielle system i forhold til virksomhedernes brug af kundernes data.

Fordelene er todelt.

For det første, hvis virksomhederne udarbejder en dataetisk politik, vil de blive nødsaget til at gøre sig dataetiske overvejelser. For det andet, hvis denne dataetiske politik bliver gjort tilgængelig for kunderne, vil kunderne blive gjort bekendt med virksomhedens politik. Denne transparens fra virksomhedens side vil fremme kundens retssikkerhed og mindske "etik-washing", som omtalt på side 15.

Udarbejder virksomhederne ikke en dataetisk politik, falder kundernes retssikkerhed tilbage på den allerede eksisterende regulering og reglerne om god skik, hvorefter *“[e]n finansiel virksomhed skal handle redeligt og loyalt over for sine kunder”*.

Selvom denne grundsætning er med til at sikre kundernes retssikkerhed, betyder dens manglende transparens, at en dataetikpolitik trods alt vil sikre dette bedre.

### **Afsluttende bemærkninger**

Danske Advokater har ikke yderligere bemærkninger til udkastet til papiret på nuværende tidspunkt.

Danske Advokater takker for muligheden for at afgive høringssvar. Vi står gerne til rådighed for uddybende spørgsmål eller dialog.

Med venlig hilsen

Susanne Bager  
Juridisk konsulent  
Danske Advokater  
sub@danskeadvokater.dk

Henrik Saugmandsgaard Øe  
Advokat  
Medlem af fagudvalg for IP,  
forbruger- og markedsføringsret

FINANSTILSYNET  
Att.: Camilla Neuenschwander  
Strandgade 29

1401 København K  
Danmark

## Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

27. juni 2023  
KBT/SMOT  
Dok ID: 210129

Dansk Arbejdsgiverforening (DA) har den 17. maj 2023 modtaget Finanstilsynets Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor.

DA vil indledningsvis kvittere for modtagelse af høringen. DA finder det essentielt, at arbejdsmarkedets parter er involveret i arbejdet med brug af AI på det danske arbejdsmarked.

DA ser overordnet set positivt på tiltag der sigter mod at styrke virksomheders arbejde med dataetik.

DA genkender dog samtidigt en række virksomheders tilkendegivelser af, at det er vanskeligt at udarbejde én model for AI-etik, som passer til alle virksomheder. Derfor ser DA positivt på disse retningslinjers frivillige og fleksible karakter, fordi retningslinjerne herigennem rummer mulighed for, at virksomheder kan arbejde med dem i overensstemmelse med virksomhedernes konkrete behov. Denne fleksible tilgang danner i DA's optik gode forudsætninger for, at virksomheder kan bruge retningslinjerne til at indfri ønskværdige målsætninger uden at dette medfører unødigt byrdefulde omstændigheder, der kan være til skade for virksomheders konkurrenceevne.

DA er meget enig i punkt 8.1, der understreger vigtigheden af medarbejderinddragelse i arbejdet med at skabe en fælles sund kultur for behandling af data.

DA mener i tillæg dertil, at vedvarende dialog mellem ledelse og et bredt udvalg af medarbejdere med forskellige perspektiver på brug af AI er afgørende for succesfuld forankring af arbejde med dataetik i organisationer

Vester Voldgade 113  
1552 København V  
Tlf.: 33 38 90 00  
CVR 16834017  
da@da.dk  
da.dk

DA mener desuden, at dialog mellem ledelse og medarbejder er særligt nyttigt ifm. implementering af AI-løsninger, som Finanstilsynet også peger i punkt 8.4. DA vil i den sammenhæng pege på Samarbejdsudvalg som et velegnet forum til at drøfte spørgsmål om dataetik i relation til AI.

Med venlig hilsen  
DANSK ARBEJDSGIVERFORENING

Sara Margarite Otmishi / Kristian Bjørnskov Thormann

**F&P's høringsvar til høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor (Jnr. 23-007824)**

27.06.2023

Tak for muligheden for at kommentere på papiret om dataetik ved brug af AI i den finansielle sektor.

F&P  
Philip Heymans Allé 1  
2900 Hellerup  
Tlf.: 41 91 91 91  
fp@fogp.dk  
www.fogp.dk

F&P ser positivt på formålet med at understøtte finanssektorens arbejde med dataetik og hilser papiret velkomment. F&P ser overordnet, at papiret udgør et godt fundament for arbejdet med dataetik.

Jakob Holm  
Chefkonsulent  
Dir. 41919176  
jho@fogp.dk

Vi har en række input til papiret, som primært går på at skabe mere tydelighed og gøre papiret lettere at gå til for læseren. Derudover har vi nogle tekstnære forslag med input til de enkelte afsnit.

Sagsnr. GES-2023-00170  
DokID 465979

Hvis noget giver anledning til opfølgende spørgsmål, så står vi gerne til rådighed for at besvare disse.

Med venlig hilsen  
Sigrid Floor Toft & Jakob Holm

---000---

## **Overordnede input**

### **Tydelighed omkring formål:**

Vi ser det som definerende for dokumentet, at papiret som det fremgår på side 7:

- Alene er inspiration til overvejelser, som virksomhederne kan gøre sig i arbejdet med dataetik
- På ingen måde er bindende for virksomhederne

Vi vil foreslå, at dette også fremgår helt tydeligt i starten af dokumentet. Samtidig er der særligt på de første sider ordbrug, som har en mere anvisende/vejledende karakter. Disse bør konsekvensrettes, så det tydeligt fremgår, at der er tale om inspiration til arbejdet med dataetik og ikke en udvidelse af kravene til god skik.

I forlængelse vil vi foreslå, at papiret udgives under en anden overskrift; eksempelvis ”**gode råd** om dataetik ved brug af AI i den finansielle sektor”, og ikke under overskriften ”**god praksis**”, som også bruges til dokumenter, der er deciderede vejledninger, hvilket vil give anledning til forvirring om papirets karakter.

Disse ændringer foreslås, da høringen hos F&P’s medlemmer viser, at der er tvivl, om hvordan dokumentet skal opfattes. Der bliver spurgt til, om papiret skal ses som inspiration, en vejledning til god praksis eller krav til virksomheder. Baseret på afgrænsningen side 7 og den mundtlige præsentation, som Finanstilsynet gav ved Fintech forum den 13. juni 2023, er det entydige formål at inspirere. Dokumentet bør derfor gennemgås for ordbrug m.v., som skaber forvirring om dette formål.

#### Et eksempel:

Side 3, 2. afsnit: ”Virksomhederne **bør** f.eks. overveje, hvilke dataetiske rammer de vil opstille, når de bruger AI til at behandle data for at levere eller administrere finansielle produkter eller tjenester.” **Bør** ændres til **kan** for at afspejle, at der er tale om inspiration.

Ændringen vil således afspejle ordbrugen senere i dokumentet, fx side 5 første linje: ”Finanstilsynet ønsker med dette papir om god praksis at give inspiration til overvejelser, som virksomheder i den finansielle sektor **kan** gøre sig om deres arbejde med dataetik.”

Side 22: I afslutningen omtales de dataetiske værktøjer, der gennemgås, pludselig som anbefalinger. Qua papirets karakter bør man fastholde brugen af ordet værktøjer og undgå ordet anbefalinger om de enkelte nedslag, da ordet anbefalinger giver en klart vejledende karakter. Fx ”Derfor anbefaler Finanstilsynet, at virksomheder i den finansielle sektor, som arbejder med dataetik, **tager stilling til ovennævnte anbefalinger** og forholder sig til, hvilke værktøjer der fungerer bedst i den enkelte virksomhed.” Markeringen med fed kunne derfor ændres til eksempelvis ”ser på ovennævnte værktøjer”.

## **Tydelig fremhævelse af eksempler og læringspointer:**

I opsætningen af dokumentet vil vi generelt foreslå, at der bliver en mere tydelig opstilling af pointer og læringspunkter, således at disse er lette at afkode fra dokumentet. Hvis papiret skal fungere som et arbejdsdokument, vil det være en fordel med et lettere format.

Et eksempel er første afsnit på side 5. Der indeholder en række enkeltstående forslag, der kunne opstilles som bullets.

Eksemplerne på dilemmaer kan også bruges som opslagsværk. Her kunne det igen lette formatet, hvis de enkelte dilemmaer opstilles med en klar ramme og de centrale overvejelser opsummeres som bullets – se også vores tekstmære kommentarer til dilemmaerne.

Yderligere forslag til denne type ændringer gennemgås i de tekstmære rettelser.

## **Samlet oversigt over inspirationsmateriale**

For at styrke dokumentet som opslagsværk vil vi foreslå, at de mange henvisninger til øvrige materialer også samles i én inspirationsoversigt, så læseren har en samlet oversigt over de materialer, der henvises til i dokumentet. Så disse ikke skal findes via fodnoterne på de enkelte sider.

I tråd med ovenstående vil vi også foreslå, at der indsættes en oversigt over relevante regler og lovgivning, som der henvises til i dokumentet. Hermed skabes en tydelig kobling mellem dokumentet og eksisterende relevant lovgivning.

## **Negativ vinkling af dilemma-eksempler**

F&P oplever generelt, at eksemplerne på dilemmaer fremstår som om databrugeren er negativ. Det finder vi uheldigt, da branchen er gået i forvejen på mange områder i at danne dagsorden og netop søgt at sikre, at kunderne får en positiv oplevelse af og har tillid til branchens databrug.

For at sikre balance bør de positive effekter af databehandlingen også nævnes, da de også er en del af en dataetisk afvejelse. Derudover kan man nævne, hvilke tiltag der kan bidrage til, at kunderne kan opfatte en given databehandling positivt (hvis den foretages ansvarligt). I tillæg kunne F&P savne eksempler, hvor dataetiske overvejelser er endt med, at kunden opfatter en databrug som positiv.

I relation til dilemmaerne vil vi samtidig foreslå, at eksempel F udgår (som dilemma), da eksemplet, som det er beskrevet, ikke illustrerer et muligt dilemma, men et reguleret scenarie.

F&P

Sagsnr. GES-2023-00170

DokID 465979

## Yderligere forklaring af Finanstilsynets holdninger

Mange steder i papiret beskrives Finanstilsynets holdninger. De fremgår dog nogle steder uden videre underbygning eller forklaring. Fx side 17 nederst: ”Det er vigtigt, at dataetisk arbejde bliver forankret ét sted i organisationen.” Da det også kan være gangbart med forankring flere steder i organisationen, ville det være godt med argumentationen for ønsket om forankring ét sted.

## Tekstnære input

Det følgende er en række tekstnære input til papirets forskellige afsnit.

### 1. Indledning

Generel kommentar: Indledningen lægger vægt på, at AI og anvendelse af data i den finansielle sektor har positive kommercielle anvendelser for virksomhederne, som øget effektivitet, målretning af produkter og indtjening.

AI og anvendelsen af data kan ikke fremstilles alene som fordele for virksomhederne, da der også er værktøjer, som kan understøtte kvaliteten i kundernes retskrav overfor virksomhederne, fx retten til at blive godkendt til et lån eller retten til udbetaling af erstatning i henhold til policens dækning. AI kan som værktøj også bidrage til at sikre mere forudsigelige og materielle korrekte afgørelser i sektoren til fordel for kunderne.

### 2. Baggrund for Finanstilsynets fokus på dataetik

Side 3, nederste afsnit: ”Men det er også væsentligt, at forbrugere kan have tillid til, at de gør det inden for rammerne af, hvad vi i Danmark grundlæggende opfatter som etisk ansvarligt.”

Vi mener, at det bør være op til forbrugerne at vurdere, hvor de ønsker at være kunde - ud fra deres holdninger til virksomhedernes adfærd. Den nuværende formulering læses således, at branchen pålægges at handle ud fra, hvad gennemsnittet af forbrugerne ser som generelt ansvarligt. Det kan være begrænsende og svært at konkretisere. Det foreslås i stedet at flytte fokus til, at forbrugerne kan afkode, om virksomhederne handler inden for rammerne, som de ser som etisk ansvarligt. Her er transparens omkring dataanvendelse en relevant parameter.

Forslag til formulering: ”Men det er også væsentligt, at forbrugere kan have tillid til og afkode, om den enkelte virksomhed handler inden for rammerne af, hvad de opfatter som grundlæggende etisk ansvarligt.”



Mulighed for øget tillid til den enkelte virksomhed, som danner basis for databrug, man ikke tiltror andre virksomheder, er netop det punkt, hvor dataetikken kan give en konkurrencemæssig fordel.

Side 4, andet afsnit: Her omtales F&P's udarbejdede principper. Et praktisk eksempel, som kan overvejes inkluderet, til gavn for resten af den finansielle sektor, er de dataetiske positioner, som F&P gør sig overvejelser om her: <https://www.fogp.dk/temaer/dataetik/paa-vej-mod-en-faelles-dataetik/>

### 3. Sammenfatning

Generel kommentar: F&P savner generelt et fokus på, at dataetik i særdeleshed er at etablere en klar overensstemmelse mellem kundernes forventning og virksomhedens ageren. Det gælder både i dette afsnit og i eksemplerne. Hvis der kan skabes en klar forventningsafstemning i forhold til databrug, så sikrer det tillid.

Side 5, første afsnit: Overvej om de punkter, som gennemgås i første afsnit, kan opstilles som bullets, så det er let at få overblikket over de enkelte inspirationspunkter.

Side 5, første afsnit: For at gøre sammenfatningen i afsnittet mere forståelig for modtageren, kan det overvejes, om det også kan fremhæves, at en datapolitik også kan nævne, hvad den finansielle virksomhed ikke vil gøre med data.

Side 5, andet afsnit: *"Alle virksomheder kan med fordel overveje, om deres arbejde med dataetik er så væsentligt, at de kan beskrive det i årsrapporten, selvom de ikke er omfattet af de nye regnskabsregler."* Denne tekst opfattes som meget normerende. Da papiret ikke er bindende, bør teksten blødes op eller udgå.

### 4. Det dataetiske landskab

Side 6, tredje og fjerde afsnit: Overvej om afsnittet med anbefalinger fra myndigheder og organisationer kan opstilles som en bullet-liste for at lette overblikket. Det kan desuden overvejes at inkludere det dataetiske etikkompas som inspiration.

Side 7, andet afsnit og citat: F&P finder, at citatet på side 7 uden kontekst kan komme til at fremstå meget negativt og foreslår enten, at citatet udgår eller udvides med mere af den kontekst, det oprindeligt fremstår i, som var en præsentation af branchestandarden for dataetik "Cool eller Creepy".

Hvis man ikke ønsker at slette citatet, så foreslås den indledende tekst udvidet således – tilføjelse med fed/kursiv: ”Dataetik forveksles ofte med databeskyttelsen efter GDPR. Dataetik går dog videre end data- og privatlivsbeskyttelse i snæver forstand. Den handler også om, hvordan databrugen, udvikling og brug af kunstig intelligens m.v. påvirker vores samfund. Det uddybes nærmere i afsnit 6. **Hos forsikring og pensionsbranchen har man indført branchestandarden ”Cool eller Creepy”, der for branchen lægger et lag oven på lovgivningen, der skal bidrage til etisk databrug.**”

Og citatet er udvidet således:

**”GDPR fjerner på ingen måde de svære etiske valg og dilemmaer, som øget digitalisering medfører. Når branchen lever af at levere tryghed, er det helt centralt løbende at drøfte branchens databrug. For du kan sagtens være GDPR-compliant og have en creepy databrug.”**

## 5. Afgrænsning

Side 7, nederste afsnit: ”Der er tale om **rent vejledende principper** og formidling af, hvad der kan blive god praksis i fremtiden i takt med, at dataetik får stigende betydning.” I tråd med papirets afsæt i inspiration foreslås den fremhævede tekst ændret til ”Der er tale om **ren inspiration** og formidling af ...” At der er tale om inspiration fremhæves i øvrigt senere i afsnittet, så rettelsen vil samtidig sikre, at der ikke opstår usikkerhed om formål.

## 6. Hvad er dataetik?

Generel kommentar: F&P vil foreslå, at det i tråd med pointerne i de forrige afsnit kort i starten af dette afsnit præciseres, at dataetik altid er knyttet til databrug. Fx fra side 7: ”**Begrebet dataetik omfatter brugen af alle former for data og omhandler virksomheders generelle tilgang til anvendelse af disse data.**” Og fra bemærkningerne til ÅRL: ”**Begrebet dataetik omfatter brugen af alle former for data og omhandler virksomheders generelle tilgang til anvendelse af disse data.**”

Generel kommentar: Det nuværende afsnit har meget fokus på, at dataetik alene er fokus på individet. Det fremgår fx som hovedformålet med dataetik i afsnittet lige under citatet fra Dataetisk Råd. F&P mener, at en samfundsvinkel også bør tydeliggøres i dette afsnit. Dataetik er *også* afvejningen mellem gevinst for fællesskabet og en potentiel omkostning for den enkelte. (Group fairness vs. individual fairness).

## 7. AI og dataetiske dilemmaer i den finansielle sektor

Generel kommentar: F&P læser eksemplerne på den måde, at man med hvert enkelt eksempel har ønsket at sætte fokus på et eller flere konkrete hensyn. Eksempelvis 'hensynet til kunden' i eksempel A. "Nytte kontra mulig oplevelse af overvågning" i eksempel B. "Graden af automatisering og sikkerhedsmekanismer" i eksempel C. Og så videre.

F&P foreslår på denne baggrund, at det tydeliggøres, hvad hvert enkelt dilemma sætter fokus på. Eksempelvis med sigende overskrift og kort indledning. Derudover foreslår F&P, at hovedpointer fra de mulige overvejelser opsummeres med 1-3 bullets for hvert dilemma. Formålet med disse tilretninger vil være at øge brugbarheden af papiret som opslagsværk.

Generel kommentar: Eksemplerne kunne godt være mere tydelige i brugen af AI. Eksempel A vil fx umiddelbart kunne gennemføres med den blotte indsamling af lokationsdata kombineret med forsikringsdata.

Generel kommentar: Lovligheden efter GDPR af flere af eksemplerne, fx eksempel A og D, kan være tvivlsom. Virksomhederne overholder GDPR og derfor bør det sikres, at eksemplerne ikke skaber forvirring om lovligheden efter GDPR.

Side 13, første afsnit: Human oversight/Human-in-the-loop beskrives som et relativt nemt håndtag at implementere, hvis man er usikker på AI'en. Men human oversight stiller enormt høje krav til den medarbejder, der skal udføre tilsynet. Artikel 14.4 i EU's AI act beskriver nogle af de krav, der er til personen, som skal udføre tilsyn.

Side 13, eksempel D: F&P savner, at det bliver tydeliggjort, at man ikke kan afsøge om en model diskriminerer (indirekte) ud fra et givent datapunkt, hvis man ikke har dette datapunkt. Etnicitet er et eksempel på en sensitiv information, som vi ifølge lovgivning og dataetik ikke må have utilsigtet bias på baggrund af. Denne bias kan dog ikke måles eller mitigeres for, hvis ikke man som databehandler har adgang til den pågældende information.

Side 14, eksempel F: F&P foreslår, at dilemmaet fjernes, da der netop redegøres for, at tilgangen anses som ulovlig. Eksemplet illustrerer dermed ikke en gråzone (et dilemma), som det er beskrevet.

## 8. Værktøj til dataetiske overvejelser i den finansielle sektor

Generel kommentar: F&P foreslår, at de vigtigste pointer for de enkelte nedslag opsummeres i 1-3 bullets efter hvert afsnit.

Side 17, udarbejdelse af dataetisk politik: F&P foreslår, at virksomhedens værdier også inkluderes som et muligt afsæt for en dataetisk politik.

Side 18, virksomhedens data og datakilder, andet afsnit: F&P foreslår, at de forskellige opmærksomhedspunkter i forhold til data listes som bullets for at øge læsevenlighed.

Side 18, virksomhedens brug af AI, første afsnit: F&P foreslår, at henvisningen til kommissionens anbefalinger fjernes, da der allerede er henvist til dem tidligere. Alternativt bør teksten blødes op, så det at forholde sig til kriterierne fremstår "frivilligt" og ikke som et krav.

Side 18, virksomhedens brug af AI, andet afsnit: Det kan overvejes at give brug af 3. partssystemer en særskilt overskrift.

Side 19: I teksten nævnes faren ved, at algoritmer trænes på baggrund af data, der af historiske årsager indeholder diskrimination. Man kunne overveje at tilføje, at hvis sådanne skævheder rettes op i algoritmerne, vil det omvendt kunne have den positive effekt, at de historiske og nuværende menneskeskabte forskelsbehandlinger (fx låneansøgninger/ansættelsesforhold for mænd kontra kvinder) vil kunne reduceres. Begrebet "positive bias" kan med fordel også uddybes yderligere. Teksten nævner også, at det er vigtigt at undgå skævheder i data. Det er dog vigtigt at understrege, at skævheder (når de er kendte) kan være ønskelige. Fx hvis man vil teste, hvordan en model performer med skæve data.

Side 21, Revision af arbejdet med dataetik: F&P vurderer, at det er for tidligt i udviklingen til at overveje revision af arbejdet med dataetik og foreslår, at forslaget om særskilt revision af dataetikken slettes.

