

(Foranalyse)

# **Er automatisk opstilling af årsrapporter muligt?** – alene på baggrund af bogføringsdata

16.03.2017

Af Jonas Midtgaard og Nicholas Meilstrup

## Indholdsfortegnelse

<b>1. Indledning</b>	<b>4</b>
<b>2. Begrænsninger</b>	<b>5</b>
2.1 <i>Koncernhåndtering</i>	5
2.2 <i>Datatilgængelighed</i>	5
2.3 <i>Datakvalitet</i>	6
2.3.1 <i>Bogføringsfejl</i>	6
2.3.2 <i>Fortolkninger</i>	7
2.3.3 <i>Branchekoder</i>	8
<b>3. Ansvarsbegrænsninger hos virksomhederne/revisor</b>	<b>8</b>
3.1 <i>Ansvar båret af ERST</i>	9
3.2 <i>Virksomhedsledelsens ansvar</i>	9
3.3 <i>Ansvar båret af revisoren</i>	10
<b>4. Bogføringsdata</b>	<b>11</b>
4.1 <i>Dataenes form</i>	11
4.2 <i>Kontoplaner på tværs af virksomheder</i>	12
4.3 <i>Detaljegrad</i>	12
<b>5. Databehandling</b>	<b>13</b>
5.1 <i>Indhentning af data</i>	13
5.1.1 <i>Mængde af data</i>	14
5.2 <i>Lagring</i>	14
5.3 <i>Hastighed</i>	15
<b>6. Strukturering af data</b>	<b>16</b>
6.1 <i>Logisk/regelbaseret opstilling</i>	16
6.1.1 <i>Strukturerede data</i>	17
6.1.2 <i>Mellem strukturerede data</i>	20
6.1.3 <i>Ustrukturerede data</i>	21
6.1.4 <i>Diskussion af notekrav</i>	23
6.2 <i>Machine learning</i>	23
6.2.1 <i>Supervised learning</i>	25
6.2.2 <i>Unsupervised learning</i>	25
6.2.3 <i>Reinforced learning</i>	25
<b>7. Opstilling af årsrapporten (numerisk)</b>	<b>26</b>
7.1 <i>Artsopdelt vs. Funktionsopdelt</i>	27
<b>8. Kvalitetskontrol</b>	<b>27</b>
<b>9. Omdannelse af årsrapport til XBRL og PDF</b>	<b>29</b>
9.1 <i>Antagelser</i>	29
9.2 <i>Validering</i>	29
9.3 <i>XBRL-dokument</i>	29
9.4 <i>XBRL til PDF</i>	29
9.5 <i>Det tekniske aspekt</i>	29
9.6 <i>Fra ERP rå data til Regnskabsbasis 2.0</i>	30
<b>10. Indsendelse af opstillede årsrapporter</b>	<b>32</b>
10.1 <i>Behandling af data</i>	32
	2

10.2	<i>Kommunikation mellem systemer</i>	32
<b>11.</b>	<b>Sikkerhed</b>	<b>33</b>
11.1	<i>Servere</i>	33
11.2	<i>Kryptering og opbevaring af data</i>	33
<b>12.</b>	<b>Mulige tiltag</b>	<b>35</b>
12.1	<i>Øge kvaliteten af den allerede tilgængelige data</i>	35
12.2	<i>Implementere en standardisere konto plan</i>	35
12.3	<i>Igangsætte systematisk indsamling af manglende data</i>	36
12.4	<i>Implementer infrastruktur der gør det muligt at hente data</i>	37
<b>13.</b>	<b>Perspektiveringer</b>	<b>37</b>
13.1	<i>Machine learning</i>	37
13.1.1	<i>Offentlig godkendt konterings motor</i>	37
13.1.2	<i>Identifikation af snyd</i>	38
13.2	<i>Kan ERST overhovedet levere en løsning der for alle?</i>	38
13.3	<i>Tredobbelt bogholderi</i>	39

Opgaveformulering: "Mellem parterne er følgende aftale indgået om en **foranalyse**, der har til **formål at undersøge, hvor langt det praktisk muligt er at komme i arbejdet med at opstille en årsrapport automatisk** – alene på baggrund af bogføringsdata."

## 1. Indledning

I processen for at omdanne bogføringsdata til en årsrapport bliver bogføringsdataene (i) vurderet, (ii) kvalitetssikret, (iii) sammenhængskorrigeret, (iv) afstemt, (v) klassificeret og (vi) opstillet til det endelige resultat - en årsrapport.

Denne rapport vil forsøge at beskrive den proces som er manuel i dag hos f.eks. revisoren for dette arbejde. Formålet er først at anskueliggøre hvilke opgaver der helt præcist indgår i processen og herefter identificere hvordan hver enkelt opgave kan, såfremt muligt, automatiseres.

Rapporten vil behandle problematikken med regnskabsklasse B virksomheder som omdrejningspunkt.

Hvor kompleksiteten muligvis er lavere ved regnskabsklasse B virksomheder, er det ikke nødvendigvis muligt at sætte samme lighedstegn til hvordan den enkelte virksomhed klassificerer dataene i sin bogføring. Klassificering kan f.eks. være bestemt ud fra hvilke krav der stilles af virksomhedsejeren, regnskabschefen, bogholderen og/eller industrien virksomheden drives i.

De brede rammer (fortolkningsmuligheder) der b.l.a. er til både bogføring og udformningen af en årsrapport til virksomheder, kan tjene som en fordel og en ulempe i forhold til målsætningen om at opstille en årsrapport automatisk på baggrund af bogføringsdata. Et bredere defineret mål, gør det nemmere at ramme - om ikke plet, så målet. De løse krav der stilles til opstilling af en kontoplan gør det sværere at strukturere dataene for den enkelte virksomhed og ensarte opstillingen af årsrapporten.

Rapporten tager udgangspunkt i en, i højere grad digitaliseret verden, hvorved nogle aktiviteter kan håndteres af maskiner fremfor menneskehænder. Digitaliseringen og systematiseringen af bogføringsdata samt den teknologiske udvikling, åbner op for nogle muligheder der førhen ikke var tilgængelige. På baggrund af de tilgængelige muligheder vil rapporten belyse løsningsmodeller og

udfordringer forbundet med at få maskiner til at opstille en årsrapport på baggrund af bogføringsdata.

## 2. Begrænsninger

### 2.1 Koncernhåndtering

Koncerner er selskabskonstellationer hvor et selskab ejer ét eller flere selskaber (datterselskaber). Datterselskaber kan ejes af moderselskabet enten helt eller delvist, med f.eks. andre selskaber (eller naturlige personer). Hvor hvert enkelt selskab skal føre en journal over sine aktiviteter (herunder bogføring), er der intet krav om at disse data skal behandles i samme system (herunder bogføringssystem). Trods der er regler om at datterselskaber skal *følge* moderselskabets eller hovedaktionærens regnskabsperiode, kan der opstå forskydninger i tallene der f.eks. gør at moderselskabets aktiver *ikke* er en sum af de underliggende selskabers relative aktiver.

Såfremt der bogføres i en koncerns underliggende selskaber i forskellige systemer vil det (i) skabe udfordringer for automatisk opstilling af en årsrapport på baggrund af bogføringsdata, da moderselskabet er afhængigt af datterselskabernes ajourføring, (ii) det vil det muligvis stikke i øjnene, såfremt summerne ikke adderer op.

Om forskydningerne skyldes systemmæssige eller manuelle processor bør undersøges nærmere af andre.

### 2.2 Datatilgængelighed

Datatilgængelighed er den allervigtigste præmis for at automatisk opstilling og indrapportering af årsrapporter er mulig.