Efter nuværende regler skal revisor foretage et konsistentstjek af ledelsesberetningen, som også omfatter, at revisor forholder sig til om reglerne for den dataetiske redegørelse er opfyldt samt at revisor gør opmærksom på det, hvis revisor er opmærksom på væsentlige fejl eller mangler i ledelsesberetningen.

Side 21, Revision af arbejdet med dataetik: Det kan overvejes at nævne D-mærket i en oversigt over værktøjer, som en eksisterende certificeringsmulighed.

Side 21-22, Rapportering på dataetik: "Virksomhederne kan f.eks. redegøre for deres overvejelser i forhold til anbefalingerne under afsnit 8.1-8.6 i dette papir. De kan altså beskrive:" Vi foreslår, at dette omformuleres, så det er i tråd med Erhvervsstyrelsens vejledning til den dataetiske redegørelse, som alene giver eksempler på nedslag. Det vil samtidig følge en overordnet tilgang om, at dokumentet skal inspirere ikke vejlede. Forslag til ny formulering: "Virksomhederne kan f.eks. redegøre for deres overvejelser i forhold til **et eller flere af eksemplerne på værktøjer beskrevet under afsnit 8.1-8.6 i dette papir. De kan altså beskrive:**"

Side 22, Rapportering på dataetik, listen med de 7 punkter: Det bør tydeliggøres, at i forhold til en politik (punkt 3), så er punkterne a, b, c eksempler på overvejelser, der kan danne afsæt for en principiel overordnet stillingtagen i en politik. Det handler ikke om at gennemgå og opremse disse elementer for hvert enkelt AI-drevet system i virksomheden.

F&P

Sagsnr. GES-2023-00170

DokID 465979

27. juni 2023

Til Finanstilsynet  
Att. Camilla Neuenschwander

*Finanstilsynets udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor – høringssvar fra Den Danske Aktuarforening*

Den Danske Aktuarforening (DDA) har med tak modtaget Finanstilsynets udkast til et papir om god praksis for dataetik ved brug af AI i den finansielle sektor i høring. Dataetik og herunder AI er et vigtigt emne for DDA, idet de teknologiske muligheder åbner for en række nye muligheder for brug af data, der – som Finanstilsynet klart gør rede for – kan rejse en række etiske spørgsmål og dilemmaer samt give anledning til etiske udfordringer.

Etiske overvejelser og hensyn er centrale for DDA og dens medlemmer. Foreningen har bl.a. til formål, at aktuarstanden skal bidrage til den generelle tillid og tryghed i samfundet. I den sammenhæng fremgår det desuden af foreningens vedtægter, at foreningen skal "virke for at fremme professionalisme, troværdighed og etiske hensyn inden for aktuarens kerneopgaver".

At dataetiske emner er af væsentlig betydning for DDA og dens medlemmer illustreres af, at foreningens bestyrelse i efteråret 2022 besluttede, at foreningens nyudviklede kursus om dataetik skal være obligatorisk for stort set alle foreningens medlemmer. Af kursusbeskrivelsen fremgår bl.a.:

"Med dataetik sættes menneskets interesser og ret til kontrol med egne data i centrum. Men dataetik handler også om at undgå uhensigtsmæssig diskrimination og andre skadevirkninger for den enkelte person, grupper af personer eller samfundet som helhed. Dataetik handler derfor om at blive opmærksom på de negative virkninger, som anvendelse af data og teknologi kan have for mennesker, miljø og samfund – og hvordan vi kan træffe bæredygtige valg, der gør gavn på alle tre niveauer."

Ønsket til medlemmerne om at gennemføre dette kursus implementeres nu over en treårig periode frem til 2025.

Aktuarforeningen afholdt desuden et medlemsmøde i maj 2023, hvor den tyske tilsynsmyndighed (BAFIN) holdt oplæg om dataetik og de dilemmaer og udfordringer, som forsikringselskaber m.fl. i den sammenhæng stilles over for. Baggrundsmateriale fra BAFIN vedlægges til orientering (med godkendelse fra BAFIN).

For Aktuarforeningen er god praksis for dataetik ved brug af AI således et højt prioriteret emne.

Vi har ikke konkrete bemærkninger til udkastet til papir om god praksis for dataetik ved brug af AI i den finansielle sektor. Vi noterer os, at der med papiret ikke er tilsigtet indført nye regler eller praksis mv.

I Aktuarforeningen er vi meget interesserede i en fortsat dialog med Finanstilsynet om brugen af AI mv. og de dataetiske – og regulatoriske – problemstillinger, der uden tvivl vil vise sig fremover.

Med venlig hilsen

Lars Sommer Hansen  
Formand for Den Danske Aktuarforening

# Machine learning in risk models – Characteristics and supervisory priorities

Consultation paper



# Table of contents

<b>I. Machine learning is advancing</b>	<b>3</b>
<b>II. Characteristics of ML</b>	<b>5</b>
<b>1. Dimensions and examples</b>	<b>5</b>
<b>2. Methodology</b>	<b>6</b>
2.1 Complexity and dimension of the hypothesis space	6
2.2 Complexity of training	7
2.3 Adaptivity	8
<b>3. Data basis</b>	<b>8</b>
<b>4. Use of the output</b>	<b>8</b>
4.1 Relevance in the model	8
4.2 Area of application	9
4.3 Degree of automation	9
<b>5. Outsourcing and IT infrastructure</b>	<b>9</b>
<b>III. Supervisory approach</b>	<b>10</b>
<b>1. Supervisory practice endures</b>	<b>10</b>
<b>2. Methods invite to „believe in data“</b>	<b>12</b>
<b>3. Focus on explainability</b>	<b>13</b>
<b>4. Adaptivity: model changes are more difficult to identify</b>	<b>14</b>
<b>IV. Outlook</b>	<b>17</b>
<b>V. Consultation</b>	<b>17</b>

# I. Machine learning is advancing

The debate surrounding the use of artificial intelligence and machine learning (collectively referred to as ML methods for short) has recently been gaining momentum, precisely in the area of financial services, fuelled by the availability of “big data” together with enhanced computing power. Standardised principles and procedures<sup>1</sup> have made the development of ML methods simple and accessible.

The use of ML methods can help to quantify risks more accurately and enhance process quality, thereby improving financial firms’ risk management.

These issues have already been addressed by both BaFin and the Bundesbank in several publications. On 16 July 2018, BaFin submitted its report entitled “Big data meets artificial intelligence – challenges and implications for the supervision and regulation of financial services”<sup>2</sup> (BDAI report for short) for public consultation, and on 28 February 2019, it published an overview of the results along with an initial assessment.<sup>3</sup> The BDAI report demarcated the field of digitalisation and identified challenges. In March 2020, BaFin explained its misgivings about a general approval requirement for algorithms and outlined how ML can be embedded in the risk-oriented supervisory approach.<sup>4</sup> These considerations are fleshed out in the “Principles for the use of algorithms in decision-making processes” published in June 2021.<sup>5</sup> In its discussion paper “The Use of Artificial Intelligence and Machine Learning in the Financial Sector”<sup>6</sup> published in November 2020, the Bundesbank set out basic theoretical considerations on how to deal with ML methods in the context of prudential supervision, in which, for instance, it derived inferences on the intensity of supervision of ML and put into context the significance of the explainability of ML methods.

Supervisors and regulators across the globe are also looking into ML methods. Noteworthy are papers published by De Nederlandsche Bank, which has developed principles for the use of ML,<sup>7</sup> and France’s ACPR, which centres on the explainability of ML.<sup>8</sup> The EBA<sup>9</sup> and the Bank

---

<sup>1</sup> These include DevOps, MLOps and software libraries which contain standard ML methods.

<sup>2</sup> BaFin, 2018, “Big data meets artificial intelligence”, available online at: [https://www.bafin.de/SharedDocs/Downloads/EN/dl\\_bdai\\_studie\\_en.pdf?\\_\\_blob=publicationFile&v=11](https://www.bafin.de/SharedDocs/Downloads/EN/dl_bdai_studie_en.pdf?__blob=publicationFile&v=11)

<sup>3</sup> BaFin, 2019, “Big data meets artificial intelligence – results of the consultation on BaFin’s report”, available online at: [https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/BaFinPerspektiven/2019\\_01/bp\\_19-1\\_Beitrag\\_SR3\\_en.html](https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/BaFinPerspektiven/2019_01/bp_19-1_Beitrag_SR3_en.html)

<sup>4</sup> BaFin, 2020, “Does BaFin have a general approval process for algorithms? No, but there are exceptions”, available online at: [https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Fachartikel/2020/fa\\_bj\\_2003\\_Algorithmen\\_en.html](https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Fachartikel/2020/fa_bj_2003_Algorithmen_en.html)

<sup>5</sup> BaFin, 2021, “Big data and artificial intelligence – principles for the use of algorithms in decision-making processes”, available online (but in German only) at: [https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl\\_Prinzipienpapier\\_BDAI.html](https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl_Prinzipienpapier_BDAI.html)

<sup>6</sup> Bundesbank, 2020, “Policy Discussion Paper, The Use of Artificial Intelligence and Machine Learning in the Financial Sector”, available online at: <https://www.bundesbank.de/resource/blob/598256/d7d26167bceb18ee7c0c296902e42162/mL/2020-11-policy-dp-aiml-data.pdf>

<sup>7</sup> De Nederlandsche Bank, 2019, “General principles for the use of Artificial Intelligence in the financial sector”, available online at: <https://www.dnb.nl/media/jkbip2jc/general-principles-for-the-use-of-artificial-intelligence-in-the-financial-sector.pdf>

<sup>8</sup> Autorité de contrôle prudentiel et de résolution (ACPR), 2020, “Governance of Artificial Intelligence in Finance”, available online at: <https://acpr.banque-france.fr/en/governance-artificial-intelligence-finance>

<sup>9</sup> EBA, 2020, “Report on Big Data and Advanced Analytics”, available online at: <https://www.eba.europa.eu/eba-report-identifies-key-challenges-roll-out-big-data-and-advanced-analytics>

of England<sup>10</sup> have likewise published their views of data-driven analyses and ML. The paper published by the Hong Kong Monetary Authority has placed a practice-oriented focus on the hurdles involved in implementing ML.<sup>11</sup>

This consultation paper builds on earlier national and international publications and makes connections between the prudential risks of ML and current supervisory practices. This approach not only simplifies the background but also follows the process that financial corporations undergo when introducing ML methods.

The focus of this paper is on solvency supervision, and specifically the application of ML methods in areas of particular relevance to supervisors. These include, on the one hand – as an exception to the principle that algorithms do not require supervisory approval – ML methods that are used in prudential inspections and approval procedures, and thus in internal models for calculating regulatory own funds requirements (Pillar 1) and, on the other hand, those that are used in risk management under Pillar 2.

In this context, consumer protection aspects and the ethical issues of ML play a relatively minor role and will therefore be disregarded in the following.<sup>12</sup>

There is no uniform definition of machine learning due, first, to the large number of different approaches and, second, to the lack of a clear dividing line to traditional techniques.<sup>13</sup> However, ML methods often have certain characteristics that are particularly strongly pronounced and which thereby set them apart from traditional techniques. The intended aim of this paper is therefore to identify such characteristic traits of ML methods which have relevant implications for supervisors and to come up with ideas on how supervisory practices could evolve in order to be able to respond to the risks involved.

This also poses the question as to whether not only supervisory practices but also the regulatory foundations themselves need to be reworked and whether or not it may be necessary to create a fundamentally new supervisory approach for ML methods. Without wanting to “jump the gun” here: since the current regulatory foundations are worded in a technology-neutral manner, they are largely transferable to the characteristics of ML methods, with only a few places where it might be necessary to adapt the regulatory foundations.

Although this paper refers below primarily to “banks”, the characteristics and prudential implications presented here can, in principle, also be applied to insurers and other enterprises engaging in the development and implementation of ML methods for regulated financial services. This paper is intended as a consultative document in order to launch a discussion process with the industry: it contains blocks of questions, the responses to which are designed to advance supervisory practice.

---

<sup>10</sup> Bank of England, 2019, “Machine learning in UK financial services”, available online at: <https://www.bankofengland.co.uk/report/2019/machine-learning-in-uk-financial-services>

<sup>11</sup> Hong Kong Monetary Authority (HKMA), 2019, “Reshaping Banking with Artificial Intelligence”, available online at: [https://www.hkma.gov.hk/media/eng/doc/key-functions/financial-infrastructure/Whitepaper\\_on\\_AI.pdf](https://www.hkma.gov.hk/media/eng/doc/key-functions/financial-infrastructure/Whitepaper_on_AI.pdf)

<sup>12</sup> Within supervisors’ prudential mandate, these topics are covered by examinations of operational risk.

<sup>13</sup> Some known definitions created by regulators and developers are listed in the annex; their variety illustrates the difficulties involved in obtaining a precise taxonomy.

The paper is structured as follows. Chapter II identifies characteristics of ML methods which could be relevant to the design of supervisory practice. On that basis, Chapter III discusses potential changes to supervisory practices. Chapter IV summarises the key findings and Chapter V lays out a roadmap of the consultative period.

## II. Characteristics of ML

This paper does not develop a universally applicable definition of ML methods. Its purpose is, rather, to create a boundary which is sufficient for supervisory purposes to assess models under Pillars 1 and 2. Depending on whether and which ML characteristics exist for a specific methodology to be examined and the extent to which they are pronounced, these characteristics will be discussed in terms of supervisory practice, inspection techniques and inspection intensity.<sup>14</sup>

### 1. Dimensions and examples

The characteristics can be grouped into the three dimensions of the **AI/ML scenario**, which is explained in greater detail in the Bundesbank discussion paper:<sup>15</sup>

The (1) **methodology and data basis** collectively describe the complexity and thus the model risk associated with the ML procedure. The (2) **use of the output** contains the importance of the procedure within risk management. The intensity of inspections is not guided by the distinction between (3) in-house development and **outsourcing** or the underlying **IT infrastructure**.

**Table 1: Characteristics of ML methods**

AI/ML scenario	Characteristics
1 Methodology and data basis	Complexity and dimension of the hypothesis space Complexity of training Adaptivity Data sources Data types Data volume
2 Use of the output	Importance in the model Area of application Degree of automation
3 Outsourcing and IT	Outsourcing

<sup>14</sup> This approach is reminiscent of the "duck test": see [https://en.wikipedia.org/wiki/Duck\\_test](https://en.wikipedia.org/wiki/Duck_test).

<sup>15</sup> Bundesbank, 2020, "The Use of Artificial Intelligence and Machine Learning in the Financial Sector", online at: <https://www.bundesbank.de/resource/blob/598256/d7d26167bceb18ee7c0c296902e42162/mL/2020-11-policy-dp-aiml-data.pdf>

ML methods and conventional approaches form a continuum. The figure below assigns characteristics to two fictitious rating system use cases.

Use case 1: A conventional IRBA rating system uses a logistical regression to estimate borrowers' probability of default (PD). A small number of variables are entered into the model. The data are structured.

Use case 2: A multilayered neural network is used to estimate the PD for an IRBA rating system. The model is rarely retrained. The dataset comprises many variables, both structured and unstructured data. The results from the neural network, together with the results of conventional models, feed into the estimation of PD.

The characteristics presented here are typically more strongly pronounced in the neural network. Other ML methods or areas of application, for their part, have their own profiles.



## 2. Methodology

### 2.1 Complexity and dimension of the hypothesis space

Models in Pillars 1 and 2 reflect a causal link assumed by the modeller between input (such as market and portfolio data) and output (such as prices of financing instruments or risk metrics). This link is also referred to as a hypothesis. The basic eponymous feature of machine

learning is the fact that the algorithm used only contains a small number of assumptions regarding the structure of the problem to be described. This structure is given, for example, in the form of fixed rules.<sup>16</sup> Alternatively, parameters describing this problem structure are also learned using the data themselves. The hypotheses available for adjustment are therefore much more varied, and the hypothesis space is not completely controlled by analytical techniques, especially as its selection is not primarily motivated by causalities.

As in many more conventional methods, the modeller only specifies the space of all hypotheses applicable to the method (hypothesis space). This space is to be searched, and the learning method selects a specific hypothesis (in the form of a set of model parameters) via optimisation. The hypotheses can have any level of complexity. The modeller defines the potential hypotheses through the selection of the ML method (e.g. neural network, random forest, k-nearest neighbours) and their specification (model design, hyperparameters).

Simple hypothesis spaces also occur in linear or logistical regressions, which are present in all conventional internal models.

A deep neural network (DNN) constitutes a much more complex model structure. Here, the hypothesis is mathematically expressed through the composition of non-linear and linear maps so that the hypothesis space has a considerably larger dimension. The functional relationship between input and output learned from a DNN can generally no longer be understood by means of a simple description using mathematical formulae, which is referred to as a “black box” characteristic.

The limited transparency of the model’s behaviour has consequences for the model development, model validation and the implications for the significance of the underlying data. It also gives rise to additional challenges with regard to the explainability of model results in order to ensure a sufficient understanding of the functional relationship and to justify the applicability of the ML methods internally and externally.

## 2.2 Complexity of training

Determining a specific hypothesis occurs in a learning process referred to as “calibration” or “training”.

One main feature of training newer ML methods is the high number of calculations due to the large number of function arguments and model parameters, the complex sequence of nested calculation instructions, and iterative procedures. This causes challenges to arise in technical implementation, such as in the availability of necessary hardware resources or the numerical stability of the calculation method.

In general, there is no unique solution to an optimisation problem involving non-linear functions in high-dimensional spaces. The selection of an optimum as a specific hypothesis depends on the training algorithm and may also depend on randomness. Choosing different optima in successive training sessions can affect the stability of the ML procedure.

---

<sup>16</sup> These do not include expert systems that, for their part, optimise the weightings of “if-then” rules.

As the complexity of training increases, these optima come under closer supervisory scrutiny.

## 2.3 Adaptivity

Some ML methods are designed to be adapted to new data very frequently or even on a virtually continuous basis. This leads to a blurring of the distinction between model development and model operation, and between model maintenance and model changes subject to supervisory assessment. This distinction is particularly relevant under Pillar 1 if model changes require supervisory approval. Furthermore, there is the question of to what extent it is possible to validate these models and reproduce model output. Ensuring the continuous, adequate quality of data likewise plays an important role. As the adaptivity of procedures increases, a clear differentiation between model maintenance and model changes becomes more important from a supervisory perspective.

## 3. Data basis

The increasing use of large data volumes is a distinguishing feature of current developments at banks, but it is also a feature of complex or new ML methods in the narrower sense.

ML frequently builds on a larger number of **data sources** and a network of different data sources. Synthetic, i.e. artificially generated data are also used, as are unstructured data (**data type**). ML methods can often handle a large number of input parameters in the model, giving rise to a large **data volume**.

The performance of ML methods is determined not least by the volumes of data available for training, by the veracity of the data and the data quality. The dataset is thus coming under greater scrutiny from supervisors.

## 4. Use of the output

### 4.1 Relevance in the model

The ML method can have different roles within the entire model. It can be integrated into a model as a supporting component, e.g. to prepare data, or as a sub-component, e.g. as a module of a rating scheme. Sometimes the ML method is also the central component of a model. Or it is used outside of the internal model, e.g. as a challenger tool or as a proxy for the "real" model within specific areas of application.

As the importance of the ML method increases within and alongside the model, the more intensely it comes under supervisory scrutiny.

## 4.2 Area of application

The area of application outlines which results the ML method provides and to what extent these are included in the enterprise's business processes. Examples of areas of application may include internal models, early warning systems for risks or credit ratings. The greater the impact of the area of application on the risk situation, the stricter supervisors set the requirements for the ML method.

## 4.3 Degree of automation

The degree of automation can be divided into algorithm-determined and algorithm-based processes; the associated operational risk can be divided accordingly. Algorithm-determined processes refer to largely automated processes that use ML results and thus entail greater risks where ML methods are insufficiently monitored. By contrast, algorithm-based processes rely on human-controlled processes to a greater extent (which in turn has its own risks).

## 5. Outsourcing and IT infrastructure

Outsourcing to specialist service providers and the use of a specific IT infrastructure are also typical for ML methods. Fintech and BigTech firms offer modular systems to create ML methods and provide an IT infrastructure tailored to high levels of performance. Difficulties may arise if ML methods are integrated into so-called "legacy" IT infrastructure.

The approach supervisors take to assess these service relationships does not change with the use of ML methods; the relevant regulations such as the prudential requirements for IT (*Bankaufsichtliche Anforderungen an die IT*, or BAIT) also cover these use cases. This is why Chapter III does not discuss these aspects specifically.



### Questions on Chapter II:

- a) Do you think it is appropriate to forgo a strict definition of ML methods and instead take an application-based approach and gear supervisory and inspection practice to the individual characteristics of the methods used?
- b) What other characteristics of ML methods do you believe could be important for supervisory practice or for internal model governance?
- c) In your opinion, which characteristics do not belong in this overview?
- d) In which relevant areas of application do you employ ML methods or where do you intend to implement them?

## III. Supervisory approach

As already mentioned in the introduction, ML methods do not generally require new supervisory practices. This chapter draws supervisory conclusions for the characteristics of the ML methods described in Chapter II. It also outlines to what extent adjustments to supervisory practice are required in certain areas. The supervisory approach is defined by the principle of proportionality and using the characteristics from Chapter II. The present chapter is structured according to the risks identified for ML and integrates these into the cycle of model development and maintenance.