Bogføring laves i mange forskellige systemer i dag – den bredeste sondring er on-premise og off-premise (f.eks. cloud accounting). Hvor der eksisterer et hav af muligheder inden for begge kategorier, har cloud accounting den fordel at de (i hvert fald de nyere systemer) at de er skabt til at kunne kommunikere med "udefra systemer". Dette gør dataene tilgængelige og dermed ønsket om automatisk opstilling og indrapportering af årsrapporter muligt.

Mange on-premise systemer kan også kommunikere med eksterne applikationer, men det kan kræve

modifikationer til det eksisterende software. Disse modifikationer kan være omkostningsforbundet og besværlige at foretage.

Cloud accounting systemer er en oplagt mulighed af flere grunde, hvoraf den mest åbenlyse er at dataene (i nogle af systemerne) er gjort tilgængelig igennem et API.

## 2.3 Datakvalitet

Bogføring er stadig en manuel proces og et håndværk. Det betyder at (i) der sker fejl (ii) det kan udføres på mange måder med hver sit resultat og udfald. Mange bilag kan konteres på flere forskellige konti, alle sammen, måske, lige *rigtige*. Ydermere kan bogholderen, og senere revisoren, vælge at udgiftsføre aktiver i resultatopgørelsen eller aktivere dem. Her vil valget truffet af den ansvarlige vil have to vidt forskellige udfald. For hver gang der kan ske enten fejl, eller der træffes beslutninger der er op til fortolkning, komplicerer og besværliggør det automatiseringsprocessen af bogføringsdata der skal resultere i en årsrapport.

### Delkonklusion

Om hvorvidt der skal laves strammere (mere konkrete) regler til hvordan en virksomhed må håndtere sine konteringer i bogføringen, må være en beslutning hos myndighederne. Mere ensartede praksis virksomheder i mellem, vil være en fordel for ERST til formålet om at autogenerere en årsrapport, men muligvis til gene for virksomhederne der vil blive frataget en fortolkningsfrihed der nu er dem tilgængelig.

#### 2.3.1 Bogføringsfejl

Antal korrektioner revisoren foretager i bogføringsdataene inden den endelige årsrapport er afhængigt af kvaliteten bogholderens arbejde samt antal bilag i det pågældende regnskab. Revisoren, såfremt det er et revideret regnskab, indestår for at regnskabet er *retvisende*, men ikke *korrekt*.

Underordnet af hvilken tilgang der tages – ”revideret” imod ”ikke revideret” – eksisterer der, eller bør der eksistere, en proces hvor den ansvarshavende for udførelsen af regnskabet, som minimum tager stikprøver i bogføringen for at berigtige konteringerne, foretager korrektioner hvor det er nødvendigt, for dermed at sikre at det endelige regnskab er retvisende.

### 2.3.1.1 Udbedring af bogføringsfejl

Udbedring af bogføringsfejl bør foretages på en eller anden vis. Når processer automatiseres kan der være en tendens til at menneskerne fralægger sig ansvaret eller i hvert fald foretager det nødvendige arbejde, man førhen gjorde, for at sikre en vis kvalitet. Det bør derfor overvejes i hvilken udstrækning en automatisk proces kan påvirke kvaliteten af fejludbedring. På sigt, i takt med udviklingen (herunder digitalisering af bilag), er der uden tvivl tiltag der kan udbedre fejlbehæftede bogføring.

### 2.3.2 *Fortolkninger*

I de tilfælde hvor bogføringskonteringer er op til fortolkning som f.eks. "køb af mad på restaurant" kan konteres som både "repræsentation" eller "personale omkostninger", må der påregnes *slør* i processen omkring automatisk opstilling af årsrapporten. Skønstillfælde som dette, vil ikke være udsalgsgivende for årets resultatet, men kan have forskellige skattemæssige udfald (moms).

Et andet eksempel kan være køb af computerudstyr til 15.000 DKK, hvor valget kan stå i mellem at aktivere (og afskrive) eller omkostningsføre udstyret i resultatopgørelsen. Her vil balancen og resultatopgørelsen se vidt forskelligt ud, efter valg af tilgang.

Et tredje eksempel hvor fortolkning kan påvirke regnskabet relaterer til branchen virksomheden befinder sig i. Erhverver en autoforhandler sig en bil, kan det konteres som varelager, hvor det vil være svært at komme afsted med for en softwarevirksomhed.

Problematikken virker måske åbenlys, men såfremt der skal føres en kontrol og berigtigelse af konteringerne forudsætter det at hver enkelt virksomhed er allokeret til den "korrekte" branchekode. På samme vis vil indkøb af cola hos en kiosk kunne være varelager og muligvis personaleomkostninger for vores softwarevirksomhed.

I takt med at bilag i højere grad bliver formidlet elektronisk mellem virksomheder, samt bogføring i højere grad i fremtiden forventes at blive automatiseret, kan mange af problematikkerne oven for løses.

### **Delkonklusion**

Det er ikke denne rapports hensigt at spå om fremtiden. Trods der er indikationer på at udviklingen

løber imod elektronisk og automatiseret bogføring, er det sandsynligvis et stykke ude i fremtiden.

### 2.3.3 Branchekoder

EU har defineret 737 branchekoder hvoraf en virksomhed skal allokeres til én af disse. Udover statistiske formål for at beskrive den økonomiske udvikling fordelt på sektorer (mm.), vil disse allokeringer kunne tjene til formål at opstille en årsrapport på baggrund af bogføringsdata. Såfremt virksomhederne er allokeret til de korrekte branchekoder, vil det gøre det betydeligt nemmere at opstille en årsrapport automatisk, da berigtigelse af hvilken konto hver enkelt bogføringspost tilhører vil være nemmere.

**Eksempelvis (som nævnt ovenfor)** vil en autoforhandlers indkøb af biler kunne blive betragtet som indkøb til varelager, hvor det næppe vil være tilfældet for en softwarevirksomhed.

Der er mange modeller hvorpå disse konteringer kan valideres, én kan avendes alene, men flere bør anvendes i kombination med hinanden. Som bemærket i dette afsnit, kan branchekoder være ét lag, der anvendes til at berigtige posteringen til den valgte konto.

*Bemærkning:* Det skal nævnes at allokeringen af virksomheder til branchekoder i Danmark trænger til en oprydning, da mange virksomheder muligvis er allokeret til en forkert kategori.

## 3. Ansvarsbegrænsninger hos virksomhederne/revisor

Efter gældende lov er det ledelsen, der via ledelsespåtegningen indestår for, at et indsendt årsregnskab giver et retvisende billede af situationen i en virksomhed. Det retvisende billede skal være tilvejebragt inden for rammerne af ÅRL, men ledelsen er overladt en række valg, skøn og vurderinger, der giver dem mulighed for at variere billedet af virksomhedens økonomiske situation.

Tager man muligheden for variation og skøn fra ledelsen, og tvinger alle selskaber til at aflægge regnskab, regelbaseret eller via avancerede machine learning-algoritmer, kan outputtet afvige fra hvad ledelsen havde forestillet sig.

Hvem har så ansvaret for, at regnskabet giver et retvisende billede?



### 3.1 Ansvar båret af ERST

På nuværende tidspunkt er ERST ikke involveret i aflæggelsen af årsregnskaber. De modtager den færdige og underskrevne årsrapport, men har ingen indflydelse på selve udarbejdelsen. At stille krav om at regnskabsaflæggere bruger bestemte systemer, regelbaseret eller machine learning-algoritmer, til udarbejdelse af årsrapporter, vil være en stor ændring af dette set-up.

Det er vores opfattelse, at det vil være i ERSTs interesse, at påtage sig mindst muligt ansvar til en start, således at ERST blot vil agere "softwareudbyder" til udarbejdelse af årsrapporter.