### 1. Supervisory practice endures

Pillar 1 includes extensive regulations for reviewing and approving internal models formulated in a technology-neutral manner and therefore also addresses the risks of ML

methods. Principle-based requirements for risk management and IT provide a sound footing in Pillar 2.<sup>17,18</sup>

Supervisors are systematically pursuing an inspection approach geared towards banking processes that establishes general, overarching inspection areas for each risk type to be inspected (e.g. credit or market risk) and is continuously adjusted to current circumstances.

Supervisory practice for ML methods can therefore also be derived from the existing framework. At the same time, an outlier analysis, also supported by this consultation, is currently surveying the areas in which the supervisory inspection approach needs to be fleshed out in order to cater to the peculiarities of using ML methods.

This not only takes account of their mathematical/methodological aspects, but also of how these ML methods are integrated into processes, which is just as important for their controlled and thus successful and efficient use.

Supervisors are focusing on any new or much more pronounced risks that arise from ML methods. These are revealed in the data basis, validation (from model development to the test procedure to operation), model changes and management.

At a more abstract level, new draft legislation such as the AI Regulation drafted by the European Commission has placed special emphasis on aspects of consumer protection relating to ML. Supervisors are facing the task of inserting new requirements into the existing requirements of Pillar 1 and 2 models in a consistent manner.

---

<sup>17</sup> The CRR and MaRisk primarily form the legal basis for Pillar 1 and 2 inspections, supported by EBA and SSM standards. In this case, Commission Delegated Regulation (EU) Nos 2014/529 and 2015/942 are to be applied for Pillar 1 models. The EBA/GL/2020/06 (Guidelines on loan origination and monitoring) states the requirements for the use of automated models in creditworthiness assessments and credit decision-making covered by the German supervisory approach.

<sup>18</sup> BaFin, 2021, "Big data and artificial intelligence – principles for the use of algorithms in decision-making processes", available online (but in German only) at:  
[https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl\\_Prinzipienpapier\\_BDAI.html](https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl_Prinzipienpapier_BDAI.html)

### Questions on Chapter III.1:

- e) In your opinion, do existing regulations already contain prudential requirements that appear to hinder the use of ML methods? Do you believe that contradictions will arise between prudential regulations for Pillar 1 and 2 models and the draft AI Regulation? Please state any relevant references to the corresponding regulations and explain the challenges.
- f) To what extent do you believe the requirements laid out in EBA/GL/2020/06 with reference to the use of automated models in creditworthiness assessments and credit decision-making are also suitable for other ML methods in Pillar 2 (MaRisk) and should be taken on?
- g) Are there any other points where you believe current supervisory practice requires adjustment in order to appropriately acknowledge ML procedures and their associated risks?
- h) Do ML methods entail specific risks for IT implementation and outsourcing management? Are “adversarial attacks” conceivable in the financial sector and should ML methods be given particular protection against such attacks?

## 2. Methods invite to „believe in data“

Data quality is already an issue of key importance in supervisory action. Here, however, the characteristics of ML methods make clear that the data basis should be viewed particularly as a starting point and as a success factor. Unstructured data can now be exploited by and for ML methods. Furthermore, ML methods allow for calculations that factor in a large number of determinants. This makes it easy for modellers to quickly scale ML to large datasets.

ML methods learn what they find in the data provided and replicate any patterns contained therein. The black box characteristic may result in problems being obscured by seemingly good performance. For example, it is possible that models may learn correlations between input data that do not actually represent real relationships but are only based on coincidental characteristics of the learning dataset (model overfitting).

ML methods can utilise large volumes of data; the quality of these data must be continually ensured. This applies not only when developing and validating models, but also in their application.

Supervisors expect banks to undertake additional efforts to guarantee the quality of the underlying data. In particular, this involves ensuring that the training data are free from any systematic bias regarding the functional relationships to be learned by the model.

### Questions on Chapter III.2:

- i) What challenges do you see when selecting data and when ensuring data quality with regard to ML methods?
- j) In your opinion, what aspects of data quality are made easier through the application of ML methods?

### 3. Focus on explainability

As the hypothesis space that can be reflected by the model becomes more complex and its dimension increases, it also becomes more difficult to describe the functional relationship between input and output (i.e. the hypothesis specified in the training) verbally or using mathematical formulae, and the details of the calculations are less transparent for modellers, users, validators and supervisors. As a result, it is more difficult to comprehend the modelling and, if applicable, to check the validity of the model output as well. User acceptance may also suffer.

This black box characteristic can be considered the price of better model performance. However, it can be entirely justified, for example by higher predictive ability. Yet this may have to be weighed up against potentially greater model risk, depending on the significance of the model within the context of the associated banking processes.

Modellers must justify why the benefits gained are worth the trade-offs in the comprehensibility of the model. The extent to which a black box could be acceptable in supervisory terms is also dependent on how the model concerned is treated in the bank's risk management.

The black box may hide the fact that, amongst other things, ML methods learn relationships in the data that lack any real basis and do not allow for any general conclusions to be drawn.

As a result, it is the explainability and plausibility of the model behaviour, rather than its comprehensibility in detail, that gain in significance overall.

The term "explainability" is multifaceted, as modellers, validators, supervisors and users have different specialist backgrounds and require different information. In order to take account of these user-specific requirements, "explainable AI" (XAI) methods were developed.

From a supervisory perspective, XAI methods are highly promising with regard to mitigating the black box characteristic. However, XAI methods themselves represent models with

assumptions and weaknesses, and, in many cases, are still in the testing phase. As a result, it is a challenge to integrate these methods into structured processes. For example, it should be determined when which methods can be used, whether global or local approaches should be employed, which selection criteria and sample sizes are needed for local explanations, and which user groups should be targeted.

Banks must employ validation methods that are tailored to the model structure and that adequately cover the model risks. In the case of machine learning procedures, there is a particular danger of overfitting, for example due to the large number of parameters. Alongside the usual approaches, such as out-of-sample validation and backtesting, XAI approaches can help to identify overfitting. In addition, synthetic or stress/extreme scenarios as well as tests against traditional procedures could support the plausibility and explainability of ML methods.

### Questions on Chapter III.3:

- k) In your opinion, what impact does the black box characteristic have on the validation of the procedure?
- l) How important is the trade-off between performance and explainability for you?
- m) Do you believe XAI methods (always) offer a way out of the black box dilemma? Which processes are most promising and for which ML methods?
- n) How do you think XAI should be incorporated into the validation process?

#### 4. Adaptivity: model changes are more difficult to identify

Institutions and enterprise are obligated to inform supervisors of any changes to Pillar 1 models and, if applicable, only implement these changes after they have been approved. There is no clear-cut distinction between regular model maintenance and model change, which continually leads to discussions with supervisors, especially as the term "model change" is also dependent on the prevailing supervisory context.<sup>19</sup>

The flexibility and, in some cases, high-frequency adaptivity – i.e. frequent adjustment due to new data, for example – of ML procedures make it more difficult to draw a clear line between

---

<sup>19</sup> For IMM models, see the table in Section 3.10 of the ECB guide on materiality assessment (EGMA), which presents illustrative examples of model maintenance and model change. For market risk, see also Commission Delegated Regulation (EU) No 2015/942, Annex III, Section 2, which lists examples of model changes.

adjustments and changes; this is, however, indispensable for supervisors. As a general rule, the need for high-frequency adaptivity should be thoroughly justified.

While this document cannot provide any ML-specific definition of the term “model change”, it uses examples to illustrate how supervisors can classify changes.

- Procedural criteria, such as the first-time use of an ML method for a given task or changes to processes in conjunction with application of the method, can, as is the case for traditional methods, be used to identify and classify model changes.
- The question of whether retraining – i.e. training following an update to the underlying data – an ML method should be considered a change cannot be generally answered on the basis of comparison with traditional approaches.

For example, in the case of counterparty credit risk, recalibrating the probability distributions of risk factors is generally classified as model maintenance and not model change.<sup>20</sup> However, the hypothesis space that can be depicted by these models is significantly smaller than those of modern ML methods, and the functional relationship between input and output is not fundamentally changed by the recalibration of such a model.

By contrast, retraining an ML method can also involve structural changes to the function reflected in the method, such as the number of layers in a DNN. Even though the functional relationship between input and output to be reflected by the model may have only marginally changed (e.g. due to market movements), the realised “hypothesis” may differ to the previous one to such a significant degree that it would have to be considered a model change in supervisory terms.

Regardless of their classification, the high flexibility and adaptivity of ML methods mean that, following successive instances of training or adjustment, they can become far removed from the original models – in the sense that the inputs flow into the outputs with entirely different weights – within just a short period of time. For this reason, this adaptability – more so than in the case of traditional methods – must be subject to checks to prevent the model structures from fundamentally changing, which may nullify their explainability and validation, without the models having to undergo renewed supervisory evaluation.

- An extension of the risk factors used in the model would still be considered a model change. However, this distinction can be difficult to make in the case of complex datasets, especially of unstructured data.

These examples also illustrate the necessary shift in supervisory focus and the approach towards model adjustments:

---

<sup>20</sup> EGMA, p. 18. For market risk, see also the simplifications for smaller model changes in rapidly changing markets listed on p. 4 of EBA/RTS/2014/10.

- **Model approval:** The possible methods of model maintenance (including ongoing internal model validation and model changes) must already be assessed within the scope of model approval. The model change policy to be drawn up by the bank must take account of the particular features of the utilised ML methods when classifying adjustments. XAI processes could potentially also be employed here.
- **Communication:** Supervisors must also be able to trace activities related to regular model maintenance in order to identify in good time whether the original model is being significantly changed through a number of incremental model changes.
- **Internal validation:** Compared with traditional methods, the internal validation of ML methods has a greater focus on monitoring and ensuring the suitability of the methods employed on a continuous (and not just annual) basis. Model validation must make clear which change to the model structure will have what effect on the model output. This places greater demands on validators. This is the only way to ensure that unintended effects arising from model changes can be identified and avoided in good time. High-frequency retraining raises the question of which version of the model should be validated (which is critical in the case of ongoing retraining, for example).

Many ML methods are used in the supervisory area of Pillar 2 where approval is not required. This results in greater flexibility for model retraining and changes, yet existing requirements in this regard (e.g. from MaRisk) remain in effect here, too. From a supervisory perspective, it is nevertheless crucial, despite this flexibility, to adapt the training cycle to the specific use case, as well as provide the necessary justification, in order to create a balance between the explainability and validation of the model and ensuring that the data are up-to-date.

### **Questions on Chapter III.4:**

- o) What questions on supervisory practice do you see arising with regard to model adjustments for ML methods?
- p) Do you believe it is necessary for certain ML methods to be retrained at very high frequencies?
- q) Do you see ML methods necessitating changes in model governance? How do traditional modelling units, validators and new "data science" units work together?

## IV. Outlook

With regard to the choice of supervisory focus, in addition to the general BDAI principles, it is essential to achieve clarity in the development and application of ML methods in the context of models relevant for supervision in Pillars 1 and 2. This will create an environment which enables enterprises to invest in ML and address the risks of these methods as early as possible. At the same time, the impetus to adopt ML must come from the enterprises themselves. Supervisors do not insist on the use of ML methods as long as “traditional” methods of meeting regulatory requirements are still seen as suitable.

The next step will be to use this consultation paper to enter into dialogue with enterprises. Supervisors also consider it necessary to harmonise international approaches as far as possible and set identical cross-sector criteria for the use of ML methods. The European Commission’s digital finance strategy<sup>21</sup> plays a role in European standardisation.

## V. Consultation

BaFin and the Bundesbank invite the industry to comment on this consultation paper and answer the questions posed within. Please respond by email to [Konsultation-11-21@bafin.de](mailto:Konsultation-11-21@bafin.de) and [ai-b3@bundesbank.de](mailto:ai-b3@bundesbank.de) by 30.09.2021.

Comments submitted will be consolidated and published in anonymised fashion.

---

<sup>21</sup> European Commission, 2020, “Digital Finance Strategy for the EU”, available online at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0591>



# Annex

## Frequently asked questions

### **Do procedures which explicitly reflect presumed functional relationships also count as ML?**

Not in the narrower sense of ML – in a statistical context, learning instead focuses on recognising functional relationships and regularities on the basis of past observations in order to make better predictions in the future.<sup>22</sup> Machine learning hinges on one specific learning problem: how to apply statistical methods to a collection of data (input-output) to learn a function which can be used to predict the output for new input.<sup>23</sup> Machine learning is used because modellers can neither keep track of the technical relationship between input and output, nor find if-then conditions for it.<sup>24</sup>

### **Do “traditional”, long-standing models already known to supervisors (e.g. logistic regression) fall under the characteristics of ML?**

Even existing internal models use self-learning procedures which, to varying degrees, display the characteristics described above. In addition to simple regression approaches for individual risk factors (e.g. short-rate models), there are also sophisticated approaches to recreating changes in own funds under combined risk factor shifts (market, credit, underwriting and operational risks) whose specification requires a similarly high amount of processing power as training a neural network, which can be found in, among others, internal models used by insurance groups. Said specification can also include “training cycles”, i.e. the iterative addition or removal of polynomial building blocks using an information criterion in combination with subsequent retraining of the expanded/reduced function via regression.

The hypothesis space these models have is generally more limited than that of the aforementioned ML methods, which is why certain problems such as the black box characteristic occur less often or not at all. However, similar limitations can arise in relation to validation concerning the progressive plausibility of calculations, making one reliant even here on checking the plausibility of results, e.g. via out-of-sample comparisons.

### **Are supervisors more sceptical of unsupervised learning than supervised learning?**

No. The use case and type of ML are directly linked, as every type solves different problems and is therefore suited to different use cases. It is not possible to decide which type should

---

<sup>22</sup> See S. Russel, P. Norvig, 2016, *Artificial Intelligence: A Modern Approach*, third edition, p. 693.

<sup>23</sup> The focus here is initially on supervised learning. With unsupervised learning, the learning task consists of pattern recognition via systematising input data (without output or labels).

<sup>24</sup> See S. Russel, P. Norvig, 2016, *Artificial Intelligence: A Modern Approach*, third edition, p. 693 ff.; Bank of England, 2019, *Machine learning in UK financial services*.

be used based on the level of risk involved. Each individual use case should be categorised and appraised according to its characteristics. Fundamentally, banking and financial supervisors in Germany use the same evaluation criterion for all types of ML. Even so, use cases with supervised learning currently outnumber all others.

**Table 2: Types of ML**

Type
Supervised learning: The algorithm is trained to map input data to given output, also known as "label". The trained model can then be used on new data.
Unsupervised learning: Finding hidden patterns or structures (clustering, association, dimension reduction).
Reinforcement learning: Learning tasks during which the model's own solution strategies are repeatedly re-used and evaluated as input parameters.

## Selected examples of definitions of the term „Machine Learning“

**Table 3: Selected definitions of ML**

Academia	Regulators	IT companies
<ul style="list-style-type: none"> <li>▪ „Set of methods that can automatically detect patterns in data, then use the un-covered patterns to predict future data, or to perform other kinds of decision making under uncertainty.“ Murphy, S. 1</li> <li>▪ „Machine Learning is the science (and art) of programming computers so they can learn from data.“ Geron, S. 2</li> <li>▪ „Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.“ Samuel in Geron, S. 2</li> <li>▪ „A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E“ Mitchell in Geron, S.2</li> <li>▪ „The objective of machine learning is the generation of ‘knowledge’ out of ‘experience’ through learning algorithms using examples to develop a complex model. This model, and thus the automatically obtained knowledge representation, can in turn be used on</li> </ul>	<ul style="list-style-type: none"> <li>▪ „The standard on IT governance ISO/IEC 38505-1:2017 defines ML as a ‘process using algorithms rather than procedural coding that enables learning from existing data in order to predict future outcomes.’“ EBA, S. 14</li> <li>▪ „ML is a methodology whereby computer programmes fit a model or recognise patterns from data, without being explicitly programmed and with limited or no human intervention. This contrasts with so-called ‘rules-based algorithms’ where the human programmer explicitly decides what decisions are being taken under which states of the world.“ Bank of England, S. 6</li> <li>▪ „Machine learning may be defined as a method of designing a sequence of actions to solve a problem, known as algorithms, which optimise automatically through experience and with limited or no human intervention.“ FSB, S. 4</li> </ul>	<ul style="list-style-type: none"> <li>▪ „ML, a subset of AI, focuses on the ability of machines to receive data and learn for themselves without being programmed with rules. ML differs from traditional programming by allowing you to teach your program with examples rather than a list of instructions. [It] enables you to “train” an algorithm so that it can learn on its own, and then adjust and improve as it learns more about the information it is processing.“ Google</li> <li>▪ „ML is a form of AI that enables a system to learn from data rather than through explicit programming. [...] ML enables models to train on data sets before being deployed. Some ML models are online and continuous [leading] to an improvement in the types of associations made between data elements. Due to their complexity and size, these patterns and associations could have easily been overlooked by human observation.“ IBM</li> <li>▪ „ML is the process of using mathematical models of data to help a computer learn without direct instruction. It’s considered a subset of AI. ML uses algorithms to identify patterns within data, and those patterns are then used to create a data model that can make predictions. With increased data and experience, the results of machine learning are more accurate – much like how humans improve with more practice.“ Microsoft</li> </ul>

#### Sources of the ML definitions

1. A. Geron, 2019, “Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow”, 2. ed. O’Reilly.
2. T. Hastie, R. Tibshirani und J. Friedman (2009), “The Elements of Statistical Learning: Prediction, Inference and Data Mining”, 2. ed. Springer-Verlag.
3. G. James, D. Witten, T. Hastie und R. Tibshirani (2013), “An Introduction to Statistical Learning”, Springer-Verlag
4. K.P. Murphy, 2012, “Machine Learning: A Probabilistic Perspective”, MIT Press.
5. EBA, 2020, “Report on Big Data and Advanced Analytics”, available online: [https://eba.europa.eu/sites/default/documents/files/document\\_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf](https://eba.europa.eu/sites/default/documents/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf)
6. Bank of England, 2019, “Machine learning in UK financial services”.
7. Bundesbank, 2020, “Policy Discussion Paper, The Use of Artificial Intelligence and Machine Learning in the Financial Sector”, available online: <https://www.bundesbank.de/resource/blob/598256/d7d26167bceb18ee7c0c296902e42162/mL/2020-11-policy-dp-aiml-data.pdf>
8. Financial Stability Board (FSB), 2017, “Artificial intelligence and machine learning in financial services, Market developments and financial stability implications”, available online <https://www.fsb.org/wp-content/uploads/P011117.pdf>
9. Fraunhofer IAIS, 2018, “Maschinelles Lernen. Eine Analyse zu Kompetenzen, Forschung und Anwendung”, available online: [https://www.bigdata-ai.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/Fraunhofer\\_Studie\\_ML\\_2\\_01809.pdf](https://www.bigdata-ai.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/Fraunhofer_Studie_ML_2_01809.pdf)
10. Google, “Using AI for a social good”, available online: <https://ai.google/education/social-good-guide/?category=introduction>
11. IBM, 2020, “Data science and machine learning”, available online: <https://www.ibm.com/hk-en/analytics/machine-learning#:~:text=Machine%20learning%20enables%20models%20to,associations%20made%20between%20data%20elements>

12. Microsoft Azure, "What is Machine Learning?" available online:  
<https://azure.microsoft.com/en-us/overview/what-is-machine-learning-platform/>

# Impressum

## Herausgeber

Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin)  
Gruppe Kommunikation  
Graurheindorfer Straße 108, 53117 Bonn  
Marie-Curie-Straße 24 – 28, 60439 Frankfurt am Main  
[www.bafin.de](http://www.bafin.de)

## Redaktion und Layout

BaFin, Interne Kommunikation und Internet

Redaktion: Rebecca Frener  
Tel.: +49 0 228 41 08 22 13

Kathrin Jung  
Tel.: +49 0 228 41 08 16 28

Layout: Christina Eschweiler,  
Tel.: +49 0 228 41 08 38 71

E-Mail: [journal@bafin.de](mailto:journal@bafin.de)

## Designkonzept

werksfarbe.com | concept + design  
Humboldtstraße 18, 60318 Frankfurt  
[www.werksfarbe.com](http://www.werksfarbe.com)

## Bezug

Das BaFinJournal\* erscheint jeweils zur Monatsmitte auf der Internetseite der BaFin. Mit dem Abonnement des Newsletters der BaFin werden Sie über das Erscheinen einer neuen Ausgabe per E-Mail informiert.

Den BaFin-Newsletter finden Sie unter: [www.bafin.de/Newsletter](http://www.bafin.de/Newsletter).

## Disclaimer

Bitte beachten Sie, dass alle Angaben sorgfältig zusammengestellt worden sind, jedoch eine Haftung der BaFin für die Vollständigkeit und Richtigkeit der Angaben ausgeschlossen ist.

Ausschließlich zum Zweck der besseren Lesbarkeit wird im BaFin Journal auf die geschlechtsspezifische Schreibweise verzichtet. Alle personenbezogenen Bezeichnungen sind somit geschlechtsneutral zu verstehen.

\* Der nichtamtliche Teil des BaFin Journals unterliegt dem Urheberrecht. Nachdruck und Verbreitung sind nur mit schriftlicher Zustimmung der BaFin – auch per E-Mail – gestattet.

# Machine learning in risk models – Characteristics and supervisory priorities

## Responses to the consultation paper

# Content

I. Feedback from the banking and insurance sector . . . . .	3
II. Characteristics of ML. . . . .	4
III. Supervisory approach . . . . .	5
IV. Outlook . . . . .	8
Annex – Selected Responses . . . . .	9
Imprint . . . . .	18



# I. Feedback from the banking and insurance sector

In July 2021, BaFin and the Bundesbank published the consultation paper entitled “Machine learning in risk models – Characteristics and supervisory priorities” and collected feedback from the insurance and banking sector.

The consultation paper focused on solvency supervision, and specifically the application of machine learning (ML) methods in areas of particular relevance to supervisors. These included, first – as an exception to the principle that algorithms do not require supervisory approval – ML methods that are subject to prudential inspections and approval procedures, i.e. that are used in internal models for calculating regulatory own funds requirements (Pillar 1). Second, ML methods were considered relevant if they are used in risk management under Pillar 2.

The consultation received feedback from associations, including central associations of the banking and insurance industry, as well as from individual insurance corporations, banks, management consultancies and technology companies.

The consultation reveals that procedures based on machine learning technologies (ML procedures) are already used in many areas of application at present. According to the feedback, current applications focus on risk management (e.g. money laundering and fraud detection, analyses in credit processes), retail and investment products, product pricing, and individual applications in other risk areas. Thus far, there are only a few instances of ML methods being used in Pillar 1 risk models, although these methods are considered highly promising in some of the responses. ML methods are also used as support or challenger tools.

*ML has seen  
hardly any use  
in Pillar 1  
models thus far*

The following section summarises the responses to the consultation.

## II. Characteristics of ML

*Broad approval  
of the  
'characteristics  
approach'*

The consultation paper did not develop a universally applicable definition of ML methods, but instead oriented supervisory practice, inspection techniques and inspection intensity around whether and which ML characteristics exist in a specific methodology to be examined and the extent to which they are pronounced. From the perspective of BaFin and the Bundesbank, this differentiation is sufficient for supervisory purposes to assess models under Pillar 1 and Pillar 2.

The characteristics were grouped based on the three dimensions of the **AI/ML scenario**: the **methodology and data basis** collectively describe the complexity and thus the model risk associated with the ML procedure. The **use of the output** contains the importance of the procedure within risk management. The intensity of inspections is not guided by the distinction between in-house development and **outsourcing** or the underlying **IT infrastructure**.

The responses to the consultation reaffirm the approach presented. There is broad consensus that no explicit definition should be prescribed: first, any definition would be insufficient given the wide variety of procedures and their continual further development. Second, no clear distinction can be drawn between ML procedures and traditional ones. Likewise, respondents are in favour of supervisory action being oriented towards the employed procedures in a technology-neutral way. The focus should not be on generalised requirements, but on specific use cases. The AI/ML scenario corresponds to this technology-neutral and case-specific perspective.

In some cases, further possible characteristics were suggested. These include the effects of model deficiencies as well as performance as ML characteristics. Individual responses proposed characteristics that, taken in isolation, cannot be considered specific to ML. These include the data basis, which is characterised more by the availability of data than the method. Other suggestions were automation, adaptivity and complexity, which can sometimes be greater in traditional procedures than in ML.

According to the consultation participants, outsourcing and IT infrastructure should not be considered ML-specific characteristics, neither would they generally form part of the selection and assessment of the ML procedure, nor be subject to their own supervisory framework independent of ML. The feedback shows that no specific risks to IT implementation or outsourcing management arising from ML could be identified, even if the fundamental significance of cloud services and outsourcing is growing with ML.

## III. Supervisory approach

### 1. Supervisory practice endures

The consultation paper explained that Pillar 1 includes extensive regulations for reviewing and approving internal models formulated in a technology-neutral manner and therefore also addresses the risks of ML methods. Principles-based requirements for risk management and IT provide a sound footing in Pillar 2.

The responses to the consultation largely agree that current technology-neutral and risk-oriented regulation also provides a sufficient framework for ML. For Pillar 1 in particular, there is already a comprehensive regulatory framework that could be applied to ML, too. In some cases, the requirements regarding the comprehensibility and measurability of ML procedures are viewed critically.

Responses from the banking sector assess the ML-related requirements of EBA/GL/2020/06 to be appropriate, technology-neutral, and risk-oriented. Accordingly, no concerns are expressed regarding the implementation and application of regulations to all ML methods in Pillar 2, either.

*The supervisory regulatory framework is considered sufficient*

### 2. Methods invite to “believe in data”

According to the consultation paper, data quality is already an issue of key importance in supervisory action. Here, however, the characteristics of ML methods make clear that the underlying data should be viewed particularly as a starting point and as a factor for successful outcomes. Unstructured data can now be exploited by and for ML methods. Furthermore, ML methods allow for calculations to be made using a large number of variables, which can, in turn, lead to the problem of “overfitting”. This makes it easy for modellers to quickly apply ML to large datasets. When using large volumes of data, the quality of the data must be ensured on an ongoing basis; in addition to model development and validation, this also applies to model application.

The responses to the consultation unanimously state that all of the data employed must be as representative as possible, and that the use of ML methods would therefore not be grounds for any new or special requirements. The feedback highlights the large amount of manual effort and expertise required for the selection of suitable data: it is becoming more and more important to have an in-depth understanding of the data employed. Insurance companies in particular state that regulatory requirements restrict the volume of usable data, e.g. due to the requirement to focus on just a small number of singular events, or that trends make historical data unsuitable for forecasting. The risk of “believing in data” is reported as already being an issue; participants say that this could be counteracted by high degrees of explainability among the models and/or a wide variety of possible models.

The feedback indicates that data quality is already of key significance and should be considered as multifaceted. The responses suggest that software tools (including ML-based tools) could help not only to ensure the quality of the data bases that are already used, but also allow new data sources, such as unstructured data, to be utilised.

*Increased focus on data basis and data quality*

### 3. Focus on explainability

One point of discussion in the consultation paper concerns the explainability of models. As the hypothesis space that can be depicted by the model becomes more complex and more highly dimensional, it also becomes more difficult to describe the functional relationship between input and output verbally or using mathematical formulae, and the details of the calculations are less comprehensible for modellers, users, validators and supervisors. As a result, it is more difficult, if applicable, to verify the validity of the model output as well. User acceptance may also suffer. While this “black box” characteristic may be justified by its higher predictive ability, for example, it leads to potentially greater model risk. In order to take this into account, explainable AI (XAI) methods were developed. In addition, even if XAI methods may seem promising from a supervisory perspective for mitigating the “black box” characteristic, these approaches themselves constitute models with assumptions and weaknesses, and, in many cases, are still in their testing phases.

From the perspective of the respondents to the consultation, the reproducibility of calculation results must be ensured; in this context, they believe that the interpretability of model output can be improved using technological means. Explainability should form the basis for selecting the modelling framework. According to the consultation feedback, a vital task of model validation is to check the plausibility of the model, whereby respondents see the complexity of the model as also determining the complexity of the validation. A favourable view is held of models with outputs that are very stable over time and for which jumps in output can be explained by clearly identifiable one-off effects. Validation using appropriate datasets and methods is considered feasible for complex models, too. The majority of respondents believe that not every intermediate step needs to be explainable, but only the final output.

Most of the responses call for the trade-off between performance and explainability to be investigated for each use case. XAI requires such a significant amount of effort to implement that the model’s performance must be extraordinarily high in order to justify it. The responses range from “XAI does not fundamentally offer a way out of the black box dilemma” to “XAI solves the black box problem”. XAI is not seen as a panacea, as it only investigates selected model characteristics; its use as a basis for expert evaluation is recognised, however. In the opinion of respondents, it is not yet possible to set out supervisory standards for XAI, as the methods continue to develop at a rapid pace.

Respondents believe that downstream XAI can only be implemented on a sample basis. For this reason, they claim that XAI should be a component of validation before the model is brought into operation as well as during ongoing operation. XAI could also serve to identify key variables and thereby construct an almost causal model.

*Explainability is a central criterion in the application of AI/ML*

*Costs and benefits of XAI must be reconciled*

## 4. Adaptivity: Model changes are more difficult to identify

Institutions and enterprises are obligated to inform supervisors of any changes to Pillar 1 models and, if applicable, only implement these changes after they have been approved. There is no clear-cut distinction between regular model maintenance and model change, which continually leads to discussions with supervisors, especially as the term “model change” is also dependent on the prevailing supervisory context. The consultation paper provided multiple examples of this. However, the flexibility and, in some cases, high-frequency adaptivity of ML procedures can make it more difficult to draw a clear distinction between adjustment and change that would be indispensable for supervisory purposes. As a general rule, the need for high-frequency adaptivity should likewise be thoroughly justified. Even if a large number of ML procedures fall within the supervisory area of Pillar 2 in which approval is not required – and there is thus greater flexibility for retraining and changing the model – existing requirements (e.g. from MaRisk) remain applicable. From a supervisory perspective, it is nevertheless crucial, despite this flexibility, to adapt the training cycle to the specific use case, as well as provide the necessary justification, in order to create a balance between the explainability and validation of the model and ensuring that the data are up-to-date.

The consultation feedback highlights the conflict between model maintenance/calibration and model change, and points out the recurring issue of model drift. Supervisory authorities are requested to further refine the definition of model change as well as generally speed up model approval processes in order to avoid creating competitive disadvantages. Possible definitions and examples for model maintenance and model change are provided: for instance, training a procedure should not be considered model change (neither should the adjustments made to hyperparameters in the process), although the iterative data science process could include trigger points for model change.

Retraining is not observed to exhibit any fundamental particularities; however, depending on data availability, these may need to be clarified in individual cases. According to the feedback, there is no clear boundary between model maintenance and model change in this context, and material changes to the output must be explainable for this reason. The insurance industry in particular does not expect risk models to be retrained more frequently than traditional approaches, although more frequent retraining could be necessary for other areas of application, such as fraud detection. Improving the quality of the model is cited as a reason as to why retraining is required.

With regard to organisational adjustments to the use of ML methods, the responses illustrate the range of vastly different concepts, going as far as merging data science and modelling units. The general consensus is that the separation of modelling and validation should be retained, but the collaboration between the two intensified. Many responses express a desire for clearer model governance.<sup>1</sup>

*Model governance growing in significance; the definition of model change must be refined*

<sup>1</sup> For more information, see also the EBA consultation <https://www.eba.europa.eu/eba-consults-machine-learning-internal-ratings-based-models>

## IV. Outlook

The results of the consultation are the foundation for starting dialogue with enterprises. Alongside the general BDAI principles, this is intended to achieve clarity in the development and application of ML methods in the context of models relevant for supervision in Pillars 1 and 2.

Further discussion is required especially in those areas where the consultation reveals a broad range of opinions, for example with regard to the issues of model explainability and model adaptivity. Certain combinations of characteristics may require more attention from supervisory authorities than others. Supervisors are focused on preserving the ability to monitor risk models.

The ongoing exchange should ensure that there is clarity regarding supervisory expectations and that these expectations are included in the existing Pillar 1 and Pillar 2 frameworks in a technology-neutral way. This will provide enterprises with a regulatory environment that enables them to invest in ML methods and address the risks of these methods as early as possible.

In order to harmonise international approaches as far as possible and establish identical cross-sector criteria for the use of ML methods, the results of the consultation will be incorporated into the implementation of the European Commission's Digital Finance Strategy<sup>2</sup> and discussed with other European supervisory authorities.

---

<sup>2</sup> European Commission (2020), "Digital Finance Strategy for the EU", available online at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0591>

# Annex – Selected Responses

## 1. Characteristics of ML

- Do you think it is appropriate to forgo a strict definition of ML methods and instead take an application-based approach and gear supervisory and inspection practice to the individual characteristics of the methods used?

*“We don’t consider a strict definition of ML methods to be appropriate.”*

*“We welcome the suggestion to gear supervisory and inspection practice to the individual characteristics of the methods used. This approach allows us to tier the supervisory and inspection practice between established statistical methods and new ML methods.”*

*“Owing to the large number of approaches and research groups formed within the ML environment and currently in development, a single definition seems to be unrealistic on the one hand and not that helpful on the other.”*

*“The strict definition of model characteristics makes no sense, as the ML methods represent only one algorithm for calculating model functions.”*

- What other characteristics of ML methods do you believe could be important for supervisory practice or for internal model governance?

*“How to deal with open source software, which is increasingly available for ML methods, should be clarified.”*

*“The model’s performance is another important feature. [...] We also consider plausibility to be an important characteristic.”*

*“We believe reproducibility is a key aspect or characteristic as well, in conjunction with explainability and adaptability.”*

- In your opinion, which characteristics do not belong in this overview?

*“Data sources and data volumes play only a secondary role for us in differentiating between conventional and modern ML methods.”*

*“The interpretability of the output is an important point which, in our view, is taken into account in your commentary. It is therefore not necessary to include this as a separate characteristic.”*

*“Outsourcing processes is a scenario that is generally possible but also already heavily influenced by regulatory requirements. Transparency and governance play a relevant role and this has to be evaluated regardless of whether ML methods are deployed.”*

- In which relevant areas of application do you employ ML methods or where do you intend to implement them?

*“Some enterprises have already integrated first ML methods as part of validation of the internal model. As far as we know, neither neural networks nor other ML methods are being used yet for official risk capital calculations. [...] These models are already used in distribution and in general operational processes. There are plans to develop elements of the internal risk model further.”*

*“The actual areas of application are currently limited, although use cases can be found throughout the bank’s operations.”*

*“In many areas we don’t see machine learning methods being used at present, but more ‘conventional’ procedures or (scientific) simulation approaches instead.”*

*“Currently, ML methods are only used to accompany the development of traditional risk models as challenger models.”*

## ■ 2. Supervisory approach

### 2.1 Supervisory practice endures

- In your opinion, do existing regulations already contain prudential requirements that appear to hinder the use of ML methods? Do you believe that contradictions will arise between prudential regulations for Pillar 1 and 2 models and the draft AI Regulation? Please state any relevant references to the corresponding regulations and explain the challenges.

*“Current inspection and application practice requires a degree of transparency and measurability that does not appear to be available.”*

*“We share the view of BaFin and the Bundesbank that, owing to its technological neutrality, current regulation and inspection practice already offers an excellent framework for employing ML in risk models.”*

*“We believe there is no need for concern regarding existing supervisory requirements when using ML methods. The draft AI Regulation is also in line with the regulations for Pillar 1 and Pillar 2 models, and no direct contradictions have been identified.”*

*“We consider it very important not to hamper the competitiveness of the insurance industry through excessive regulation.”*

*“Currently, we have not found any regulations that would prevent the use of ML methods in principle.”*

- To what extent do you believe the requirements laid out in EBA/GL/2020/06 with reference to the use of automated models in creditworthiness assessments and credit decision-making are also suitable for other ML methods in Pillar 2 (MaRisk) and should be taken on?

*“The scale of the relevant requirements laid out in chapter 4.3.4 of EBA/GL/2020/06 for the use of automated models in creditworthiness assessments and credit decision-making have been deemed as appropriate.”*



“The requirements under EBA/GL/2020/06 with reference to the use of automated models in creditworthiness assessments and credit decision-making should provide a sufficient safeguard. The required measures generally provide a good starting point for applying ML methods in practice.”

- Are there any other points where you believe current supervisory practice requires adjustment in order to appropriately acknowledge ML procedures and their associated risks?

“No.”

*“The protection of training, validation and test data as well as the AI application’s source code all play an important role in ML processes. [...] Other points that could be relevant to supervisors include a potentially strongly changing data scope, the volatility of the output and a lack of documentation on algorithms.”*

*“We do not believe that there are any other points where existing supervisory practice requires adjustment.”*

*“We consider this to have been covered by training/expertise/standards and directives.”*

- Do ML methods entail specific risks for IT implementation and outsourcing management? Are “adversarial attacks” conceivable in the financial sector and should ML methods be given particular protection against such attacks?

*“From the perspective of numerical analysis, the implementation of ML algorithms has a strong analogy to previously used methods, meaning that there is no need to reclassify any other specific IT risks besides the aforementioned data protection topic (existing risks may need to be adjusted if necessary).”*

*“We see no risks specific to ML for IT implementation. The current requirements already apply to conventional ML models. [...] Adversarial attacks are generally possible on all ML systems. The degree of protection is derived from the sensitivity of the data represented in the model and the relevance of the processes supported by the ML system.”*

*“As the group of people with access to the data streams relevant to risk calculation is generally limited and clearly defined, and a large number of controls and checks have been set up, we do not consider there to be a greater risk from adversarial attacks. Furthermore, and regardless of the methods used, the significance of IT risks is growing due to the rise in data exchange and processing.”*

## 2.2 Methods invite to „believe in data“

- What challenges do you see when selecting data and when ensuring data quality with regard to ML methods?

*“Data credibility is not just a phenomenon for ML/AI but also exists in the conventional environment.”*

*“The requirements for data quality are not specific to certain methods and have already been addressed previously in model development processes. Data modellers’ expertise and experience are crucial when selecting data and ensuring data quality, especially for complex models. We do not believe ML methods simplify the selection and quality of data.”*

*“It is important that the datasets used match the real data as far as possible to ensure that they remain representative for the “use case”. The more that data are anonymised (primarily for data protection reasons), the further datasets and the models trained on them become removed from the real data, which may lead to modified ML models and additional risks.”*

*“The possibility of using unedited raw data in many ML methods and the efforts of many institutions to automate processes across all areas may tempt them to include variables in models without any manual interaction regardless of the underlying data quality or – even more worryingly – without checking the meaningfulness of the variables to the underlying problem.”*

*“When using methods subject to strict data protection, such as machine learning, critical questions should be asked as to whether the data are actually suitable for extrapolating scenarios relevant for capital requirements.”*

- In your opinion, what aspects of data quality are made easier through the application of ML methods?

*“Quality assurance and data protection are already in focus.”*

*“We consider it advantageous if many data (including metadata) are used for AI/ML, as they can strengthen the robustness of ML models. ML methods can be used to intelligently replace missing datasets and thereby improve the model’s quality.”*

*“The broader the dataset used, the lower the model’s expected loss of performance should a feature/data source ‘fail’.”*

*“ML methods are capable of identifying disturbances or distortions in the data to a certain extent and of calculating meaningful output even when the data are incomplete or disrupted. Nevertheless, the data preparation process prior to the actual process of developing the model is extremely important, requiring a consistently high level of quality.”*

### 2.3 Focus on explainability

- In your opinion, what impact does the black box characteristic have on the validation of the procedure?

*“The black box characteristic is based on the model’s high level of complexity, meaning that causality relationships cannot be shown transparently. However, we assume that a result which is reproducible rather than driven by randomness must be guaranteed going forward when using ML methods. The interpretability of the output can be improved through suitable methods (e.g. a sensitivity analysis).”*

*“For good reason, pure black box models that do not attain at least a basic level of explainability using XAI techniques should never be used, especially in the context of risk models.”*

*“As long as the model output is also stable and verifiable over time, and one-off effects (temporal and technical) have been sufficiently trialled, the black box characteristic is viewed as unproblematic.”*

*“The “black box concept” is not correct in this respect. We believe a better image would be a complex clock mechanism in which it is only clear which cog is touching which, but the entire mechanism and the interactions within it cannot be readily understood. In a similar manner, an ML model can be approximated locally by using a simple linear model. The validation of the black box characteristic can be made easier by using XAI approaches.”*

*“The effect of the black box characteristic on validation should be borne in mind for each individual case.”*

*“ML/AI-based output or the suitability of methods have to be explainable or plausible for specialists but not every individual operation or individual interim results. Even conventionally produced statistics are not meaningful when viewed in isolation, their output and methods have to be interpreted first.”*

*“When using modern ML methods, the process of making the black box explainable (opening the black box) should be both part of the initial validation process directly after the model has been developed and part of the regular validation process while the model is up and running.”*

- How important is the trade-off between performance and explainability for you?

*“Depending on the use case, either the explainability or the performance of the model can be deemed more important. The real challenge lies in adapting the complexity in the ML method to the complexity of the problem. We don’t see a trade-off between explainability and performance.”*

*“Weighing up performance and explainability is always a good idea when selecting a model irrespective of whether it is an ML model or a conventional model.”*

*“The trade-off between performance and explainability is very important to practice and this has to be weighed up depending on the application.”*

- Do you believe XAI methods (always) offer a way out of the black box dilemma? Which processes are the most promising and for which ML methods?

*“As XAI techniques can indeed require considerable effort to validate conventional models in practice, this effort should be offset by a significant gain in performance.”*

*“In our view, XAI techniques are not a way out of the black box. Various XAI techniques can only highlight those features that were primarily crucial to a prediction.”*

*“XAI techniques are suitable tools for enhancing the interpretation of models.”*

*“It remains for experience to show which methods are best suited to assessing the behaviour of algorithms. The development of other methods to analyse algorithms’ behaviour is a dynamic area currently – especially for algorithms from the BDAI area. Accordingly, we believe it is too early to look at setting standards for this area.”*

*“Every approach harbours the risk of trusting the output too readily. In any case, focusing on models and input data as well as intensively verifying the output of XAI methods will be essential.”*

- How do you think an XAI integrated downstream of the method should be incorporated into the validation process?