Vi mener, at ERST som "softwareudbyder" må påtage sig ansvaret for, at beregninger og opstillinger er lavet korrekt og i overensstemmelse med angivne principper og regnskabspraksis. Tager ERST ikke dette ansvar, vil servicen være overflødig, da andre systemer (Caseware, Wolters Kluwer, osv.) vil være nødvendige som kontrol.

Vi mener, at virksomhedsledelsen samt revisoren stadig bør bære ansvare for, at regnskabet giver et pålideligt og retvisende billede. Det kræver dog mulighed for at lave ændringer eller gøre indsigelser ift. det af ERST genererede regnskab.

### 3.2 Virksomhedsledelsens ansvar

Efter gældende regler indestår ledelsen for:

- At årsrapporten er aflagt i overensstemmelse med ÅRL og danske regnskabsstandarder.
- At valgt regnskabspraksis er hensigtsmæssig og giver et pålideligt og retvisende billede af virksomhedens resultat og økonomiske situation.
- At de interne processer og kontroller er tilstrækkelige for udarbejdelse af et pålideligt og retvisende årsregnskab.

Vi mener fortsat at ledelsen skal underskrive på det endelige resultat, men det efterlader nogle udfordringer ift. regnskaber genereret af ERST, som ledelsen ikke ønsker at underskrive.

Hvordan håndteres situationer hvor ledelsen ikke mener, at ERSTs genererede regnskaber giver et retvisende billede?

Ledelsen bør skrive under på følgende:

- At alle transaktioner og begivenheder der har fundet sted i løbet af regnskabsperioden, er repræsenteret i resultatopgørelsen eller balancen.
- At alle aktiver og passiver i regnskabet er til stede og ejet af virksomheden på balancedatoen.
- At alle transaktioner er registreret nøjagtigt og i korrekt periode.
- At alle aktiver og passiver er værdiansat korrekt.

### 3.3 Ansvar båret af revisoren

Det er pt. den godkendte revisors opgave at sandsynliggøre, at det af ledelsen underskrevne regnskab giver et pålideligt og retvisende billede af virksomhedens økonomiske situation. Denne sandsynliggørelse sker gennem en række kontroller af det fremlagte regnskab. Det kan være kontrol af varelager, kontrol af tilgodehavender, at listede aktiver faktisk tilhører virksomheden, værdiansættelse af listede aktiver, mm.

- At vurdere om ledelsens valg af regnskabspraksis er passende.
- At vurdere ledelsens regnskabsmæssige skøn.
- Revisor er selvsagt ansvarlig for revisionen, og dermed ansvarlig for konklusionen.
  - o Dette er dog ikke det samme som en garanti for, at årsrapporten er uden væsentlig fejlinformation.
  - o Revisionen består i praksis af en række stikprøver, hvilket i sig selv efterlader mulighed for fejl.

Med henvisning til at en gennemført revision ikke er en garanti, mener vi at ERST på lang sigt bør implementere automatiserede kontrolsystemer. De automatiserede kontrolsystemer behøver ikke være på revisorernes "garantiniveau", for at det er økonomisk forsvarligt at implementere dem. Endvidere er virksomheder i regnskabsklasse B ikke underlagt revisionspligt.

Da der ikke er lovkrav om revisorpåtegning, og en revisorpåtegning ikke er en garanti, mener vi, at revisorens rolle og ansvar bør begrænses til rådgivning af ledelsen i udarbejdelse af et pålideligt og retvisende regnskab.

Typisk vil revisoren i praksis også stå for udarbejdelsen af selve årsregnskabet, hvilket i dette set-up vil være overflødig.

## 4. Bogføringsdata

### 4.1 Dataenes form

De to overordnede datasæt, der er nødvendig fra bogføringssystemer for at opstille en årsrapport er kontoplanen og journalposterings. Denne data varierer i både format og struktur afhængig af kildesystemet. Bogføringssystemer har hver deres måde at abstrahere deres data på, og det har effekt på hvordan hentet data ser ud.

Interessant indhold for kontoplanen er for hver konto et id, kontonavn, evt. kontonummer og om kontoen er en del af resultatopgørelsen eller balancen.

Interessant indhold for posterings er for hver postering et id, konto id, mængde posteret, side posteret til, valuta og posteringsdato.

Givet herunder er to eksempler på hvordan kontoplanen og posterings ser ud i to forskellige bogføringssystemer.

1. e-conomic, REST API (kilde: <https://restdocs.e-conomic.com/>)

Data hentet fra e-conomics REST API er i JSON format.

*Kontoplan* – Eksempel på struktur af en konto, se Bilag 1.1. Interessant data hentes fra felterne `accountNumber`, `name` og `accountType`.

*Posterings* – Eksempel på indhold af en postering, se Bilag 1.2. Interessant data læses fra felterne `entryNumber`, `accountNumber` under `account`, `amount`, `currency` og `date`.

2. Billy (kilde: <https://www.billy.dk/api/>)

Data hentet fra Billys API er i JSON format.

*Kontoplan* – Eksempel på struktur af en konto, se Bilag 1.3. Interessant data læses fra felterne `id`, `name`, `accountNo`, og `reportType` under `nature`.

*Posterings* – Eksempel på struktur, se Bilag 1.4. Interessant data læses fra felterne `id`, `accountId`, `amount`, `currencyId`, `side` og `entryDate`.

Begge systemer brugt i eksemplerne ovenfor leverer data i JSON format og har lignende abstraktioner af både konti og posterings. På trods af det, er der stadig forskelle der skal tages højde

for. Andre systemer kan have andre dataformater og strukturer der er væsentligt anderledes end det eksemplerne viser. Hos CrediWire er vi bekendt med mere end 10 forskellige datastrukturer fordelt over formaterne JSON, XML og CSV.

## 4.2 Kontoplaner på tværs af virksomheder

Det er tilladt for virksomhedsejeren at opstille sin egen kontoplan, allokere struktureret numre til de valgte konti (det modsatte er også tilladt), navngive konti som virksomhedsejeren finder det relevant mm. De løse krav, eller mangel på samme, er en klar fordel for den enkelte virksomhed, som har brug for og forstand til, at arbejde fleksibelt med sit regnskab. Det er dog ikke formålstjenesteligt for den "uerfarne" virksomhedsejer der opstiller sin egen "spaghetti" kontoplan med f.eks. en vilkårlig nummerorden.

Kontering af, eksempelvis telefonregninger, kan gøres til en selvopfundet konto - "3" - denne konto behøver ikke tilhøre nogen "overkonto" eller på anden vis indikere hvad kontoen beskriver. Et menneske kan måske udlede (og revisoren gør muligvis) at "3" er et mobilselskab og (efter virksomhedens formål og branche) berigtige at konteringen er en telefonregning der tilhører "drift" i resultatopgørelsen. Det er dog ikke lige så nemt for en maskine at drage den samme konklusion. I princippet er der intet til hinder for at bogholderen fejlagtigt konterer mobilregninger til en salgskonto heller. Disse fejl bør revisoren opdage i dag.

## 4.3 Detaljegrad

Hvor der eksisterer krav til detaljegraden af konteringer i bogføringen, kunne disse med tiden gøres endnu skarpere for at hjælpe med at identificere bogføringsfejl. Mere data og mere detaljeret data vil uden tvivl være gavnligt i jagten på validering af konteringer.

### Delkonklusion

Såfremt ERST (eller anden myndighed) ikke opdager fejl i årsrapporten, kan de "leve" i bedste velgående i dag. Når processer automatiseres, vil man muligvis stille det maskineri (eller de personer) der udfører arbejdet til ansvar for fejl der accepteres i systemet. Det bør derfor overvejes hvordan denne problematik håndteres.

## 5. Databehandling

### 5.1 Indhentning af data

Posteringsdata vil være det der udgør størstedelen af det nødvendige data. Kontoplanen for en virksomhed vil typisk udgøre en relativt lille del af datamængden der skal overføres og behandles i forholdet til posteringer. Data indhentes typisk igennem en form for API men filoverførsler kan også være nødvendige. (Eksempler kan ses under afsnittet "dataenes form" punkt 1. Og 2. Billys og e-conomic).

Givet at data tilgås igennem et API vil det være nødvendigt at få autorisering for at tilgå en virksomheds data. For de to regnskabssystemer Billy og e-conomic er det nødvendigt at få en token der skal bruges i API kald for at hente data. Denne token skal lagres og bruges hver gang data hentes. Det kan variere i andre regnskabssystemer hvordan der opnås adgang til data og hvordan adgang beholdes efter at have fået den første gang.

Typisk adgang til data i praksis opnås ved at datamodtageren anmoder brugeren om adgang. Brugeren kan så her give datamodtageren en adgang til sit regnskab igennem regnskabssystemet. Herefter vil aftalen om adgangen være lagret hos datamodtageren men varetaget af regnskabssystemet, hvor den ligeledes vil være lagret. Brugeren kan via regnskabssystemet fjerne adgangen til enhver tid.

Det er eksempelvis brugerne der giver tilladelse til at Billy må videregive deres data. Tilladelsen gives engang i form af en token, der placeres i det eksterne software der er integreret til Billy gennem API'et. Herefter kan data hentes så ofte det ønskes. Denne token vil fungere indtil brugeren lukker den igen. Hovedparten af Billys brugere flytter data ud af Billy på denne måde. Det handler om at de enten kan se administrative fordele i det, eller at der er krav de skal leve op til.

API med sensitiv data (generel for alt web, ikke kun API) bør have et SSL certifikat, og kun tillade udveksling af data over HTTPS. Brugere bør ikke bekymre sig om certifikater, men adspurgt om hvad de vægter højest ved valg af ERP system svarer de "sikkerhed".

Hos Billy har vi eksempelvis erfaret at bankdata skal være underlagt certifikat udveksling, men vi ser

det ikke umiddelbart nødvendig. Kigger man på API'er fra Facebook, Google, Github etc., vil man ikke se noget om certifikater.

*Note: Certifikater* kan anvendes til autorisering men bruges sjældent i åbne API'er. Vi er endnu ikke stødt på et API der bruger den metode.

### 5.1.1 Mængde af data

Mængden af data der skal hentes vil variere meget imellem virksomheder. Nogle kan have få tusinde posteringer på et regnskabsår hvor andre kan have mange hundrede tusinde over samme periode. Her antages der relativt små virksomheder med få millioner i omsætning årligt. Større virksomheder vil typisk have betydeligt flere posteringer. Med den betydelige varians i datamængde kan det være problematisk at forudsige hvor mange ressourcer, det vil kræve at indhente data fra alle virksomheder.

## 5.2 Lagring

Indhentede data skal lagres og her vil man bruge en database. Typisk vil man vælge at lagre i en relationel database eller en dokument-baseret database. En udførlig beskrivelse af fordele og ulemper ved de to er uden for omfanget af denne rapport. En meget simpel generalisering er, at en relationel database typisk vælges for data integritet hvor en dokument-baseret database vælges for hastighed.

De faktorer der vil bestemme hvordan data bedst lagres, samt hvor mange ressourcer der skal bruges til at lagre samt skrive og læse til og fra lagret data, vil primært være datamængde samt hvad dataene skal bruges til og hvor hurtigt.

Datamængden er problematisk at vurdere da mængden for hver virksomhed ikke pålideligt kan forudsiges, som beskrevet ovenfor.

Hvad dataene skal bruges til er kendt: Generere årsrapporter samt være tilgængelig for at være i stand til at kunne undersøge enkelte posteringer. Et åbent spørgsmål kan være hvor langt tilbage posteringsdata skal være tilgængeligt?

Hvor lang tid ERST (eller anden offentlig myndighed) bør være i stand til at gå tilbage i tiden, hos en

virksomhed i dens bogføringsdata, er et spørgsmål kun ERST (eller anden offentlig myndighed) kan svare på. Det kommer an på hvilke krav der stilles. Skal ERST altid være i stand til at gå tilbage til ethvert regnskabsår og se posteringsdata?

Hvis ikke "uendeligt" hvor lang tid skal ERST så kunne gå tilbage? (er en fem årsgrænse tiltrækkeligt?)  
Er der andre der skal have adgang til posteringsdata (f.eks. SKAT)? Hvad er deres krav?

### **Delkonklusion**

En del af dataopbevaringsbyrden kunne med fordel lægges ud til ERP udbydere. Gældende regler foreskriver, at man skal opbevare regnskaber og tilhørende bilag 5 år bagudrettet. Overgår man til et system hvor regnskabsafloggeren er forpligtet til at indberette digitalt, vil det være forholdsvis nemt for ERP udbydere at opbevare bilag med tilhørende posterings data i 5 år. Dermed vil information altid kunne hentes igennem API'et. Dog vil det stille nogle krav til sikring af dataenes integritet.

Billys har eksempelvis en del cases hvor brugere har rettet i et regnskabsår efter, at perioden er afsluttet, det udgøre ikke et større problem, da ændringer altid vil kunne spores gennem posteringsjournaler, men man vil ikke umiddelbart kunne stole på en fuld P&L eller balance.

Vil ERST være 100% sikker på dataenes korrekthed, bør det overvejes at opbevare den selv. Hvorvidt denne historik skal opbygges efter projektets start, eller om ERST ønsker at indhente data fra projektets start og frem vil have stor betydning for efterfølgende fremgangsmåde.

Såfremt ERST ønsker at gemme posteringsdata for alle regnskabsår. Bør ERST være opmærksom på at trods en virksomhed har eksisteret i 10 år, kan der blot være 2 års bogføringsdata tilgængeligt i f.eks. bogføring systemet. Virksomheder skifter systemer fra tid til anden. Man vil ofte tage et skift ved en årsregnskabsafslutning, og ikke nødvendigvis overføre tidligere data fra system A til system B.

## **5.3 Hastighed**

Hvor hurtigt data skal bruges åbner op for yderligere spørgsmål der vil være nødvendige at afklare:

Hvor stor en periode fra data er tilgængelig for et regnskabsår til at et årsregnskab er genereret er acceptabel? Skal posteringsdata være tilgængeligt øjeblikkeligt per anmodning? Hvis ikke, hvor lang tid skal data så bestilles i forvejen?

Som reglerne er i dag, har virksomheder 5 måneder efter endt regnskabsår, til at aflægge regnskab. Der kan opstå et problem hvis ERST skal hente data fra 200.000 virksomheder inden for samme uge. Mange indrapportere dog tidligere. Med det i mente kunne man med fordel, implementere webhooks<sup>1</sup>, så den enkelte ERP bruger selv kan melde klar, og dermed igangsætte upload til ERST.

Dog må det forventes at der kan opstå flaskehalse ved deadline. En løsning kunne være at dataene skal være tilgængelig i minimum 3 måneder efter deadline.

### **Delkonklusion**

De problematikker lagring af data åbner op for vil typisk kunne afhjælpes med tilstrækkelig mange ressourcer, men de kan have betydelig effekt på langtidsudgifterne for et system. Hvor der er spørgsmål, er der naturligvis også løsninger. Løsningsmodeller kræver blot at der tages stilling til hvilke antagelser der bør tages her i blandt punkterne rejst ovenfor.