*“A downstream XAI can be used to check whether a model’s decision was correct. [...] This is just a manual check that could be employed at random in order to document that a certain quality has been attained.”*

*“The regular application of XAI techniques is recommended in order to understand the ML models better. XAI techniques can be included both during development and as an ex post tool.”*

#### **2.4 Adaptivity: model changes are more difficult to identify**

- What questions on supervisory practice do you see arising with regard to model adjustments for ML methods?

*“Updating the data or changing the (hyper)parameters that result solely from training the model in ML models do not constitute a model adjustment as these instead represent components of applying the model under the current market/business conditions.”*

*“Even previously, a new calibration, provided that it met the established quality standards, did not necessarily constitute a model change. It would therefore be reasonable to state that there is no model change for ML methods either, provided that they meet certain quality standards. This too, of course, depends on the application.”*

*“The existing regulatory framework and model approval are sufficient; they do not require extension. Banks guarantee certain aspects through their individual governance.”*

*“Supervisors should outline clear criteria to delineate model maintenance from model changes subject to approval.”*

*“Training (including tuning hyperparameters) is a regular process and should not be classified as a model change subject to supervisory regulations. It is part of the (annual) recalibration.”*

*“Shouldn’t changes in the weights of the information types used rather than changes to the risk factors used be considered when distinguishing model maintenance from model changes? Switching a feature should not necessarily be recorded as a model change if the new feature provides very similar information.”*

- Do you believe it is necessary for certain ML methods to be retrained at very high frequencies?

*“Regular model retraining is considered useful regardless of the methodology chosen, i.e. even in the case of conventional regression models.”*

*“It should also be borne in mind that “data science” consists of iterative cycles. The question arises as to the point at which a cycle is considered completed, and when a new one begins. This could then be a trigger point for a model adjustment by the supervisory body. We believe that retraining at very high frequencies is not required. The driving force behind changes to model output tends to be the underlying data and their volatility/expansion.”*

*“The models are recalibrated regularly, meaning that the distinction between model maintenance and a model change subject to supervisory assessment is already relevant nowadays. The necessity of frequent retraining strongly depends on the use case in question. It should not be more frequent than for existing models. Furthermore, model changes for ML methods should not focus solely on individual parameters, meta parameters or the model’s output, but on the learning process in its entirety, including all meta parameters, i.e. optimisation across various modelling approaches should be viewed as part of the learning process. Only when a learning process has been modified should it be classified as a model change.”*

*“As such, retraining is only carried out when a higher level of quality is expected as a result. Retraining is not continued if the retrained model does not show a higher quality level.”*

*“For Pillar 1 and Pillar 2 applications, we believe that retraining would not be necessary that often under normal circumstances (unlike applications that include customer service, for example).”*

*“For methods in particular that require retraining the models frequently, we believe it is important to ensure that this does not produce any major jumps in the output variables. If jumps do occur, this should be identifiable and justifiable in the model monitoring.”*

- Do you see ML methods necessitating changes in model governance?

*“Due to the technological neutrality of the regulations and supervisory practice, the use of ML methods in internal models does not generally require changes to the supervisory approach.”*

*“In model governance such as this, all parties are involved, as the entire process from data load to model deployment has to be explained and documented so that each party can understand what is input for the model and what output it provides.”*

*“Model governance can remain unchanged.”*

*“We believe that ML methods will cause pre-existing aspects of model governance to become more clearly apparent than before.”*

- How do traditional modelling units, validators and new “data science” units work together?

*“The purpose of the validation function is to perform an independent review of the models, processes and calculations. This should also include reviewing the current methodology. ML methods can be used by both modelling units and validation units. Ideally, there should be a lively exchange between these two units, and this may also include other “data science” units (if these have already been implemented).”*

*“The time-tested division of labour between modelling and validation should be maintained. We do not consider it necessary to make adjustments. That said, traditional modelling units should not work separately from new “data science” units. Generally, a common understanding of the new methods should be cultivated on the development and validation side.”*

*“Data science and modelling units will merge with one another when an ML model is used productively, while validators will continue to act independently as before. It is conceivable that knowhow for successful validation will increasingly be transferred via the data science unit.”*

*“Regardless of the ML methods, close cooperation between modelling units and validation units is recommended. The validation unit should constructively review the appropriateness of the developed models, processes and output. [...] “Data science” is a new field for some insurance companies, which means that their own “data science” teams may not necessarily be firmly anchored yet. In any case, there should be a regular exchange of views between all departments regarding best practices and areas of application (e.g. as part of “use cases”).”*

# Imprint

## **Publisher**

Deutsche Bundesbank  
Postfach 10 06 02, 60006 Frankfurt am Main  
[www.bundesbank.de](http://www.bundesbank.de)

Federal Financial Supervisory Authority  
(Bundesanstalt für Finanzdienstleistungsaufsicht – BaFin)  
Graurheindorfer Straße 108, 53117 Bonn  
Marie-Curie-Straße 24-28, 60439 Frankfurt am Main  
[www.bafin.de](http://www.bafin.de)

## **Disclaimer**

Please note that great care has been taken in compiling all of the information contained herein. However, BaFin accepts no liability for the completeness and accuracy of this information.

BaFin Journal uses gender-neutral language solely for the purpose of improving readability. All personal designations are therefore intended to be gender-neutral.

## Finanstilsynet

Sendt pr. e-mail: CNE@ftnet.dk

Att.: Camilla Neuenschwander.

Den 27. juni 2023

### Hørings svar – udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Den danske Fondsmæglerforening takker for muligheden for at komme med input til Finanstilsynets udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor.

#### Generelle bemærkninger

Den danske Fondsmæglerforening er meget positiv over for brugen af kunstig intelligens til at skabe værdifulde løsninger for virksomhederne, kunderne og samfundet. Foreningen anerkender også fuldt ud, at det kan være hensigtsmæssigt, at virksomheder forholder sig til de etiske dilemmaer, der kan opstå ved brugen af kunstig intelligens.

Den danske Fondsmæglerforeningen finder det derfor også positivt, at Finanstilsynet har til hensigt at inspirere virksomheder til at arbejde med dataetik.

#### Specifikke bemærkninger

##### Understøtte brugen af AI

Det er meget positivt, at Finanstilsynet generelt ønsker at **understøtte brugen** af AI i den finansielle sektor og specifikt nævner, at AI kan være en stor fordel for både virksomheder, kunder og samfundet. Det er derfor også særdeles vigtigt, at papiret udformes med dette for øje, så danske virksomheder kan være konkurrencedygtige i forhold til virksomheder i andre lande og udnytte de digitale fordele, som vi har i Danmark. Der kan derfor med fordel tilføjes et afsnit under pkt. 2 'Baggrund for Finanstilsynets fokus på dataetik', hvoraf det fremgår, at det kan have alvorlige konsekvenser af at undlade at bruge de nye teknologier.

Under pkt. 3 'Sammenfatning' fremgår følgende: *'Virksomhederne kan have mange fordele af at bruge AI, men det er vigtigt, at virksomheder under Finanstilsynets tilsyn tager arbejdet med dataetik alvorligt'*. Budskabet kan med fordel ændres, så det ikke fremstår som om, at der er et modsætningsforhold mellem brugen af AI og dataetik.

---

#### Den danske Fondsmæglerforening

Børsen  
1217 København K  
Telefon 3374 6000  
[www.fondsmæglerforeningen.dk](http://www.fondsmæglerforeningen.dk)  
[info@fondsmæglerforeningen.dk](mailto:info@fondsmæglerforeningen.dk)





Under afslutningen i pkt. 9 fremgår det, at *'Det øger også betydningen af dataetikken, når flere virksomheder bruger AI og derved gør behandlingen af data mindre gennemskuelig'*. Den danske Fondsmæglerforening finder det uheldigt, at det fremgår som om, at det er en naturlig følge af brug af AI, at behandlingen af data bliver mindre gennemskuelig. Dette behøver ikke at være tilfældet.

### Inspiration

Den danske Fondsmæglerforening finder det positivt, at Finanstilsynet ønsker at **inspirere** virksomhederne. Det er afgørende, at denne hensigt forfølges gennem hele papiret, da der – som det også fremgår – *ikke* er tale om regulatoriske krav. Det vil heller ikke være hensigtsmæssigt med danske særregler på dette område. Det kan i den forbindelse være hensigtsmæssigt at ændre overskriften, så *'God praksis'* ikke indgår. Betegnelsen *'god praksis'* kan opfattes som en praksis, der bør følges af virksomhederne, da det modsatte kan opfattes som *'dårlige praksis'* eller *'mindre god praksis.'* Papiret kunne f.eks. hedde *'Inspiration til dataetiske overvejelser ved brug af AI i den finansielle sektor'*. Det kan anbefales at gennemlæse papiret og sikre, at der ikke er eksempler og sprogbrug, der kan indikere, at der er tale om andet end inspiration.

Som eksempel på dette kan fremhæves sætningen under afsnit 4 under pkt. 5 *'Afgrænsning'*, hvoraf det fremgår: *'Der er tale om rent vejledende principper og formidling af, hvad der kan blive god praksis i fremtiden i takt med, at dataetik får stigende betydning.'* Den sidste del af sætningen kan med fordel udgå, da det kan læses som om, at der *er* tale om principper, der har – eller kan få – karakter af mere end inspiration.

Den danske Fondsmæglerforening er i forlængelse af dette meget enig i, at virksomhederne selv skal have frihed til at definere, hvad god dataetik er, og hvordan de vil håndtere eventuelle dilemmaer, som det fremgår af side 8 nederst. Der kan være mange individuelle forhold, der spiller ind i virksomhedernes drøftelser og håndtering, herunder strategi, produkter, kunder, kommunikation m.v.

Et andet eksempel er på side 15, hvoraf det fremgår omkring prisvandring, at Finanstilsynet ville *'i det mindste have vurderet adfærden som uetisk'*. Denne sætning bør udgå, da det er u hensigtsmæssigt, at Finanstilsynet i dette inspirationspapir vurderer adfærd som værende etisk eller uetisk.

### Detaljegraden

Den danske Fondsmæglerforening finder det – som nævnt ovenfor – meget positivt, at Finanstilsynet gerne vil inspirere virksomhederne til at drøfte dataetik. Det er dog vores opfattelse, at afsnit 8 bliver for konkret og peger på bestemte måder at agere på i virksomhederne.

Som eksempler på dette kan nævnes, når der tales om den rette kultur – og hvordan denne skal skabes. Et andet eksempel er budskabet om, at det er vigtigt at inddrage alle medarbejdere for at skabe en fælles sund kultur. Et tredje eksempel er, når det fremgår: *'Hvis en virksomhed beslutter sig for at arbejde seriøst med dataetik, er det vigtigt, at den forankrer opgaven og udpeger en ansvarlig medarbejder eller funktion for bl.a. at sikre, at opgaven bliver prioriteret.'*

---

#### **Den danske Fondsmæglerforening**

Børsen  
1217 København K  
Telefon 3374 6000  
[www.fondsmæglerforeningen.dk](http://www.fondsmæglerforeningen.dk)  
[info@fondsmæglerforeningen.dk](mailto:info@fondsmæglerforeningen.dk)



Vi vil opfordre Finanstilsynet til at tage ovennævnte input med i det videre arbejde med inspiration af virksomheder – både til at benytte AI og forholde sig til dataetik.

Med venlig hilsen

Den danske Fondsmæglerforening

Marianne Settnes  
Formand

---

**Den danske Fondsmæglerforening**

Børsen  
1217 København k  
Telefon 3374 6000  
[www.fondsmaeglerforeningen.dk](http://www.fondsmaeglerforeningen.dk)  
[info@fondsmaeglerforeningen.dk](mailto:info@fondsmaeglerforeningen.dk)



Finanstilsynet  
Strandgade 29  
1401 København K

Dato 27.6.2023

Sendt til email: Camilla Neuenschwander, CNE@FTNET.DK

## **Dataetisk Råds høringssvar på Finanstilsynets udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor**

### **Baggrund for afgivelse af høringssvar**

Dataetisk Råd er et regeringsnedsat uafhængigt forum, som bl.a. skal rådgive Folketinget, ministre og offentlige myndigheder om dataetiske spørgsmål ved brugen af data og digital teknologi. Rådet forholder sig herunder til etiske spørgsmål om fordelene ved anvendelse af ny digital teknologi og hensynet til retssikkerhed, samfundsmæssige værdier og borgernes grundlæggende rettigheder.

Dataetisk Råd finder det relevant at afgive høringssvar til Finanstilsynets udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor som led i rådets opgave med at levere sparring til relevante myndigheder, og idet udgivelsen netop angår dataetik og kommer fra en central myndighed på erhvervsområdet.

### **Dataetisk Råds bemærkninger til papir om god praksis for dataetik ved brug af AI i den finansielle sektor**

Dataetisk Råd hilser Finanstilsynets papir om god praksis for dataetik ved brug af AI i den finansielle sektor velkommen. Der er stigende bevidsthed i befolkningen om digital ansvarlighed blandt andet i lyset af den hastige teknologiske udvikling og udbredelse af kunstig intelligens mv. Rådet finder det derfor særdeles positivt med en udgivelse, der sætter fokus på dataetik, og hvordan etisk brug af data kan hjælpe til at sikre tilliden fra kunder såvel som borgere til en sektor af bred betydning for samfundet som den finansielle.

Rådet bifalder også Finanstilsynets initiativ til en sektororienteret anvendelse af dataetikken, der bringer dataetiske principper i spil med case-specifikke eksempler. Udgivelser som denne kan efter Dataetisk Råds opfattelse være et vigtigt skridt på vejen mod en større dataetisk bevågenhed på tværs af erhvervslivet. Finanssektoren har flere eksempler på virksomheder, der allerede arbejder seriøst med dataetik og digital ansvarlighed, og papiret kan bidrage til at gøre dataetikken relevant og vedkommende for endnu flere i den finansielle sektor.

Dataetisk Råd bifalder tillige papirets centrale anbefaling, at virksomheder bør tage arbejdet med dataetik seriøst og tage stilling til, hvordan de som virksomhed bruger data – eventuelt i form af en dataetisk politik. Erfaringsmæssigt kan udarbejdelse af en dataetisk politik være et rigtig godt udgangspunkt for at arbejde med dataetik i virksomheder.

Dataetisk Råd har nedenfor formuleret nogle overordnede bemærkninger til, hvor papiret med fordel kan udbygges for at skabe et endnu bedre afsæt for finansvirksomheder, der vil arbejde med emnet. Papiret nævner selv, at dataetik kan være 'et vanskeligt begreb at arbejde med', og Dataetisk Råd håber, at bemærkningerne kan understøtte Finanstilsynets arbejde med papiret:

## **1. Definitionen af dataetik som begreb.**

Dataetisk Råd værdsætter papirets nuancerede tilgang til dataetik som en særlig tankegang og refleksion, der kan fremmes i virksomheder. Dette er på linje med rådets eget arbejde for at bringe dataetik i øjenhøjde for virksomheder. Det prisværdige formål kan dog efter rådets opfattelse styrkes igennem en mere klar definition og anvendelse af dataetik som begreb.

Det anføres i papiret, at der ikke er nogen 'universel' definition af dataetik. Det er korrekt, at der i eksempelvis fagkredse er forskellige opfattelser af, hvad man skal forstå ved udtrykket dataetik. Men en faglig diskussion behøver ikke at skygge for en mere klar stillingtagen til, hvordan Finanstilsynet mener, at begrebet skal forstås i virksomhedernes arbejde med det. En klar definition vil være en stor hjælp for, at virksomheder skal kunne arbejde med dataetik i praksis. På den ene side nævnes det i papiret, at dataetik handler om, hvad man *bør* gøre i en given situation – med andre ord en forståelse af etik, som kan opfattes sådan, at rigtigt og forkert kan defineres uanset den specifikke

kontekst. På den anden side nævnes, at 'moral og normer varierer afhængigt af person, virksomhed, situation og kultur'. Her defineres etik som afgrænset af det folk *opfatter*, man bør gøre i den givne situation uden nærmere kriterier for, hvordan én opfattelse skal afvejes imod en anden. Rådet er enig i, at konteksten har en betydning for, hvad der opfattes som rigtigt og forkert. Men rådet finder det centralt at fastholde, at der er visse typer af dataindsamling og behandling, som *ikke* er dataetisk forsvarlige – også selvom de opfattes sådan af (nogle) folk i visse sammenhænge – hvis dataetikken skal kunne bruges som rettesnor for refleksioner over egen brug af data. I den forbindelse finder Dataetisk Råd det tillige væsentligt, at virksomheder i den finansielle sektor bør inddrage dataetikken ved alt databehandling, og ikke kun snævert ved anvendelse af kunstig intelligens. Dataetisk Råd henviser her til det arbejde, som rådet lavede ved sin etablering, der afdækkede hvad man skal forstå ved dataetik<sup>1</sup>.

## **2. Anvendelsen af konkrete dataetiske principper og værdier såsom fairness, bias, privacy mm til at opstille en systematisk og anvendelig definition af dataetik i forbindelse med AI.**

I tillæg til formuleringen af en tydeligere definition af dataetik opfordrer Dataetisk Råd Finanstilsynet til at overveje i højere grad at uddybe, hvordan dataetik hænger sammen med konkrete principper og værdier som bias, fairness, privacy mm. Sådanne principper og værdier er naturligvis komplekse og skal forklares og foldes ud i en kontekst af den finansielle sektor. Men sådanne principper og værdier kan være et godt redskab til at skabe en systematisk forståelse af, hvad dataetik betyder, fordi det giver lejlighed til at ekspliciterer hvad dataetik kan indebære mere præcist. Ved anvendelse af principper og værdier kan dataetik udfoldes til et begreb med substans, hvilket – alt andet lige – forhindrer, at diskussioner om dataetik bliver for abstrakte og derved uhåndterlige i praksis.

Papiret nævner flere gange sådanne principper, herunder i reference til Finanstilsynets papir om 'God praksis ved brug af superviseret machine learning', og principper anvendes løbende til at vurdere eksemplerne. Men brugen kan efter rådets opfattelse

---

<sup>1</sup> Dataetisk Råds analyse af 'Dataetisk metode & teori' kan findes [her](#).

med fordel systematiseres og uddybes som led i en klar definition af dataetik. I publikationen 'Dataetik – Sådan gør du', som papiret i øvrigt henviser til, har Dataetisk Råd udvalgt ti principper og værdier, som Finanstilsynet kan lade sig inspirere af.<sup>2</sup>

En klar og systematisk begrebsafgrænsning af dataetik vil være et godt afsæt for virksomhedernes formulering af en politik for dataetik, som netop er et hovedformål med Finanstilsynets papir. Ligeledes kan konkrete principper bidrage til at undgå, at en politik for dataetik bliver så overordnet, at det ikke reelt kan efterprøves, om virksomheden efterlever den, hvad enten efterprøvelsen sker ved virksomheden, hos kunderne eller ved (ekstern) revision.

### **3. Kobling mellem dataetiske overvejelser og konkrete handlinger, der sikrer dataetik i praksis.**

Papiret beskriver overvejende dataetikken som et refleksionsværktøj til at vurdere den etiske dimension af en given databehandling. Hvordan etiske dilemmaer herefter håndteres er der mindre fokus på. Efter Dataetisk Råds opfattelse kan dataetik med fordel beskrives og anvendes som et praktisk redskab til at sikre etisk brug af data i konkrete situationer, som også er målet med papiret.

Denne tilgang er i tråd med Dataetisk Råds perspektiv på dataetik, som udfoldes i rådets publikation 'Dataetik – Sådan gør du'. Her er dataetiske refleksioner kun første skridt i en struktureret proces for dataetisk praksis. De næste skridt handler om at pege på, hvordan man kan anvende data på en måde, som reducerer eller helt fjerner etiske betænkeligheder.

Værdien af denne tilgang kan ses ved at undersøge papirets eksempel A. I dette eksempel overvejer et forsikringselskab, om det er dataetisk at bruge lokationsdata indsamlet med samtykke til at tilbyde kunderne skræddersyede forsikringer – eksempelvis rejseforsikringer, hvis de pludselig dukker op i Østrig i vinterferien. Det er et relevant eksempel at fremdrage, og i den pågældende situation vil en ligetil løsning f.eks. kunne være at spørge kunderne om deres præferencer – ønsker de at selskabet skal give dem uopfordrede tilbud om ekstra dækning på baggrund af data, herunder lokations-

---

<sup>2</sup> Find 'Dataetik – Sådan gør du' og relateret materiale [her](#).

data? Og hvordan kan man indrette en samtykkeformular hertil, der indhenter et tilstrækkeligt informeret samtykke, således at kunden ikke finder det grænseoverskridende, at forsikringsselskabet bruger data på denne måde?

Refleksionerne bliver med udvidelsen af eksemplet brugt som afsæt til at træffe velovervejede beslutninger og i sidste ende guide imod konkrete løsninger. Sidstnævnte er vigtigt, idet man ofte kan opnå det samme formål med en given databehandling på flere forskellige måder. Således gøres dataetikens indgangsvinkel til et spørgsmål om, *hvordan* man kan foretage databehandling på etisk vis fremfor en vurdering af, *hvorvidt* en given databehandling er forsvarlig eller ej.