Lagring af store mængder af data til generering af årsrapporter, åbner op for spørgsmål til sikkerhedshensyn, som der vil blive diskuteret senere i rapporten.

## **6. Strukturering af data**

Når dataene er indhentet og lageret efter hensyntagen til ovenstående problematikker, er næste opgave at strukturere dataene på en vis der gør det muligt at opstille en årsrapport. Denne rapport opstiller tre muligheder hvorpå det kan løses: logisk, machine learning samt regler/lovkrav. En løsningsmodel hvor disse tre muligheder anvendes i kombination af hinanden er fjerde mulighed.

### **6.1 Logisk/regelbaseret opstilling**

Som nævnt i afsnittet "*kontoplaner på tværet af virksomheder*" eksisterer der helt frie rammer til konstruktion af kontoplaner. Dette afsnit vil beskrive "logiske/regel baseret" modeller hvorpå

---

<sup>1</sup> Læs apendix API polling vs. webhooks



bogføringsdata kan opstilles til en årsrapport på baggrund af eksempler hvor håndtering af data der er i) strukturerede, ii) mellem strukturerede og iii) helt ustrukturerede. Fordele, ulemper samt begrænsninger ved en logisk/regelbaseret tilgang vil afslutningsvis blive diskuteret.

### 6.1.1 Strukturerede data

Nedenfor er et eksempel vist, på hvordan en strukturerede kontoplan med en god orden kan opstilles til en årsrapport på baggrund af en regel baseret model.

En *strukturerede* kontoplan bør være opbygget omkring, samt indeholde, en logisk nummerorden, med tilhørende beskrivende kontonavn felt, konto type samt angivet om kontoen er momsbelagt. God bogføring indeholder fyldestgørende beskrivelser i de tilhørende tekstfelter til konteringerne.

Hvor hver konto har tilhørende konteringer, kan man udlede fra et granulerede datasæt hvor hver konto hører til, samt (pga. en høj granulering) relevant data, der evt. kan anvendes til validering af dataenes korrekte kontering. Dette vil blive diskuteret i et senere afsnit – *kvalitetskontrol*.

#### Eksempel på kontoplan (struktureret)

## 21000 Resultatopgørelse

Omkostninger:	Indtægter:
<b>Debet</b>	<b>Kredit</b>
2100 Vareforbrug	1100 Varesalg
3100 Salgsfremmende omkostninger	6100 Renteindtægter
3200 Lokaleomkostninger	
3300 Kassedifferencer	
3400 Bilers driftsomkostninger	
3600 Salgsfragt	
3800 Tab på tilgodehavender	
3900 Øvrige omkostninger	
4100 Lønafregning	
4200 ATP-bidrag	
5100 Afskrivninger på biler	
5200 Afskrivninger på inventar	
7100 Renteomkostninger	
7200 Selskabsskat	

## 22000 Balance

Aktiver:	Passiver:
<b>Debet</b>	<b>Kredit</b>
11120 Biler 11121 Akk. afskrivninger på biler 11130 Inventar 11131 Akk. afskrivninger på inventar 12110 Varelager 12210 Varedebitorer 12230 Periodeafgrænsningsposter 12310 Kasse 12320 Bank, folio 12330 Giro 14261 Købsmoms (indgående afgift)	13110 Egenkapital 13111 Overført resultat 14220 Varekreditorer 14230 Skyldig ATP 14240 Skyldig AM-bidrag 14245 Skyldig SP-bidrag 14250 Skyldig A-skat 14260 Momsafregning 14262 Salgsmoms (udgående afgift) 14290 Andre kreditorer

Opstilling af kontoplan på enkelt række

Opstilling af kontoplan til årsrapport



Figur 1.

En regelbaseret model kan læne sig op af den numeriske orden allokeret til div. Konti. I ovenstående eksempel allokeres eksempelvis konti 1100 til 3900 over **Bruttofortjeneste** og summen af posterne udgør derfor resultatet på denne mellemregningskonto. Såfremt der under kontoen **Salg** skulle konstrueres en underkonto som f.eks. "**1110 Salg uden moms**", vil en regelbaseret model kunne tage højde for dette og allokere kontoen korrekt.

Der kan uden problemer arbejdes med flere standarder og grupperinger til kontoplaner. Gøres dette, vil det være en fordel hvis hver standard identificeres med en kode, det muliggør at systemet kan genkende den valgte kontoplan og dermed anvende den korrekte "mapping". Mapping er betegnelsen for den proces der finder sted, når konti fra en kontoplan indpasses under de kategorier der er defineret til årsrapporten.

### 6.1.2 Mellem strukturerede data

En række virksomheder anvender en kontoplan med fortløbende allokering af numre til konti. Dette kan være et eksempel på en *mellem strukturerede* kontoplan. I denne model opretter virksomheden konti uden der eksisterer en egentlig struktur. Eksempelvis:

- 1 Salg
- 2 Vareforbrug
- 3 Varelager
- 4 Egenkapital
- 5 Bankkonto

Hvor ovenstående kan være tilstrækkeligt for at etablere en virksomhed, samt facilitere det første salg, kan virksomhedens krav til kontoplanen udvikle sig i takt med virksomheden vokser. Der kan f.eks. komme et behov for at tilføje "**6 Salg med moms**". Da der anvendes en fortløbende allokering af numre til konti, kan det være svært for en maskine at forstå at **konto 6** knytter sig til **konto 1** og dermed opstår der udfordringer. Der kan laves en regelbaseret model på disse kontoplaner, men den skal forventes at laves for hver enkelt virksomhed, da det ikke kan antages at virksomheder på tværs har allokeret de samme numre til de samme konti. Derfor vil det påhvile den enkelte virksomhed at mappe disse konti "korrekt". Det vil være en engangsøvelse for hver konto der oprettes.

### 6.1.3 Ustrukturerede data

Der findes en række virksomheder der har givet sig i kast med at lave en kontoplan selv, uden at have skænke struktur en tanke. Det er muligt i visse systemer at oprette konti uden at allokere et nummer til kontoen, mere præcist er det valgfrit. Det vil sige at en kontoplan kan se ud som vist nedenfor:

1100 Salg

Vareforbrug

1110 Varelager

310 Egenkapital

Bankkonto

Disse konstruktioner er hvad vi betegner som "spaghettikontoplaner". Det er yderst vanskeligt at udlede en struktur fra disse planer er derfor yderst vanskeligt (for ikke at sige umuligt) at lave en regelbaseret model til at opstille en årsrapport på baggrund af bogføringsdata med udgangspunkt i en numerisk opstillet kontoplan.

Man bør overveje, at tage udgangspunkt i en standardiseret kontoplan.

Der vil være en lang række fordele ved dette.

Det kan f.eks. være typen af virksomhedsform og/eller branchekoden, der er styrende for, hvilken konti i den standardiserede kontoplan der skal anvendes, herunder virksomhedens registrering af moms- og/eller lønsumspligt m.v.

Kontoplanen vil ud fra disse oplysninger kunne levere som slutprodukt et regnskab, hvor Danmarks Statistik og brancheforeningen får fornuftigt materiale - og særligt ensartet materiale. SKAT vil få et produkt, hvor linjerne vil være enklere, da nøgletal vil afsløre unøjagtigheder.

Hvis kontoplanerne låses fast, vil det altid være nemt at opdatere disse fra Erhvervsstyrelsens side og læserne vil altid vide, hvad der ligger i de forskellige poster og sammentællinger ved at læse en given vejledning på en offentlig hjemmeside ([www.Erst.dk](http://www.Erst.dk)). Vejledningens beskrivelse af kontoplanens indhold samt instruks til anvendelse heraf vil lette meget (måske alt) omkring regnskabspraksis. Dette vil lette hele regnskabets opstilling meget, særligt selvfølgelig på klasse B og C.

Med udgangspunkt i en standardiseret kontoplan, vil det udover at gøre det lettere for det offentlige at udsøge diverse oplysninger, også være muligt at opstille regnskabet lettere. Med en standardiseret kontoplan vil det være muligt ved en valgmulighed at ændre et regnskab fra at være artsopdelt til at være funktionsopdelt, da kontoplanen holder styr på de enkelte elementer i regnskabet.

Med indførelsen af Mikrovirksomheder åbner Erhvervsstyrelsen selv op for en ensartethed af regnskabspraksis, da det ved anvendelse af disse regler kun er muligt at måle finansielle aktiver/forpligtelser og investeringsejendomme til kostpris.

Det anbefales, at der findes afkrydsningslister efter hvilke poster, der indgår i resultatopgørelsen og balanceposter, hvorved beskrivelsen af anvendt regnskabspraksis i de enkelte regnskaber kan blive standardiseret.

Når regnskabet er indberettet til Erhvervsstyrelsen, bør det overvejes, om der skal kvitteres med en side, der opstiller virksomhedens nøgletal. Derved kan den administrative omkostning til levering af XBRL fil fra en ekstern leverandør spares helt i virksomhederne, samtidig med at regnskabslæseren får et hurtigt overblik.

Det kan overvejes om SKAT og Erhvervsstyrelsen, via et regnskabslogin, skal have adgang til diverse kontokort i bogføringen, således at data kan udtrækkes, og analyser kan trækkes direkte. Dermed kan angivelser af indberettede og betalt moms, A-skat, aktionærlån m.v. kontrolleres.

Derudover ville en adgang til virksomhedernes kontokort i bogføringen kunne afkorte sagsbehandlingstid ved f.eks. skattesager.

Det ville for eksempel forkorte sagsbehandlingstiden ved blandt andet konkursager, hvor det forholdsvis nemt kunne være gennemskueligt for kurator og/eller dommeren at se, hvilke aktiver der forefindes i virksomheden – eller fandtes på tidspunktet for konkursen.

Indføres en automatisk indberetning af alle fakturaer på B2B handler, gives der her virkeligt grundlag for muligheden for en meget enkel kontrol af omsætning og udgående moms.

Begreber som tilknyttede virksomheder og associerede virksomheder bør ændres til et begreb som forstås af menigmand.

Måske er det ambitiøst at få selskaber med ejerskab af andre virksomheder med, udover holdingselskaber og driftsselskaber, da man ofte vil løbe ind i problemer med skattelovgivningen i form af sambeskatning m.v.

Skat af årets resultat fremgår ikke automatisk af bogføringen, men beregnes årligt på baggrund af bogføringen. Skat skal indgå i resultatopgørelsen, da det er en post i regnskabet, uanset om dette opstilles som funktionsopdelt eller artsopdelt. Det bør derfor overvejes, hvor stor en del af skattelovgivningen, der skal medtages i automatikken.

Regnskabelementer som salgsfremmende omkostninger indeholder bl.a. poster som repræsentation, hvor der skal foretages skattemæssige korrektioner. Ved bogføring på den rigtige konto vil der automatiske korrigeres for dette skattemæssigt.

Det samme vil være muligt for ikke fradragsberettigede omkostninger, tilgang af driftsmidler hvor der

skal foretages afskrivninger fremfor straksfradrag m.fl.

Jo mere integreret og digitaliseret et regnskabssystem vil være med andre registrer, jo lettere vil det være at opstille et tilfredsstillende regnskab med få valgmuligheder. Selskabets registrerede kapital kan kontrolleres til det offentlige register hos Virk.dk og ejerforholdene kan kontrolleres til ejerregistret.

Bankengagementer burde indberettes til et offentligt register, og ikke kun banksaldi. Derved vil garantistillelser, forpligtelser og hæftelser udover bankindestående og -lån kunne afstemmes til det bogførte, inden regnskabet endeligt opstilles.

En mere digitaliseret platform vil kunne forenkle indsamlingen af information til brug for opstilling af regnskabet, samt forenkle kontrollen af regnskaberne for både regnskabslæseren og de offentlige instanser.

#### *6.1.4 Diskussion af notekrav*

Der er forskellige noter som bør/kan indgå i en årsrapport som f.eks. note til egenkapital, personale (såfremt der er ansatte), kapitalandele i andre selskaber (såfremt disse haves), operationelle leasing aftaler (såfremt disse haves), tilbageværende periode på lejeaftaler (såfremt disse haves) mm. De to sidstnævnte eksempler er tekst noter, og kan være vanskelige at genere, særligt da informationerne ikke eksisterer i bogføringsdataene. De tre første nævnte noter kan opstilles under samme forudsætninger som diskuteret ovenfor.

ERST bør tage stilling til om hvordan notekrav til tekst bør håndteres. Det er yderst relevant for beskrivelsen af virksomhedens situation om der eksisterer både leasingaftaler og lejekontrakter af f.eks. betydelig størrelse.

## **6.2 Machine learning**

Machine learning åbner op for nogle helt nye tiltag mht. til læsning af bogføringsdata, opstilling af årsrapporter samt validering af data. Grundstenen for hvilken som helst machine learningsmodel er store mængder af data. Machine learning kan opdeles i tre underkategorier, der vil kunne anvendes til det ønskede formål her i rapporten: i) Supervised learning, ii) Unsupervised learning, iii) Reinforced learning.

Det vil være svært at anonymisere data så den enkelte virksomhed, så den ikke vil kunne genkendes.

Billy har tidligere stillet data til rådighed for andre virksomheder, til machine learningsformål. I disse tilfælde har Billy dog "ansat" en eller flere udviklere der har arbejdet på servere sat op og ejet af Billy. De har altså ikke fået noget data ud af Billy, men midlertidigt fået adgang til Billys, udelukkende for at kunne træne deres computere. Årsagen til at det ikke er muligt at anonymisere dataene, er at ML skal bruge en række data på bilagene, til at afgøre hvordan et bilag skal behandles. Eksempelvis vil branchekoder på de enkelte, både leverandører og kunder have stor indflydelse på hvordan et bilag skal bogføres. Derfor er CVR.nr. på begge nødvendige at have på bilaget.

Endvidere vil indholdet af Billys databaser langt fra være nok til at lave en pålidelig model. Først og fremmest må Billy ses som et nicheprodukt, da det kun dækker 4% af markedet.

Billys kunder vælger systemet pga. brugervenlighed og ikke avancerede funktioner. Dvs. at de regnskaber der er data på i Billy er meget lidt komplicerede. ERST som offentlig udbyder skal kunne favne alle, hvilket vil stille enorme krav til datasættet.

Billy har i dag ikke anvendt ML til at udarbejde regnskaber. ML har Billy udelukkende brugt med henblik på at automatisere bogføring.

Vores forståelse var oprindeligt at Billys evt. kunne stille rå bogføringsdata til rådighed, ikke et datasæt til ML formål.

Det er vores opfattelse at de datasæt ERST har brug for et kombineret datasæt med eksempelvis 10.000 sæt rå bogføringsdata og hvordan hver enkel postering er håndteret i f.eks. CaseWare. Med andre ord, Billy har ikke et datasæt der kan lære en computer at lave årsregnskaber. ERST skal have fat i CaseWare, men der vil man mangle bilagene, så det vil ikke kunne lade sig gøre at vurdere om en given postering er korrekt.

Vores anbefaling vil være at starte ERST egen dataindsamling eller alliere sig med en udbyder som indsamler disse data, hvor dataintegriteten kan sikres.