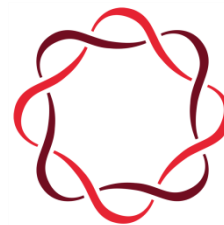
Med en sådan mere eksplicit sammenhæng mellem refleksioner og adfærd kunne papiret løfte dataetik ud af primært at være en refleksionsorienteret øvelse. De mange gode eksempler ville derved efter rådets vurdering i højere grad få den tiltænkte virkning – nemlig at anspore til digital ansvarlighed på et konkret niveau.

Samlet set finder Dataetisk Råd mange positive takter i Finanstilsynets papir. Rådet støtter ønsket om at fremme dataetik blandt finanssektorens virksomheder og håber, at Finanstilsynet vil finde rådets bemærkninger konstruktive som udtryk for et oprigtigt ønske om, at finansvirksomheder får de bedste forudsætninger for arbejde med dataetik og netop føre papirets anerkendelsesværdige ambitioner ud i livet.

Dataetisk Råd står til rådighed for uddybning og yderligere rådgivning.

Med venlig hilsen på vegne af Dataetisk Råd.

Johan Busse  
Formand  
Dataetisk Råd



# Høring over dataetik ved brug af AI i den finansielle sektor

## Resumé

---

### Resuméoverskrift

Finans Danmark er optaget af, at tilsynets vejledning om kunstig intelligens er tilpasset den kommende europæiske regulering, og udstedes på grundlag af fælles europæiske regler.

Finans Danmark foreslår, at tilsynet afventer den endelige tekst til AI-forordningen, EDPB-vejledning for AI samt det frivillige EU-USA baserede adfærdskodeks for dernæst at vurdere, om der er behov for, at man udsteder en særskilt dansk vejledning, der supplerer den allerede eksisterende danske best practice for brug af superviseret machine learning.

---

## Høringsvar

26. juni 2023  
Dok: FIDA-1232191906-51062-v1  
Kontakt Mette Stürup



# Høring over god praksis for dataetik ved brug af AI i den finansielle sektor

Tak for muligheden for at kommentere på Finanstilsynets "God praksis for dataetik ved brug af AI i den finansielle sektor".

Finans Danmark er optaget af, at tilsynets vejledning om kunstig intelligens er tilpasset den kommende europæiske regulering, og udstedes på grundlag fælles europæiske regler. Der foregår i øjeblikket drøftelser mellem EU og USA om en frivillig adfærdskodeks for kunstig intelligens for at sikre ensartede principper på tværs af EU og USA.

Endvidere er Datatilsynet på vej med en kommende EDPB-vejledning for AI. Finans Danmark foreslår derfor, at tilsynet afventer den endelige tekst til AI-forordningen, EDPB-vejledning for AI samt det frivillige EU-USA baserede adfærdskodeks for dernæst at vurdere, om der er behov for, at man udsteder en særskilt dansk vejledning, der supplerer den allerede eksisterende danske best practice for brug af superviseret machine learning."

Vi savner desuden præcisering af hvilken lovhjemmel, som den er baseret på. Best praksis fra 2019 guidance, fokuserer på bankernes risikostyring og brug af machine learning der. Risikostyring ligger indenfor institutternes forpligtelser efter den finansielle lovgivning. Dette dokument fokuserer derimod ikke på brug af AI i risikostyring og de risikostyringsmæssige risici, der kan være ved det.

Såfremt Finanstilsynet vurderer, at der er behov for yderligere vejledning i dansk regi, vil Finans Danmark foretrække, at det i så fald baseres på EBA's nuværende vejledning på området [Final Report on Big Data and Advanced Analytics.pdf \(europa.eu\)](#) samt Ekspertgruppen på højt niveau om kunstig intelligens etableret af EU- Kommissionen vejledning om etiske retningslinjer for pålidelig kunstig intelligens. [https://www.europarl.europa.eu/meetdocs/2014\\_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI\\_DA.pdf](https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_DA.pdf)

I forhold til det foreliggende udkast har vi en række overordnede bekymringspunkter:

- Vi er bekymrede over et eventuelt overlap med Forbrugerombudsmandens kompetenceområde (§ 1 "Indledning" henviser til udbud af produkter, der er let tilgængelige og målrettet den enkelte kundes behov)

## Hørings svar

26. juni 2023

Dok. nr.:

FIDA-1232191906-51062-v1



- Det er uklart hvilken status vejledningen har. Titlen "bedste praksis" og terminologi som "bør" antyder, at anbefalinger er bindende. Imidlertid fremgår det på side 7, at "dette er vejledende principper og forklaring på, hvad der kan blive god praksis i fremtiden, efterhånden som dataetikens betydning vokser". Vi anbefaler, at begrebet "bedst praksis" ikke anvendes og der ryddes op i sprogbrugen. En model kunne være at det ikke er et offentligt dokument.
- Vi finder desuden at de nævnte eksempler i vejledningen bør nøje overvejes. De aktuelle eksempler synes at antyde en meget aggressiv brug af data i danske finansielle institutioner, som ikke kan genkendes. Det er desuden tvivlsomt om den behandling, der angives i eksemplerne, alle er lovlig. Eksempler med ulovlige brug af data bør ikke medtages. En mulighed kunne være at undlade eksemplerne. Alternativt at involvere Datatilsynet iff. de anvendte eksempler.
- Endelig forekommer det uklart, hvad Finanstilsynet mener forskellen er på machine learning og AI. Eksemplerne synes at være regler som bruger data mere end egentlig AI teknologi. Det handler mere om dataetik end om AI.

## Hørings svar

26. juni 2023

Dok. nr.:

FIDA-1232191906-51062-v1

### Specifikke bemærkninger

Den retlige status af dokumentet bør fremgå tydeligt i indledningen. På side 7 i toppen nævnes beslægtede regelsæt og tilsvarende på side 11 top. Er der en særlig årsag til, at den kommende AI forordningen ikke er nævnt disse steder?

På side 5, afsnit 4 beskrives det dataetiske landskab, hvor reglerne vedrørende politik for dataetik nævnes, og at finansielle virksomheder er underlagt reglerne i en række bekendtgørelser og ikke årsregnskabsloven. Dette bør afspejles under sammenfatningen i afsnit 3.

På side 7 i toppen forklares, at dataetik går videre end data- og privatlivsbeskyttelse. Dog synes der at være en gennemgående fokus i dokumentet vedrørende brug af kundedata. Det bør bemærkes, at dataetik jo netop også gå videre end data- og privatlivsbeskyttelse, fordi dataetik omfavner ikke-persondata, herunder at machine learning/AI også benytter data, der ikke er karakteriseret som persondata.

På side 7, afsnit 5 defineres AI. Det er problematisk, at der ikke anvendes en definition der stemmer overens med den kommende regulering. Man burde i stedet for overveje blot at henvise til definitionen i den kommende regulering.



På side 9, afsnit 7.1 står følgende "Det er generelt bredt anerkendt, at potentialet ved AI er stort, også i den finansielle sektor, og at brug af teknologien kan øge virksomhedernes effektivitet og indtjening." Et lige så vigtigt aspekt for brug af machine learning/AI er tillige til hjælp for de finansielle virksomheder til overholdelse af retlige forpligtelser.

I afsnit 7.1 skrives omkring brugen af AI i den finansielle sektor, herunder open-source, hvor der skrives lidt om in-house vs. tredje-part modeller. En tredje mulighed er at gøre brug af tredjeparts modeller men bygge videre på dem in-house. Det kunne være ønskeligt, at dokumentet også satte lidt ord på den situation.

På side 18 og 19 beskrives lidt omkring god og dårlig bias, vi er meget enige i, at man skal undgå skævheder i data og gøre sig overvejelser omkring rimelighed (fairness) er et essentielt parameter. Med tekniske briller på, ville en model der skal forudsige noget bestemt dog siges at have en bias, da modellen netop skal hjælpe frem til det der ønskes forudsagt.

Med venlig hilsen

**Mette Stürup**

Direkte: +4527152020

[ms@fida.dk](mailto:ms@fida.dk)

## Hørings svar

26. juni 2023

Dok. nr.:

FIDA-1232191906-51062-v1



**From:** Digitaliseringsklar lovgivning - høring <klarlovgivning@digst.dk>  
**Sent:** 27-06-2023 15:05:54 (UTC +02)  
**To:** Camilla Neuenschwander (FT) <CNE@FTNET.DK>  
**Cc:** Nini Katrine Bertelsen <ninber@digst.dk>  
**Subject:** RE: Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Kære Camilla Neuenschwander

Vi har blot en enkel kommentar:

På side 5 skriver i: Desuden indeholder Erhvervsstyrelsens "Virksomhedsguiden" både artikler og videoer om emnet".

Indholdet på denne side er lavet af Digitaliseringsstyrelsen, hvor den erhvervsrettede indsats for dataetik også hører til. Derfor må dette gerne beskrives tydeligt. F.eks. "Desuden indeholder "Virksomhedsguiden" både artikler og videoer om dataetik som er skrevet af Digitaliseringsstyrelsen."

Med venlig hilsen



**DIGITALISERINGSSTYRELSEN**

**Mikkel Bo Jensen**

Fuldmægtig

Kontor for analyse og strategi

Sekretariatet for digitaliseringsklar lovgivning

M: +45 20 12 49 54

E: [mikjen@digst.dk](mailto:mikjen@digst.dk)

Digitaliseringsstyrelsen

Landgreven 4, Postboks 2193

1301 København K

[www.digst.dk](http://www.digst.dk) – fordi hverdagen er digital

*Du bedes venligst være opmærksom på, at Digitaliseringsstyrelsen ved modtagelsen af din meddelelse behandler de personoplysninger om dig, der fremgår af beskeden. Dette sker som led i Digitaliseringsstyrelsens myndighedsudøvelse. Du kan læse mere om, hvordan Digitaliseringsstyrelsen behandler dine personoplysninger, og hvilke rettigheder du har som registreret mv. i styrelsens privatlivspolitik <https://digst.dk/om-os/privatlivspolitik/>*

---

**From:** Camilla Neuenschwander (FT) <CNE@FTNET.DK>

**Sent:** 17. maj 2023 14:01

**To:** Fintech <Fintech@FTNET.DK>

**Subject:** Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Til rette vedkommende

Hermed sender Finanstilsynet udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor i høring.

Finanstilsynets hensigt med god praksis-papiret er ikke at indføre nye forpligtelser for eller krav til virksomhederne. Baggrunden for udarbejdelsen af papiret fremgår i øvrigt af vedhæftede høringsbrev.

Finanstilsynet beder om eventuelle bemærkninger til udkastet senest den 27. juni 2023. Bemærkninger kan sendes pr. e-mail på [cne@ftnet.dk](mailto:cne@ftnet.dk) eller pr. post til den i høringsbrevet nævnte adresse.

Med venlig hilsen

**Camilla Neuenschwander**

Chefkonsulent, cand.jur.  
Kontor for Fintech, Betalingstjenester og Governance



Strandgade 29, 1401 København K  
Tlf.: +45 33 55 82 82  
Direkte tlf.: +45 33 55 82 12  
<mailto:cne@ftnet.dk>  
[www.finanstilsynet.dk](http://www.finanstilsynet.dk)

---

Finanstilsynet er ansvarlig for behandlingen af de personoplysninger, vi modtager om dig. Du kan læse mere om, hvordan vi behandler dine personoplysninger på vores hjemmeside <https://www.finanstilsynet.dk/Kontakt/Privatlivspolitik>

Finanstilsynet gør opmærksom på, at denne e-mail og eventuelle vedhæftede filer er fortrolige. Hvis du har modtaget denne mail ved en fejl, bedes du straks oplyse Finanstilsynet herom ved at besvare denne e-mail og derefter slette e-mailen. Vi gør opmærksom på, at hvis du har modtaget denne e-mail ved en fejl, kan enhver form for kopiering, offentliggørelse eller distribution af denne e-mail være ulovlig.

27 June 2023

Danish Financial Supervisory Authority  
Strandgade 29  
1401 Copenhagen K  
Denmark

## Nasdaq's comment letter to Danish FSA's draft Good Practice Paper on ethical use of data in AI

Nasdaq Copenhagen A/S welcomes the opportunity to comment on the Danish FSA's draft Good Practice Paper on the ethical use of data in AI.

Nasdaq operates several regulated markets and MTFs across the Nordics and the Baltics, which are home to more than 1,200 listed companies. The Nasdaq trading venues drive the European and global economy and provide investment opportunities for both institutional and retail investors. At Nasdaq we use Artificial Intelligence to improve capital markets by designing smart systems and mechanisms that aim to increase market transparency, enhance market efficiency, and facilitate price discoverability. Besides enhancing our internal processes and solutions, the Nasdaq Group, which includes all our markets globally and our financial technology and information services businesses, uses AI and machine learning to help financial institutions increase their efficiency and effectiveness. A prominent example is how our Anti-Financial Crime Division uses AI to help over 2,400 financial institutions in the US and Canada to prevent and detect potential financial crime such as money laundering and financing of illegal activities.

Against this background, we would like to share our experiences and work on ethics in the field of AI development and provide reflections on important considerations, which we believe should be taken into account when developing best practices, codes of conduct and future regulation on AI.

Nasdaq commends the Danish FSA on the initiative to put forward a set of non-binding considerations on ethical data use for AI for the financial sector. The need for considerations on ethics when using AI solutions is evident in the current environment - with the rise of generative AI models such as ChatGPT, bringing a significant shift and expansion in the use of AI technology worldwide.

Overall, we support the Danish FSA's efforts to promote considerations on proper and safe use of data in AI solutions in the financial sector. We believe this could ensure trust in AI and thus the best possibilities for further growth of AI in an accountable manner. The draft Good Practice Paper is also of importance in light of the EU-US Trade and Technology Council's announced work on a voluntary code of conduct on AI.

Nasdaq works actively with both ethics and governance with regard to AI, having focus on an ethical AI culture within the organisation. As an example, Nasdaq has created a set of AI ethical principles, which reflect Nasdaq's core values<sup>1</sup>, and which we apply when

---

<sup>1</sup> Available at <https://www.nasdaq.com/docs/nasdaq-culture-book>

developing and using technology and services that incorporate AI. Our AI ethical principles focus on amongst others transparency, fairness, accountability and reliability. We recognise those same elements as areas emphasized in the draft Good Practice Paper as elements of importance and to be considered when using data in AI.

At the same time we recognize the need for anchoring of ethical principles within the organisation and the importance of setting the tone from the top, as highlighted in the draft Good Practice Paper. In addition to being connected to the organisation's core values, our approach is operationalized via a risk-based internal governance framework, with procedures in place as part of the product development lifecycle process. This includes oversight by an internal AI Governance Committee, focused specifically on the AI field.

The draft Good Practice Paper aims at highlighting the need for considerations around ethical use of data, and it does not in itself create regulatory obligations. We are of the view that such ethical considerations in the use of AI go hand in hand with the current national, Danish regulatory requirements on data ethics disclosures in annual reports and supplement upcoming regulatory requirements at the EU level, i.e. the AI Act.

Therefore, we are sharing our reflections with regard to core AI policy and regulatory principles, which in our view should be taken into account in future AI regulation.

Overall, Nasdaq believes that regulation should encourage AI accountability in a way that considers the following guiding principles:

- **Reasonably Tailored.** Regulation should be reasonably tailored to the nature and degree of associated risk presented in an AI system and utilize existing legal and regulatory frameworks to avoid potential overlap with existing regulation. Additionally, regulation should not entail affirmative licensing but instead be based on self-regulatory organization (SRO) models in terms of accountability and oversight from supervisory authorities.
- **Application Focused.** Regulatory frameworks should be use-case and context specific with a focus on the application and human impact of AI technologies. The role of human oversight and approval should follow from this application-focused approach, bearing in mind that the appropriate level of human oversight should strike the right balance. In this regard, we welcome the Danish FSA's approach, encouraging considerations on human oversight and the proper level of human oversight for ethical use of AI. Additionally, regulation on AI should account for the full scope of AI application, including cases of AI interaction with AI.
- **Transparent.** The draft Good Practice Paper highlights the importance of financial actors' general understanding of and the ability to understand and explain how AI solutions work and how systems make decisions. In this regard, we recognize the need for transparency in the field of AI, and that regulatory frameworks should include clear definitions with case study examples informed by continued dialogue with industry stakeholders.
- **Enable Crime Fighting.** Regulations must protect individuals from abusive data practices and privacy violations. In doing so, they should enable financial institutions in Europe and other regions to leverage data both from within and

outside their own networks for limited, compelling purposes such as to identify criminal activity. Thus, we welcome that the draft Good Practice Paper highlights that AI can support and reinforce risk management of organisations, e.g. via identifying money laundering and terror financing. We are of the view that AI regulatory frameworks must allow for fighting financial crime, which we see as one of the prominent use cases for AI.

- **Ensure Responsible Data Use.** Responsible management and use of data, including good governance and risk management, is essential when developing regulatory frameworks. Thus, we recognize the importance of proper data use and such proper data use as supporting the financial sector in remaining robust as emphasized in the draft Good Practice Paper. While the ability to analyze information is vital for crime fighting, and regulation must allow for it, using data for other purposes may prove problematic.
- **Encourage Modernization.** Regulation on AI should encourage and not overburden efforts to modernize our capital markets, including leveraging the latest capabilities in cloud, AI and machine learning. Moreover, regulatory frameworks should be flexibly formulated to allow for modernization and innovation in AI technologies.
- **Ensure Fairness.** Lastly, AI Regulation must protect against technology or models that treat individuals unfavorably on the basis of characteristics such as race, gender, disabilities, social status, and sexual or political orientation.

We believe that these guiding principles can help designing and providing AI enabled products and services that are respectful of ethical values. As AI use cases and application evolve at a fast pace, Nasdaq reassesses periodically its guiding principles on AI. We would be pleased to discuss any aspects of the above, if you would find it useful.

We are in the advent of generative AI and a significant shift, where emerging technologies can bring growth and effectivisation. Thus, now is the time to ensure proper tools for growth of responsible AI.



Finanstilsynet  
Strandgade 29  
1401 København K

Att.: Camilla Neuenschwander  
Pr. email: [CNE@ftnet.dk](mailto:CNE@ftnet.dk)

27. juni 2023

## Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Tak for muligheden for at kommentere denne høring. Vi sender hermed vores bemærkninger til Finanstilsynets udkast til vejledning.

Den finansielle sektor er langt fremme – i forhold til andre sektorer – i forhold til anvendelsen af kunstig intelligens (AI). Dette skyldes bl.a. de store mængder af data og den høje grad af digitalisering, som er kendetegnet ved denne sektor, ligesom økonomiske gevinster for banker, forsikringsselskaber mv. i mange tilfælde er ikke-ubetydelige. Det er i den forbindelse vores vurdering, at der for netop den finansielle sektor er et stort og til dels uforløst potentiale i brugen af AI, hvilket gør sig gældende i forhold til effektivisering, udvikling af nye produkter og forretningsmodeller, men også i forhold til mere kontrolbetonede opgaver.

Det er vores opfattelse, at der er et udtalt behov for, at den finansielle sektor (som en betydelig repræsentant for den private sektor) bør vise vejen for en ansvarlig, fair og etisk brug af AI.

På den baggrund er vi også enige i, at der er et behov for, at Finanstilsynet understøtter en sådan etisk brug af AI – eksempelvis via en vejledning som den foreliggende.

### Generelle bemærkninger

Vi mener, at udkastet til vejledningen indeholder mange gode og relevante etiske overvejelser ift. brugen af store datamængder og AI. Når det er sagt, bærer vejledningen imidlertid også præg af at have et overvejende fokus på generel dataetik og kun i mindre grad på konkrete etiske overvejelser ift. brugen af AI. Således har vejledningen mere karakter af en dataetisk vejledning med sporadisk eksemplificering ift. AI. Det er således illustrativt for det samlede indtryk, at Finanstilsynet i det afsluttende afsnit nævner 'dataetik' syv gange og 'AI' én gang.

Det er ligeledes illustrativt, at eksemplerne A, B og E – efter vores vurdering – ikke forudsætter brug af AI – hverken som eksemplet er gengivet, eller som det i praksis behøver at blive implementeret. Det er i øvrigt også for eksemplerne om tilbud om tilkøb af forsikringsprodukter baseret på kundens lokationsdata hhv. data om kundens helbred (eksempel A og B) uklart, hvori det etiske dilemma består, i det omfang kunderne som beskrevet har givet samtykke til, både at de relevante data indsamles og til at blive

FSR – danske revisorer  
Slotsholmsgade 1, 4. sal  
DK - 1216 København K

Telefon +45 7225 5703  
fsr@fsr.dk  
www.fsr.dk

CVR. 55 09 72 16  
Danske Bank  
Reg. 9541  
Konto nr. 2500102295



kontakten. For så vidt angår eksempel D (der både angiver, at der overvejes at indsamle data om kundernes etnicitet, samtidig med at virksomheden vil anvende AI til at foretage en kategorisering af kunderne i forhold til etnicitet – vi har svært ved at forstå, at begge dele skulle være nødvendigt) er det påfaldende, at de etiske overvejelser ikke knytter sig til brugen af AI, men derimod til indsamlingen af data om etnicitet. Vi skal derfor foreslå, at titlen ændres til f.eks. "Dataetik i den finansielle sektor", eller at indholdet af vejledningen tilpasses, jf. nedenstående bemærkninger, som i vidt omfang har en AI-vinkel.