ERST skal kunne sammenholde alle rå bogføringsdata med det endelige regnskab, for at computeren skal have en chance for at finde mønstrene.



ERST skal have et enormt datasæt, da i skal kunne dække alle cases, ikke kun de mest normale.

Selv om ERST skulle kunne erhverve sig store mængder af data fra f.eks. et eller flere cloud accounting systemer, bør ERST være opmærksom på at der er rod i branchekoderne, som bemærket tidligere. Branchekoder er DEN vigtigste faktor når du skal automatisere bogføring vha. machine learning. ERST bør fokusere på at få allokeret virksomhederne til de "rigtige" branchekoder før ERST igangsætter læringssæt med afsæt i bogføringsdata. Såfremt datagrundlaget der dannes læring på ikke er godt, vil læringen ikke blive god.

### 6.2.1 *Supervised learning*

**Supervised learning** er hvor et eller flere mennesker "underviser" computeren i den ønskede adfærd. Det kunne i denne sammenhæng være at ERST udstyrer en algoritme med en eller flere revisorer der bliver præsenteret for bogføringsdata datasæt, hvor han/hun eller de foretager den "korrekte" mapping af posterne til en årsrapport. Computeren får her mulighed for at følge de handlinger der foretages og kan "lære" af dette. Processen dikterer store krav til dokumentation af processen til revisoren eller revisorerne, for at maskinen får noget værdifuldt ud af det. Det umildbare problem er stadig, som bemærket tidligere, at to revisorer med samme bogføringsdatasæt vil muligvis opstille to forskellige årsrapporter. Sådanne hændelser vil "forvirre" computeren, da det er sværere at lave korrelationer mellem datagrundlaget og resultatet.

### 6.2.2 *Unsupervised learning*

**Unsupervised learning** er hvor ERST f.eks. føder computeren med en masse datasæt – her bogføringsdata SAMT de årsrapporter (det endelige resultat) og gerne med specifikationerne.

Computeren kan muligvis identificere sammenhængene og bruge denne metode til at genere fremtidige årsrapporter. Problemet her kan dog være at, de eksisterende datasæt er lavet under en bestemt lovgivning, såfremt lovgivningen ændre sig, burde opstillingen af årsrapporten også ændre sig relativt til de relevante ændringer.

### 6.2.3 *Reinforced learning*

**Reinforced learning** er kort sagt en kombination af de to modeller ovenfor. Computeren får lov at genere resultater på baggrund af de data den fødes med, hver gang den producerer et "korrekt"

resultat får den at vide af en revisor at det er korrekt, hver gang den producerer et forkert resultat får den at vide den er "forkert". Denne guidende proces arver både fordelene og ulemperne fra supervised og unsupervised machine learning.

### **Delkonklusion**

Det er Billys overbevisning at opgaven med at fremskaffe, strukturere og træne computeren på nuværende tidspunkt er for stor til at være ét projekt. Opgaven bør udføres over tid i delopgaver der fokuserer på at opsamle data, strukturere dem mm. Det har været nævnt tidligere, er der en del uregelmæssigheder i CVR registret. Der bør fokuseres på hvordan ERST kan skaffe et datagrundlag for senere ML fokus.

En start kunne være at lave tvungen regelbaseret regnskabsafregning, hvor rådata uploades til ERST servere (eller anden udbyder), efterfølgende præsenteres brugeren for en række spørgsmål, der sammenholdt med virksomhedens stamdata, vil danne grundlag for udarbejdelse af regnskabet. Det er dette Billy har lavet i samarbejde med Letregnskab. Det er vores overbevisning at det med en målrettet indsats kunne udvikles og testes inden for en årrække.

Data fra denne proces ville være velegnet til senere ML. Man bør være opmærksom på at ML er baseret på historisk data, hvilket vil betyde at data, fra før evt. lovændringer vil være forældet, og computeren/algoritmen vil have brug for ny træning, der kun kan ske som supervised ML, da der ikke findes historisk data.

## **7. Opstilling af årsrapporten (numerisk)**

Årsregnskabsloven åbner op for at virksomhederne selv kan vælge hvordan de opstiller deres årsrapport – artsopdelt eller funktionsopdelt. Fra og med driftsomkostninger er de to forskellige modeller præcis ens og resultat efter skat er upåvirket af modellen anvendt. Trods resultatet er enslydende metoderne imellem spiller det naturligvis en væsentlig rolle hvorledes virksomheden klassificerer sine indtægter, omkostninger aktiver mm. Sammenligningen virksomheder i mellem er noget mere besværligt såfremt der anvendes to forskellige metoder. Det vil ydermere tilføje noget mere kompleksitet at acceptere både artsopdelt og funktionsopdelt indberetning til machine learning, såfremt ERST ønsker at udforske mulighederne i den retning.

Det kan antages at få virksomheder i regnskabsklasse B vil vide hvad forskellen er mellem artsopdelt og funktionsopdelt regnskab og derfor ikke vil vide hvilken form de skal angive at opstille regnskabet efter. Det kan derfor diskuteres om hvor vidt det er relevant for virksomhederne at træffe den beslutning eller ej.

## 7.1 Artsopdelt vs. Funktionsopdelt

Figur 1. ovenfor illustrerer hvordan bogføringskonti kan opstilles til en artsopdelt årsrapport. Princippet bag en opstilling til en funktionsopdelt årsrapport er meget sammenlignelig.

Bogføringen kan deviere mellem de to opstillingsformer. For at kunne opstille regnskabet korrekt bør ERST kunne identificere hvilken opstilling der er at foretrække. Principielt bør opstillingen være bestemt ud fra type af virksomhed, men der indspiller sig også kulturelle forskelle. Hvor Danmark har tradition for at artsopdele regnskabet, har Tyskland præcis den modsatte historisk. Hvis formålet er at angive den korrekte skat, er opstillingen som bemærket tidligere, underordnet. Såfremt formålet er at få en grundning forståelse af virksomhedernes udviklingen i aggregeret forhold, bør der tages stilling til de ovenstående problematikker rejst.

## 8. Kvalitetskontrol

Kvalitet er en vigtig parameter for et årsregnskab. Ansvar for at årsrapporten er retvisende ligger, som bemærket tidligere, hos ledelsen. Dog kan der ligge nogle muligheder i store datamængder for at kvalitetstjekke dataene hos virksomhederne.

Cloud accounting systemerne gør sig umage med at hjælpe virksomhederne med at bogføre korrekt. Disse systemer med hjælpeværktøjer udbedres løbende. Der er derfor allerede i dag iværksat private initiativer for at sikre en kvalitetsoptimeringsproces, som kan virke til fordel for ERSTs ønske om en automatiserede proces omkring opstilling af en retvisende årsrapport.

Såfremt ERST ønsker at bistå i denne proces, er der mulige tiltage der kan foretages betinget af at der fremskaffes et tiltrækkeligt stort datasæt.

ERST kan iværksætte forskellige tilgange til at kvalitetssikre dataene. Dette kan f.eks. gøres ved at

bede en computer stille en række spørgsmål til hver enkel bogføringskontering. Disse spørgsmål kunne f.eks. være til en salgskonto til en specifik kunde (1100):

- 1) Er dette en eksisterende kunde?
- 2) Hvis ja, ligner salget noget, som kunden normalt køber (pris, produkt, volumen mm.)?
- 3) Hvordan køber kunden almindeligvis over året (er der sæsonudsving)? Hvordan har kunden købt de sidste par perioder (f.eks. uger/md) (er der tilgang med f.eks. 10%, eller købes der mindre)?
- 4) Såfremt det er en "storkunde" der altid køber 50% mere en gennemsnitlig kunde, holder dette aktuelle køb sig inden for denne grænse?