Vejledningen har i øvrigt et relativt stort fokus på produktudvikling og anvendelse af data til glæde for kunderne. Således fremhæves dette særligt i det indledende afsnit (afsnit 1, s. 3). I praksis er en ikke-ubetydelig del af AI-anvendelsen imidlertid drevet af kontrolbehov og ønsket om at reducere kommercielle risici, hvilket slet ikke nævnes i indledningen. På s. 10 gengives dog en undersøgelse fra EBA, hvoraf det fremgår, at de fire mest udbredte anvendelsesformer knytter sig til forskellige former for kontrol og profilering. Dette kontrolperspektiv er efter vores opfattelse et af kerneelementerne i de etiske overvejelser og kan med fordel have en mere fremtrædende rolle i vejledningen.

#### **Den kommende AI-act**

I afsnittet om det dataetiske landskab (afsnit 4, s. 5-7) bør det efter vores opfattelse nævnes, at der i EU-regi arbejdes på en ny regulering af AI (AI-act), der i vidt omfang forventes at regulere forhold, som ellers kunne falde ind under det, der nu betegnes som AI-etiske overvejelser. Uagtet, at AI-act ikke er vedtaget, og at den konkrete udformning af retsakten således endnu ikke er endeligt fastlagt, kunne det med fordel overvejes at medtage konkrete, relevante eksempler fra den ny regulering i vejledningen.

#### **De etiske dilemmaer**

Det er som udgangspunkt vores opfattelse, at eksempler (afsnit 7, s. 9-15) kan være nyttige i forhold til at konkretisere de AI-etiske dilemmaer. Det kan imidlertid være vanskeligt at udlede nogle mere generelle og i praksis anvendelige principper fra de angivne eksempler og vejledningen i øvrigt.

Vi skal derfor foreslå, at vejledningen suppleres med et afsnit om mere generelle og tværgående principper, som kan danne grundlag for organisationernes videre arbejde med AI-etik, eksempelvis:

- Kontrol- eller servicedrevet brug af AI: En organisation bør gøre sig etiske overvejelser i forhold til, om AI kan anvendes til begge formål, og i givet fald om disse områder skal håndteres forskelligt i forhold til kundeinddragelse, muligheden for indsigelser, interne kontroller osv.
- Modstridende interesser: Brugen af AI er i praksis ofte (men ikke altid) drevet af organisationernes interesse i at udvikle nye forretningsmodeller, opnå effektivisering/besparelser eller undgå tab/svindel. Dette kan efter omstændighederne ske på bekostning af kundernes integritet og vil i så fald ofte kunne være i strid med kundernes interesser. En organisations AI-etiske overvejelser kan med fordel adressere denne mulige interessekonflikt, hvor organisationens økonomiske interesser måske må vige for hensynet til kundernes integritet.
- Formålsbestemthed i forhold til data: På databeskyttelsesområdet anvendes princippet om formålsbestemthed, idet en videreanvendelse af data som udgangspunkt ikke må være uforenelig med formålene for indsamlingen af data. Brugen af AI vil ofte ligge uden for eller i periferien af de oprindelige indsamlingsformål, og det kan i hvert fald ikke

udelukkes, at en ellers lovlig videreanvendelse af data kan opleves som intimiderende for kunder. AI-etiske overvejelser kan med fordel inkludere overvejelser om formålsbestemthed – dvs. hvordan organisationen generelt forholder sig til, at data anvendes i videre omfang end det, som kunder på forhånd måtte kunne forvente.

- AI eller menneske: Det er vores vurdering, at der i befolkningen som helhed er en ikke-ubetydelig modvilje (måske endda aktuelt en voksende frygt) for, at man bliver underlagt beslutninger, der træffes baseret på AI-modeller. Dette kan eksempelvis være begrundet i en frygt for, at AI-modellen tager fejl eller ikke forstår lige netop 'min særlige situation'. Tilsvarende kan det være forbundet med en følelse af umyndiggørelse eller ligegyldighed på områder, som kan have store personlige (økonomiske) konsekvenser for det enkelte individ. En organisation kan med fordel forholde sig til sådanne dilemmaer, herunder om der på udvalgte forretningsområder skal stilles større krav til, at kunderne modtager særlige informationer, eller at der på kundens foranledning kan inddrages personer i sagsbehandlingen. Alternativt at organisationer på særligt indgribende områder helt afholder sig fra at anvende AI.
- Afgørelser eller beslutningsunderstøttende aktiviteter: Uagtet at der i databeskyttelsesforordningens artikel 22 findes udtrykkelige regler om automatiserede afgørelser, kan man som organisation gøre sig overvejelser om graden (og konsekvenserne) af en automatiseret brug af AI. En egentlig afgørelse med negative konsekvenser for kunden er selvsagt mere indgribende end f.eks. en AI-baseret screening af kunden mhp. at tilvejebringe en oversigt over relevante produkter, som en medarbejder kan rådgive kunden på basis af. Således kan det være relevant som organisation at afholde sig fra mindre indgribende aktiviteter, indtil organisationen (og de relevante medarbejdere) har oparbejdet en tilstrækkelig modenhed og awareness ift. brugen af AI.
- Forretningsområder: I tilknytning til de ovenstående punkter vil en sondring mellem specifikke forretningsområder også kunne være relevant at inddrage i AI-etiske overvejelser. Således kunne man måske antage, at der ud fra et samfundsmæssigt perspektiv er en større accept af, at kreditvurderinger sker på et mere automatiseret (og bredt databaseret grundlag), hvorimod en meget indgribende profilering af kunder med henblik på at identificere snyd i "små" forsikringssager kan synes at være disproportional og uetisk, uanset om det måtte ske med kundens samtykke.
- AI-anvendelse, der gavner nogle kunder på bekostning af andre: Brugen af AI giver potentielt mulighed for at udarbejde meget detaljerede kunde profiler og -kategoriseringer, hvilket også kan give et grundlag for at målrette og prissætte produkter til meget specifikke kunde grupper. Dette er til dels berørt i eksempel F (s. 14), men på et mere generelt plan giver det også anledning til etiske overvejelser i forhold til, om en ellers lovlig detailkategorisering, der vil gavne nogle kunder på bekostning af andre kunder, ligger inden for virksomhedens etiske rammer. Som en variation af denne problemstilling kan også nævnes det eksempel, at kunder, der er villige til at afgive ekstra mange personoplysninger (og som dermed kan profileres i en højere grad), kan opnå potentielle økonomiske fordele, som andre kunder ikke ville kunne opnå.
- Forankring af ledelsesansvar for kontrol og styring: Organisationen bør overveje, hvordan styrings- og kontrolaktiviteter i forhold til anvendelsen af AI forankres i forhold til organisationens eksisterende kontrolapparat. I visse tilfælde kan den tværgående styring og kontrol med AI-aktiviteter formentlig indlejres i eksisterende styringskoncepter (f.eks. knyttet til data governance, risikostyring, god skik eller jura), men i alle tilfælde må der overvejes, om nye roller og fora må etableres med henblik på at løfte ledelsesansvar og beslutningskompetencer i forhold til anvendelsen af AI i organisationen.

Ovenstående er blot eksemplificeringer, men en sådan mere tematisk tilgang giver efter vores opfattelse et mere anvendeligt begrebsapparat, når de etiske overvejelser skal nedfældes og forankres i organisationen.

### **Den dataetiske tankegang og brugen af AI**

Afsnittet om den dataetiske tankegang (s. 16) virker efter vores opfattelse en smule misvisende – eller i hvert fald for lidt målrettet anvendelsen af AI.

Det generelle indtryk af afsnittet er, at ikke alene relevante interessenter i virksomheden, men alle medarbejdere skal optrænes i dataetik, hvilket muligvis (muligvis ikke) er korrekt. I forhold til brugen af AI synes det imidlertid at være af større betydning, at f.eks. ”menige” medarbejdere forstår output af AI-modeller (herunder at der kan være tale om sandsynligheder og ikke sikre svar), at de er uddannet i at identificere bias eller andre former for ’skred’ i resultaterne over tid, og at de er bekendt med de interne kontroller og sikkerhedsmekanismer, der løbende skal sikre korrekt anvendelse af AI. En AI-model er således typisk ikke blot et it-system på linje med alle andre it-systemer, og dette indebærer andre og nye risici, som medarbejderne skal forstå og forholde sig til. Der er imidlertid en risiko for, at et sådant fokus med tiden udvandes i organisationen – en risiko, som man eventuelt også kan forholde sig til i forbindelse med mere AI-etiske overvejelser. Dermed ikke sagt, at ”en fælles sund kultur for behandling af data” ikke også kan være relevant i et mere generelt dataetisk perspektiv.

### **Kundeinddragelse**

Vejledningen berører enkelte steder, at dataetik bl.a. handler om samfundsmæssige værdinormer mv., og i afsnittet om ansvaret for arbejdet med dataetik i virksomhederne (s. 16-17) nævnes f.eks. oprettelsen af etikkomiteer i forhold til at sikre, at den enkelte virksomheds forskellige interesser bliver bredt involveret i arbejdet med dataetik. Det er i den forbindelse vores vurdering, at der – som supplement til afsnittet om virksomhedens kunder (afsnit 8.3.3, s. 19) – med fordel kan peges på, at kunderepræsentanter kan inddrages i sådanne udvalg og komiteer med henblik på at have større føling med kundernes etiske værdisæt og med henblik på at tilvejebringe en større transparens.

–ooOoo–

Vi står til rådighed for evt. uddybning af ovenstående.

På vegne af FSR – danske revisorerers Cybersikkerhedsudvalg.

Med venlig hilsen

Thomas Krath Jørgensen  
Chef for revision og rapportering, statsaut. revisor

## Finanstilsynet

Sendt pr. e-mail: CNE@ftnet.dk

Att.: Camilla Neuenschwander.

Den 27. juni 2023

## Høringssvar – udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Dansk Erhverv takker for muligheden for at komme med input til Finanstilsynets udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor.

### Generelle bemærkninger

Dansk Erhverv er generelt meget positiv over for brugen af kunstig intelligens og nye teknologier til at skabe værdifulde løsninger for virksomhederne, kunderne og samfundet.

Dansk Erhverv er helt enig i vigtigheden af, at virksomheder løbende forholder sig til de etiske dilemmaer, der kan opstå ved brugen af kunstig intelligens. Det er derfor også positivt, at Finanstilsynet har til hensigt at inspirere virksomheder til at arbejde med dataetik. Dansk Erhverv har også lanceret 'Gode råd til virksomheders brug af kunstig intelligens', hvor ansvarlig anvendelse af kunstig intelligens også indgår i rådet omkring udarbejdelse af en politik for virksomheders brug af kunstig intelligens.

Dansk Erhverv bakker op om, at Finanstilsynet sætter fokus på dataetiske dilemmaer i den finansielle sektor for at sikre den tillid, der er helt afgørende for sektoren og dansk erhvervsliv. Det er afgørende for dansk økonomi, danske virksomheder og arbejdspladser, at Danmark har en sund finansiell sektor baseret på tillid fra danske borgere og virksomheder. Det gælder også i arbejdet med kunstig intelligens.

Det er vigtigt, at papiret har karakter af inspiration til virksomhederne, da det vil være særdeles u hensigtsmæssigt med danske særregler på dette område, og herudover at papiret understøtter danske virksomheders brug af kunstig intelligens, jf. nedenfor.

### Specifikke bemærkninger

Dansk Erhverv hilser meget velkomment, at Finanstilsynet ønsker at inspirere virksomhederne den finansielle sektor i det vigtige arbejde med dataetik.

Den finansielle sektor har en stor betydning for dansk økonomi. En sund og stabil finansiell sektor er afgørende for at kunne skabe vækst og udvikling i samfundet. Dette forudsætter også, at der er en generel opbakning og tillid til virksomheder i den finansielle sektor. Det er i den forbindelse vigtigt, at virksomhederne løbende forholder sig til etiske dilemmaer, herunder i forhold til brugen af kunstig intelligens.

Det er afgørende, at hensigt om at inspirere forfølges gennem hele papiret, da der – som det også fremgår – *ikke* er tale om regulatoriske krav. Det vil heller ikke være hensigtsmæssigt med danske særregler og enegang på dette område, da danske virksomheder er i skarp konkurrence med udenlandske virksomheder. Det bemærkes i den forbindelse, at der er EU-regulering på vej, herunder bl.a. AI-forordningen.

Det kan i den forbindelse anbefales at ændre overskriften, så 'God praksis' ikke indgår. Betegnelsen 'god praksis' kan opfattes som en praksis, der bør følges af virksomhederne, da det modsatte kan opfattes som 'dårlige praksis' eller 'mindre god praksis.' Papiret kunne f.eks. hedde 'Inspiration til dataetiske overvejelser ved brug af AI i den finansielle sektor'.

Dansk Erhverv læser også papiret som om, at det *ikke* har karakter af vejledning. Dette bakker vi op om og anbefaler, at det både af papiret og den efterfølgende brug og kommunikation tydeligt fremgår, at papiret har karakter af inspiration til virksomhedernes vigtige arbejde med dataetik.

Dansk Erhverv opfordrer til, at papiret gennemgås med henblik på at sikre, at eksemplerne og sprogbrugen tydeliggør, at der er tale om inspiration til virksomhederne.

Som eksempel på dette kan nævnes sætningen under afsnit 4 under pkt. 5 'Afgrænsning', hvoraf det fremgår: *'Der er tale om rent vejledende principper og formidling af, hvad der kan blive god praksis i fremtiden i takt med, at dataetik får stigende betydning.'* Den sidste del af sætningen bør udgå, da teksten kan læses som om, at der er tale om principper, der har – eller kan få – karakter af mere end inspiration til virksomhederne. Det vil ikke være hensigtsmæssigt i dag at forsøge at angive, hvad der vil kunne blive god praksis i fremtiden.

Dansk Erhverv deler fuldt ud synspunktet om, at virksomhederne selv skal have frihed til at definere, hvad god dataetik er, og hvordan de vil håndtere eventuelle dilemmaer, som det fremgår af side 8 nederst. Der kan for den enkelte virksomhed være mange individuelle forhold, der spiller ind i virksomhedernes drøftelser og håndtering, herunder strategi, produkter, kunder, kommunikation m.v. Det væsentligste er, at virksomheder er opmærksomme på de etiske dilemmaer og forholder sig til dem.

Et andet eksempel er på side 15, hvoraf det fremgår omkring prisvandring, at Finanstilsynet ville *'i det mindste have vurderet adfærden som uetisk'*. Denne sætning bør udgå, da det er uhenigtsmæssigt, at Finanstilsynet i dette inspirationspapir vurderer adfærd som værende etisk eller uetisk.

Det er meget positivt, at Finanstilsynet generelt ønsker at understøtte brugen af kunstig intelligens i den finansielle sektor og specifikt nævner, at AI kan være en stor fordel for både virksomheder, kunder og samfundet. Dansk Erhverv er helt enig i dette. Det er derfor også særdeles vigtigt, at papiret udformes med dette for øje, så danske virksomheder kan være konkurrencedygtige i forhold til virksomheder i andre lande og udnytte de digitale muligheder, som vi har i Danmark.

Papiret kan med fordel gennemlæses for at sikre, at papiret generelt understøtter brugen af kunstig intelligens. Som eksempler kan nævnes:

Under pkt. 3 'Sammenfatning' fremgår følgende: *'Virksomhederne kan have mange fordele af at bruge AI, men det er vigtigt, at virksomheder under Finanstilsynets tilsyn tager arbejdet med dataetik alvorligt'*. Budskabet kan med fordel ændres, så det ikke fremstår som om, at der er et modsætningsforhold mellem brugen af AI og dataetik.

Under afslutningen i pkt. 9 fremgår det, at *'Det øger også betydningen af dataetikken, når flere virksomheder bruger AI og derved gør behandlingen af data mindre gennemskuelig'*. Dansk Erhverv finder det uheldigt, at det fremgår som om, at det er en naturlig følge af brug af AI, at behandlingen af data bliver mindre gennemskuelig. Dette behøver ikke at være tilfældet. Tværtimod ses flere eksempler på, at kunstig intelligens kan være med til at skabe gennemsigtighed og transparens.

Dansk Erhverv finder det som nævnt positivt, at Finanstilsynet gerne vil inspirere virksomhederne i den finansielle sektor til at drøfte dataetik. Det er i den sammenhæng vigtigt, at det er op til virksomhederne at definere, hvad god dataetik er, og hvordan de vil håndtere eventuelle dilemmaer. Det er derfor vores opfattelse, at afsnit 8 bliver for konkret og peger på bestemte måder at agere på i virksomhederne.

Som eksempler på dette kan nævnes, når der tales om den rette kultur – og hvordan denne skal skabes. Et andet eksempel er budskabet om, at det er vigtigt at inddrage alle medarbejdere for at skabe en fælles sund kultur. Et tredje eksempel er, når det fremgår: *'Hvis en virksomhed beslutter sig for at arbejde seriøst med dataetik, er det vigtigt, at den forankrer opgaven og udpeger en ansvarlig medarbejder eller funktion for bl.a. at sikre, at opgaven bliver prioriteret'*.

Dansk Erhverv opfordrer til, at papiret, herunder særligt afsnit 8 ændres, så det mere får karakter af temaer, der kan inspirere virksomhederne i deres arbejde med dataetik.

Hvis ovenstående giver anledning til spørgsmål, står vi naturligvis til rådighed for uddybende bemærkninger.

Med venlig hilsen

**Poul Dahlgard**

Chefkonsulent og advokat

## Hørings svar til Finanstilsynet vedr. god praksis for dataetik

Idet Forbrugerrådet Tænk takker for modtagelsen af høringsbrevet af 17. maj 2023 skal vi hermed afgive vores bemærkninger til Finanstilsynets udkast til God praksis for dataetik ved brug af AI i den finansielle sektor.

Indledningsvist vil vi gerne rose Finanstilsynets initiativ til at inspirere virksomheder underlagt Finanstilsynet til at arbejde med dataetik. Det er højaktuelt at udbygge papiret om God praksis for machine learning til et decideret fokus på dataetik ved brug af AI. Dog gør vi opmærksom på, at dataetik også kan være relevant, når der indsamles persondata, uden at der nødvendigvis anvendes machine learning eller AI.

Brug af store datamængder, sårbarhed ved cyberangreb samt udviklingen indenfor kunstig intelligens og nye tjenester til følge forudsætter en høj grad af teknisk robusthed og forbrugertillid. Det er helt rigtigt, som det påpeges i udkastet, at dataetik handler om at tilegne sig en særlig tankegang, fremme en sund virksomhedskultur og forankre politikken på tværs af organisationen.

Ligeledes bemærkes, at reguleringen som fx databeskyttelseslovgivningen og den kommende AI-forordning ikke i sig selv er tilstrækkelig til at understøtte en bæredygtig udvikling i brugen af data, som bygger på danske og europæiske værdier og som kan styrke danske virksomheders konkurrenceevne i forhold til BigTech. Endelig finder vi, at finansielle virksomheder i kraft af deres særlige rolle i Danmark påtager sig et vigtigt samfundsvar ved at arbejde målrettet med dataetik.

Med andre ord er dataetik blevet et afgørende parameter for virksomheder at arbejde med og yderst relevant at sætte på dagsorden, når Finanstilsynets 2025-vision er at sikre virksomheder og forbrugernes tillid til det finansielle system.

Vi er enige i, at dataetik overordnet handler om, hvad virksomheder *bør eller ikke bør gøre* og ikke blot, hvad de kan og må efter lovgivningen. Det er også rigtigt, at der ikke findes en entydig definition af dataetik, men derimod flere måder at forstå begrebet på. Ikke desto mindre er der hjælp at hente i Dataetisk Råds værktøj "Dataetik - Sådan gør du" i forhold til at få defineret, hvordan dataetik skal forstås i den finansielle sektor, og det vil vi opfordre Finanstilsynet til at gøre for at lette virksomhedernes arbejde med dataetik i praksis.

Endvidere kan Finanstilsynet med fordel folde de gode, oplyste eksempler på dataetiske dilemmaer mere ud ved at anvende de principper og værdier, som ligeledes fremgår af Dataetisk Råds værktøj. Derved kan man komme den dataetiske vurdering nærmere og hjælpe virksomheder til at se, hvorvidt en given praksis er farbar eller ej.



Vi forstår på den ene side godt, at Finanstilsynet ikke har mandat til at tvinge virksomheder til at arbejde med dataetik og dermed er forsigtige med at melde klart ud i papiret, men på den anden side, er det vores opfattelse, at finansielle virksomheder ikke længere kan komme udenom dataetik, og at det vil gøre det praktiske arbejde med dataetik lettere med en mere klar linje fra Finanstilsynets side. En mulighed kunne derfor være at opfordre relevante brancheforeninger, som Finanstilsynet samarbejder med, til at anbefale deres medlemmer at udarbejde politikker for dataetik samt sikre undervisning, forankring og løbende opfølgning.

Vi skal endvidere opfordre Finanstilsynet til at definere bestemte metoder eller bestemt brug af AI, som Finanstilsynet ikke finder acceptabel ud fra en dataetisk tankegang, og på den måde sætte en klar retning og pege på, hvad der ikke er acceptabel praksis i branchen. Som I gør opmærksom på i papiret, er AI endnu ikke så udbredt i branchen, hvilket måske også kan være et udtryk for en usikkerhed i forhold til, hvad der er god etik eller ej.

Afslutningsvis skal Forbrugerrådet Tænk opfordre Finanstilsynet til kigge nærmere på, hvordan den danske mærkningsordning for IT-sikkerhed og ansvarlig dataanvendelse (D-mærket) kan understøtte de finansielle virksomheders tekniske robusthed og arbejde med dataetik. Vi har med stor glæde noteret os, at PensionDanmark har fået D-mærket, og vi synes det er oplagt for Finanstilsynet at pege på D-mærket i selve papiret som en relevant løsningsmodel i forhold til at sikre dataetikken i branchen.

Med venlig hilsen

**Uffe Rabe Krag**  
Politisk chef

**Anette Høyrup**  
Chefjurist

## Høring over udkast til god praksis for dataetik ved brug af AI (kunstig intelligens) i den finansielle sektor

29. juni 2023  
[flnit]  
[sNr]

Finansforbundet takker for muligheden for at kommentere på udkast til god praksis for dataetik ved brug af AI i den finansielle sektor.

Digitaliseringen og data spiller en afgørende rolle i den finansielle sektor, og AI er gennem årtier blevet en naturlig komponent i sektorens opgavevaretagelse og forretningsmodeller. Der er et generelt højt niveau ift. datasikkerhed og -beskyttelse, men der er et behov for, at sektoren – og det omkringliggende samfund i øvrigt – diskuterer på ny, hvordan vi i det videre arbejde kan sikre ansvarlighed og ordentlighed – på ikke-konkurrenceforvridende eller urimelige præmisser.

Finansforbundet er overordnet meget positive overfor, at Finanstilsynet anerkender behovet for, at virksomhederne går et skridt videre end alene at leve op til gældende regler og lovgivning, og påpeger behovet for, at virksomhederne også tager stilling til de etiske spørgsmål, der opstår i takt med, at data får stigende betydning. Finansforbundet ser et potentiale i og indgår gerne i drøftelser om, hvordan papirets principper og forslag i fremtiden kunne få en mere forpligtende karakter.

Finansforbundet vil samtidig opfordre Finanstilsynet til også at lægge vægt på de etiske dimensioner, muligheder og faldgruber forbundet med opsamling og brug af data i samspillet med AI, der er indlejret i HR- og ledelsesopgaven. Dataetik skal være til stede og udleveres både på inder- og ydersiden af virksomheden.

I afsnittet "8.1 Den dataetiske tilgang" lægges op til, at virksomheden bør tage stilling til, hvordan den vil arbejde med dataetik, herunder hvordan den valgte tankegang fremmes hos medarbejdere og ledere samt hvordan medarbejdere kan inddrages.

Der lægges i papiret vægt på, at medarbejdere skal trænes i at kunne opdage dataetiske dilemmaer samt i at kunne reflektere og agere dataetisk i relevante situationer. Dette kunne med fordel udfoldes nærmere, så det står tydeligere for virksomhederne, hvordan de konkret kan fremme dataetisk tækning hos medarbejdere.

Der kunne samtidig peges på, hvordan dette kunne opfyldes, samt gives konkrete forslag til fx hvordan medarbejderne får det nødvendige grundlag, gennemsigtighed, til at kunne forholde sig reflekterende og afsøgende ift. evt. dataetiske dilemmaer eller problemstillinger.

"Human-in-the-loop" princippet handler om at indsætte en menneskelig gatekeeper, medarbejder/leder, i en AI-understøttet anbefaling/beslutningsproces. Det fodrer, at virksomheder gør sig overvejelser om, hvordan medarbejdere/ledere klædes på til at

varetage en sådan rolle, og at rammerne understøtter det ansvar, herunder beskyttelse af ledere/medarbejder ift. et evt. modpres eller teknologiafhængighed.

Finansforbundet mener dertil, at en ambitiøs dataetisk politik bør have opmærksomhed på, hvordan data bruges i de interne virksomhedsprocesser. I takt med, at digitale medarbejderdata fra fx arbejdstelefoner, systemer, sensorer mv. i højere og højere grad inddrages i HR- og ledelsesudøvelsen, bør der stilles krav til, at såvel ledere som medarbejdere får den nødvendige indsigt, værktøjer og rammer til at kunne præstere og agere på den (databaserede) præmis. Fx bør såvel ledelse som medarbejdere kunne forstå, hvordan data påvirker medarbejdernes enkeltes muligheder for at avancere og kompetenceudvikle samt hvilken rolle data spiller i forhold til lønudvikling, opgavefordeling mv.

En udfordring er, at AI mange gange ligger gemt inde i maven af andre systemer. Det kan derfor være svært at gennemskue, hvilke data der ligger til grund for en beslutning, hvordan data sammensættes og hvilke kriterier, der anvendes til en konkret vurdering.

Et grundlæggende tiltag, der bør følge en dataetisk politik og satsning, er derfor uddannelse og kompetenceudvikling af både ledere og medarbejdere.

Virksomhederne bør tilsvarende arbejde med at skabe klare procedurer for inddragelse af medarbejdere. En omfattende undersøgelse af OECD tegner et klart billede af, at chancerne for succes med implementering af AI øges markant, hvis man lykkes med at inddrage medarbejderne. Alligevel sker inddragelsen af medarbejdere oftest usystematisk og har en orienterede karakter. Finansforbundet mener, der bør sættes et langt større fokus på at skabe klare procedurer for, hvornår og hvem, der skal inddrages i hvilke processer, fx ifm. indkøb, udvikling eller løbende vurdering af, hvordan en specifik AI performer. Inddragelse kan ske på flere niveauer og være mellem den enkelte leder og medarbejder, mellem virksomhedsledelse og medarbejdere og/eller repræsentanter. Inddragelse bør ske uanset om data og AI bruges til at levere/administrere finansielle produkter eller tjenester, eller til at levere/administrere interne medarbejderrettede processer.

Finansforbundet er derfor særligt positive over for, at Finanstilsynet sætter fokus på behovet for, at dataetik bør udmøntes i konkrete tiltag og ikke alene pakkes ind i politikker. Finansforbundets gennemgang af de banker, der er underlagt den lovpligtige redegørelse om dataetik i ledelsesberetningen, efter følg eller forklar-princippet, viser, at meget få virksomheder har iværksat reelle tiltag, der understøtter politikken. Dertil synes at være en overordnet tendens til, at politikkerne sjældent går længere end til den grundlæggende lovgivning, her særligt GDPR.

Dette billede understøttes af Erhvervsstyrelsens evaluering af virksomhedernes arbejde med dataetik. Finansforbundet mener derfor, at der er behov for at sætte øget fokus på tiltag til at arbejde dataetisk, herunder forankring, procedurer mv.

Finansforbundet bakker derfor op om vejledningens forslag om, at virksomheder, der formulerer en dataetisk politik bør introducere tiltag, der sikrer implementering og opfølgning på politikkerne, samt inddragelse af medarbejdere og ledelse i de dataetiske dilemmaer. Særligt tiltag vedrørende gennemsigtighed og inddragelse af ledere og medarbejdere ifm. brug af data samt overvejelser, implementering og evaluering af

en AI bør fastlægges. Dette bør igen ske uanset om AI-løsning anvendes i interne processer, ifm. ledelse af medarbejdere, eller i kundevedtne processer.

Finansforbundet er helt enige i, at dataetik skal forankres ét sted. Ansvar for dataetik er i dag fragmenteret, og mange virksomheder betragter dataetik og GDPR som to sider af samme sag, men dataetik bygger ovenpå GDPR. GDPR stiller krav til indsamling, opbevaring og forvaltning af personoplysninger. Dataetik har et bredere perspektiv, der handler om at vurdere og drøfte forskellige hensyn – også når der er tale om data, som ikke (længere) er personrelaterede. Dataetik handler kort sagt ikke kun om at vurdere, hvad virksomheden kan (teknologisk) eller må (juridisk), men handler tilsvarende om at tage stilling til, hvad virksomheden vil med data og AI (værdier, kultur og strategi).

Finansforbundet mener, at dataetik skal udbredes til hele organisationen, og det stiller krav om tydelig opbakning fra toppen i virksomhedsledelsen, og at der etableres en styring via politik og understøttede tiltag samt tydelighed om, hvor ansvaret ligger.

Finansforbundet står til rådighed for at uddybe overstående bemærkninger.



Digital

Copenhagen  
Fintech



FINANSFORBUNDET

## Hørings svar om papir vedrørende dataetik ved brug af AI (kunstig intelligens) i den finansielle sektor

### – DI Digital, Finansforbundet og Copenhagen Fintech

Med de store teknologiske landvindinger som kommer i kølvandet på AI, kommer også med nye og svære dilemmaer. Det stiller helt nye krav til virksomheder, der skal være på forkant med udvikling ift. at sikre konkurrenceevne, og samtidig være i stand til at forstå og tilpasse organisationerne til teknologiernes indbyggede begrænsninger og risici. Det er samtidig afgørende, at vi er i stand til at indfri potentialerne ved teknologi og dataanvendelse.

Trods mange bud er der ingen, der med sikkerhed ved, hvordan den digitale verden kommer til at udvikle sig. Netop derfor har vi brug for et øget fokus på at kunne drøfte, reflektere og tage stilling til de muligheder, faldgruber og dilemmaer, der opstår – vi har brug for dataetik

DI Digital, Finansforbundet og Copenhagen Fintech takker derfor for et godt udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor.

Den finansielle sektor og Fintech-området har en lang tradition for at sikre høj compliance og ordentlighed i de digitale finansielle løsninger. Over de senere år er særligt spørgsmålet om dataetik og ansvarlighed blevet et omdrejningspunkt i udviklingen af nye Fintech AI-løsninger. Der er derfor positivt, at der samles forslag og inspiration til, hvordan sektoren bedre kan ensrette og rammesætte dataetik, uden der er tale om indførelsen af nye regler og krav. Dataetik er en proces, som virksomhederne skal igennem for at skabe en dataetisk kultur, og altså ikke noget som let kan, eller nødvendigvis bør omdannes til regulering. I afgrænsningen fremgår det i overensstemmelse hermed, at papiret alene er inspiration til overvejelser som virksomhederne kan gøre sig i arbejdet med dataetik, og at det på ingen måde er bindende for virksomhederne. Dette kan med fordel rykkes frem, og inkluderes i indledningen, så formålet med papiret er tydeligt for modtageren fra starten af.

Det er positivt, at Finanstilsynet sætter fokus på at skabe en stærk finansiell stemme på den dataetiske dagsorden, og at papiret dermed ikke er afgrænset til virksomheder, der er omfattet af de krav i årsregnskabsloven, som er nævnt i afsnit 4, men adresserer hele det finansielle område bredt.

Internationalt skal der være tillid til danske finansielle digitale løsninger, på samme måde som der er høj tillid til danske fødevarer og design. Dataetik kan skabe nye eksportmuligheder og være med til at bane vejen for, at fremtidige digitale løsninger til fx kritisk infrastruktur bygger på danske værdier. Ambitionen er, at "Made in Denmark" bliver ensbetydende med, at digitale løsninger er "Made with Ethics". Etik i en forretningsmæssig kontekst kan give værdi for forretningen i form af konkurrencemæssige fordele og er altså ikke alene en idealistisk indsats. Det kan bl.a. skabe en fordel i den globale kamp om at tiltrække og fastholde de eftertragtede digitale

29. juni 2023

[flnit]

[sNr]

kompetencer og it-specialister, innovative løsninger til kritisk finansiel infrastruktur af høj kvalitet og baseret på danske og nordiske værdier.

DI Digital, Finansforbundet og CPH Fintech er positive overfor de vejledende principper og Finanstilsynets forslag til, hvad der kan blive god dataetisk praksis i fremtiden, men må samtidig påpege at dataetik ikke alene er relevant i relationen mellem kunder, virksomheder og samfundet. Det handler ligeledes om opsamling, opbevaring og brug af medarbejderdata mv.

I takt med, at virksomheder i stigende grad bruger medarbejderdata ifm. rekruttering, performance, udvikling og karriere-vurderinger mv., så rejses tilsvarende behov for dataetiske procedurer og praksis i virksomhederne, der sikrer transparens, refleksion, kompetenceudvikling og inddragelse af medarbejdere. Selvom dette ikke er Finanstilsynets naturlige interesse- og kompetenceområde, så er medarbejderne en central aktør i den digitale transformation, og vi mener dermed at papiret også bør adressere medarbejderperspektivet, herunder også de dataetiske rammer virksomheder opstiller, når AI bruges som led i det interne arbejde, fx strategiprocesser, rekruttering, ledelse mv.

For at give et mere velbalanceret billede af både de muligheder og udfordringer, der er forbundet med sektorens brug af AI, bør der også tilføjes eksempler og cases der fremhæver potentialerne ved brug af data og AI. Der er nemlig en overvejende negativ vinkling i de nuværende cases. Finanstilsynet kunne overveje at henvise til konkrete afgørelser og positive use-cases, der parallelt illustrerer fordele og muligheder, som dataetik kan medføre samt operationaliserbare tiltag. Det anbefales dertil, at der henvises til eksisterende gode værktøjer, fx D-mærket, som er en mærkningsordning, der hjælper virksomheder med at arbejde struktureret med dataetik, og som stiller en række krav til pålidelige algoritmer og AI. Som den første finansielle virksomhed har PensionDanmark opnået D-mærket, som dokumenterer ansvarlig og transparent håndtering af medlemmernes data. Det er vigtigt med en fælles ramme og forståelse – også på tværs af sektorer – for hvordan man arbejder med dataetik i praksis, derfor bør de samme allerede eksisterende gode værktøjer anbefales.

Det kunne være en fordel, hvis papiret bidrog med et bedre overblik over, snitflader til eksisterende relevant lovgivning. Der eksisterer nemlig allerede en lang række forskellige reguleringer, som et frivilligt optag af dataetik kan stå på skuldrene af. Det kunne også være relevant og brugbart for virksomhederne, hvis papiret forholder sig til dataetik i relation til kommende AI-forordning, herunder hvilken rolle papiret spiller frem mod ikrafttrædelsen, men også når den er trådt i kraft.

DI Digital, Finansforbundet og CPH Fintech står til rådighed for at uddybe overstående bemærkninger.



Digital

Copenhagen  
Fintech

FINANSFORBUNDET

Finanstilsynet  
Strandgade 29  
1401 København K

29. juni 2023

J.nr. 2023-12-2076  
Dok.nr. 610523  
Sagsbehandler  
Anna Carolina Jensen

Sendt med Digital Post

---

## Høring over udkast til papir om god praksis for dataetik ved brug af AI i den finansielle sektor

Datatilsynet  
Carl Jacobsens Vej 35  
2500 Valby  
T 3319 3200  
dt@datatilsynet.dk  
datatilsynet.dk  
CVR 11883729

Finanstilsynet har den 17. maj 2023 anmodet om Datatilsynets eventuelle bemærkninger til ovenstående udkast.

Datatilsynet har indledningsvis noteret sig, at det fremgår af udkastet, at det er en forudsætning for at kunne foretage dataetiske overvejelser ved brug af AI i den finansielle sektor, at virksomheden i forvejen overholder gældende lovgivning, herunder bl.a. reglerne i databeskyttelsesforordningen<sup>1</sup> og databeskyttelsesloven<sup>2</sup>.

Udkastet giver umiddelbart Datatilsynet anledning til følgende bemærkninger:

1. Udkastets afsnit 7.2 indeholder en række eksempler, hvor en påtænkt behandling af personoplysninger gøres til genstand for dataetiske overvejelser. Nogle af de valgte eksempler indeholder også databeskyttelsesretlige problemstillinger. Det drejer sig om følgende eksempler:

Eksempel C: Et pengeinstitut overvejer at udvikle en algoritme, der kan stå for kreditvurderinger. Det kan lette arbejdsbyrden for medarbejderne og optimere forretningen. Pengeinstituttet vil træne algoritmen på et stort og nøje udvalgt dataset, der er repræsentativt for pengeinstituttets praksis for kreditgivning.

Eksempel C og følgeteksten bør efter Datatilsynets opfattelse revideres, så det fremgår tydeligere, at den registrerede har krav på menneskelig indgriben i tilfælde af automatiske afgørelser omfattet af databeskyttelsesforordningens artikel 22. Det følger af databeskyttelsesforordningens artikel 22, stk. 3.

Eksempel D: En finansiell virksomhed overvejer systematisk at indsamle data om sine kunders etnicitet for at kunne dokumentere, at virksomheden ikke diskriminerer på baggrund af race eller lignende. Virksomheden vil bruge AI til at kategorisere kunderne.

Eksempel E: En finansiell virksomhed vil bruge AI til svindlbekæmpelse. Virksomhedens teori er, at bestemte etniske grupper svindler oftere end andre. Den sorterer kundehen-

---

1 Europa-Parlamentets og Rådets forordning (EU) 2016/679 af 27. april 2016 om beskyttelse af fysiske personer i forbindelse med behandling af personoplysninger og om fri udveksling af sådanne oplysninger og om ophævelse af direktiv 95/46/EF (databeskyttelsesforordningen).

2 Lov nr. 502 af 23. maj 2018 om supplerende bestemmelser til forordning om beskyttelse af fysiske personer i forbindelse med behandling af personoplysninger og om fri udveksling af sådanne oplysninger (databeskyttelsesloven).

vendelser for at identificere kunder, der bor i områder, som optræder på Indenrigs- og Boligministeriets liste over parallelsamfund. Kun disse henvendelser underkastes nøjere granskning.

Side 2 af 2

Det er Datatilsynets umiddelbare opfattelse, at de beskrevne behandlingsaktiviteter ikke vil kunne ske i overensstemmelse med databeskyttelsesforordningens artikel 5, stk. 1, litra a, b og c, samt artikel 11, uanset om den registrerede har givet sit udtrykkelige samtykke. Virksomhederne skal dokumentere modellens eventuelle bias eller mangel på samme på anden vis.

Datatilsynet skal derfor henstille, at disse eksempler revideres.

**2.** Derudover har Datatilsynet følgende bemærkninger til udkastets afsnit 8.3:

Afsnit 8.3.1 indledes således:

God dataetik starter fra det øjeblik, en virksomhed tager skridt til at indsamle data. Virksomheden bør først nøje overveje, hvilke data der er nødvendige for at levere det konkrete finansielle produkt. Nogle oplysninger er blot bekvemme at have og bruge, hvorimod andre oplysninger er nødvendige for f.eks. at kunne tilbyde en ydelse eller overholde lovgivningen.

Datatilsynet bemærker i den forbindelse, at det ikke er lovligt at behandle personoplysninger alene under henvisning til, at oplysningerne er "bekvemme at have og bruge". Der skal altid være et formål med behandling af personoplysninger, og behandlingen skal altid være nødvendig af hensyn til formålet med behandlingen – her levering af det konkrete finansielle produkt. Det følger af databeskyttelsesforordningens artikel 5, stk. 1, litra b og e.

Af afsnit 8.3.1 fremgår herefter bl.a. følgende:

"Som nævnt i Finanstilsynets papir "God praksis ved brug af superviseret machine learning" bør virksomheden sikre datakvalitet og stabilitet i indhentning af data, herunder ved indhentning fra tredjeparter. Finanstilsynet opfordrer til, at virksomhederne generelt er bevidste om og kritiske overfor deres datakilder. Hvor stammer data fra, og er data valide? Virksomhederne bør definere, hvad de anser som en valid datakilde. Hvilke kilder ønsker virksomhederne at bruge og til hvilke formål?"

Datatilsynet skal hertil bemærke, at rigtighed er et krav efter databeskyttelsesreglerne, jf. forordningens artikel 5, stk. 1, litra d. Det er ikke valgfrit at sikre, at data er valide og korrekte.

Afsnit 8.3.2 indeholder følgende ordlyd:

Det er vigtigt, at virksomhederne forholder sig til, om en AI-model er rimelig. Rimelighed handler ofte om at undgå skævheder, f.eks. bias i modeller eller i data. Det er vigtigt, at virksomhederne undgår skævheder i den benyttede data. Det er også vigtigt, at virksomhederne forholder sig aktivt til, hvordan de kan mindske effekten af bias i modellens tilblivelse og brug.

Datatilsynet bemærker, at det heller ikke er valgfrit for virksomheden at sikre sig, at de modeller, som virksomheden udvikler eller indkøber, er fri for bias og diskrimination. Dette er også et krav efter databeskyttelsesreglerne, jf. forordningens artikel 5, stk. 1, litra a, om lovlighed, rimelighed og gennemsigtighed.

Datatilsynet anbefaler derfor, at afsnit 8.3 tilpasses med henblik på at tydeliggøre ovenstående.

Udkastet giver ikke i øvrigt Datatilsynet anledning til bemærkninger.

Med venlig hilsen

Anna Carolina Jensen