Altså, bør ERST prøve at identificere en trend i købet/konteringen. Såfremt det er muligt, er der mindre grund til at slå alarm, end hvis det ikke er muligt.

Der kan på sin vis opstilles en række valideringsspørgsmål der tager udgangspunkt i historikken. Denne metode har naturligvis sine begrænsninger, da blot fordi ting er sket før behøver de ikke ske igen og omvendt, blot fordi hændelser ikke har indtruffet før betyder det ikke at de vil indtræffe igen, som f.eks. salg til en ny kunde.

En anden række spørgsmål der bør stilles ved årsrapportsafslutningen er til konti så som: løn, A-SKAT, moms, varelager samt anlægsaktiver.

Et lønsystem sender løbende afregninger til SKAT, ved afslutningen på året kunne det være fordelagtigt hvis ERST fik autoriseret en adgang til SKAT af virksomheden, der gjorde det muligt at tjekke om den løn der er bogført stemmer overens med den løn der er angivet til SKAT.

På samme vis kan man tjekke A-SKAT, AM bidrag og moms. Stemmer bogføringen overens med de løbende indberetninger til SKAT?

Ved gennemgang af konteringer der er (eller ikke bør være) momspålagt, bør ERST tjekke for korrekt angivet momskoder i bogføringssystemet. Almindeligvis vil en revisor tage nogle stikprøver for at berigtige bilagene.

## **9. Omdannelse af årsrapport til XBRL og PDF**

### **9.1 Antagelser**

Når den fornødne data til årsrapporten er indsamlet og tilgængeligt i systemet, vil det være muligt at generere årsrapporten i både XBRL- og PDF-format. Heraf bør det understreges at det er virksomhedens ansvar at dataene er stillet til rådighed og klar til behandling.

### **9.2 Validering**

Den tilgængelige data skal være afgrænset til den specifikke virksomhed, samt det pågældende regnskabsår. Heraf skal der udføres validering af om hvorvidt alle krav om en indleveret årsrapport kan opfyldes, ud fra det data-grundlag der er tilgængeligt, og om hvor vidt der skal foretages yderligere indrapportering af påkrævet data. Såfremt at der skulle være mangler, skal disse først opfyldes, før end at der udfærdiges en årsrapport til indberetning.

### **9.3 XBRL-dokument**

Først og fremmest bør XBRL dokumentet genereres, og opfylde standarderne herfor. Her kan der benyttes et eksisterende værktøj til formålet, afhængigt af at brugt platform, for at undgå at bruge tiden på at skrive et værktøj der kan generere XBRL. For at kunne generere dokumentet, er det i høj grad vigtigt at kontekst på den tilgængelige data er kendt, så det kan indskrives i dokumentet uden menneskelig handling. Dette er dog ikke et problem der skal håndteres ved generering af dokumentet, men i en tidligere proces, når dataene indhentes. Så snart dokumentet er genereret, er XBRL-dokumentet klar til at blive videresendt.

### **9.4 XBRL til PDF**

Med XBRL-dokumentet, har vi allerede et veldefineret overblik over hvilken data der er hvad. Herfra kan vi generere en årsrapport i PDF-format, opstillet helt efter eget ønske. Her har vi allerede oplysningerne om de påkrævede afsnit, og kan opstille et PDF dokument ud fra disse fakta.

### **9.5 Det tekniske aspekt**

Hvor vidt det kan lade sig gøre at bruge en eksisterende XBRL generator, afhænger af den valgte tekniske platform (programmeringssprog). Om det kan svare sig at tage højde for, er svært at sige. XBRL dokumenter vil også blot kunne genereres med en simpel template, der benyttes til at generere XML-dokumentet. Dette kan hurtigt være en bedre løsning, afhængig af kompleksiteten af de genererede dokumenter.

Ønsker man at generere et dokument programmatisk, kan dette også gøres med en såkaldt XMLWriter, hvor du kan tilgå din DOM (Document Object Model), og manipulere denne. Der findes typisk native API'er til dette, i de forskellige programmeringssprog, som vil kunne understøtte de behov der er, for at generere et XBRL dokument.

Alternativt, som tidligere nævnt, kan der skrives et template, som evt. kan renderes af en template engine. På den måde gør man det muligt at sende en kontekst (din data), til din renderer, som så kan parse dit template, og indsætte de fornødne data, baseret på forskellige kriterier og anden logik.

Til generering af PDF-dokumenter, findes der open-source løsninger, som kan håndtere dette. Herunder findes bl.a. [GhostScript](#), som udbyder et bredt interface til at kunne generere PDF-dokumenter, og gør det let at arbejde med.

Generering af filer, kan potentielt blive en tung proces, så det er også vigtigt at indberegne at denne proces, skal ske over tid, så serverne ikke bukkes under, lige når det er vigtigt at disse kan bruges.

## 9.6 Fra ERP rå data til Regnskabsbasis 2.0

At hente data fra ERP systemet er ligetil. Vi vil ikke komme med en uddybende beskrivelse, da det ligger frit tilgængeligt. API'et er den præcise beskrivelse af hvordan man får data ud af systemet. Det er fyldestgørende med komplet oversigt over hvordan dette gøres. <https://www.billy.dk/api/>

Som sagt sker indhentningen af data gennem ERP API'et. For at få adgang til Billy eller Economic brugere skal ERP brugeren give adgang, gives denne adgang vha. en token, som brugeren kan hente i deres ERP system. Denne token "copy pastes", til et felt oprettet til dette, i ERP brugerens regnskabsbasis 2.0 konto. Med denne token kan Regnskabsbasis 2.0 hente alle regnskabsdata

gennem API'et<sup>2</sup>. Dette gør Credewire allerede.

Regnskabsbasis 2.0 skal ses som en simpel kontoplan, tricket med at få den udfyldt, er at oversætte ERP brugerens kontoplan til Regnskabsbasis 2.0 format. Det er den proces vi igennem rapporten kalder "mapping". Mapping skal laves for hver enkel ERP bruger, og ikke blot for ERP systemet, og hver gang en kontoplan ændres skal ændringen ajourføres i mappingen. At det er nødvendigt at mappe hver enkel kontoplan, skyldes at ERP brugere i stort omfang tilpasser deres kontoplan til deres behov. Eksempelvis kan vi nævne at kun 2% af Billy's brugere benytter sig af Billy ERP standardkontoplan. Når man opretter sig som bruger i CredWire, har CredWire allerede lavet denne mapping, læs afsnit 4. Derfor kan Regnskabsbasis 2.0 med fordel hente dataen fra CredWire, og undgå at skulle mappe 200.000 kontiplaner.

Her er et eksempel på en "reallife" resultatopgørelse-kontoplan i Billy, hurtigt (ikke nødvendigvis korrekt), mappet mod regnskabsbasis 2.0. Det skal også gøres for balancen.

Account Name	Type	Account type	
Indtægter	Group	Revenue	
Salg	Account	Revenue	
Varesalg	Account	Revenue	
Yaresalg	Account	Revenue	
<b>Total indtægter</b>	<b>Group sum</b>		
<b>Salgsomkostninger</b>	<b>Group</b>	<b>Expense</b>	
Andre direkte omkostninger	Account	Expense	
Yareindkøb	Account	Expense	
<b>Total salgsomkostninger</b>	<b>Group sum</b>		
<b>Bruttoresultat</b>	<b>Sum from</b>		
<b>Andre indtægter</b>	<b>Group</b>	<b>Revenue</b>	
Andre indtægter	Account	Revenue	
Kontant indbetaling	Account	Revenue	
Realiserede kursgevinster og -tab	Account	Revenue	
Rykkergebyrer	Account	Revenue	
<b>Total andre indtægter</b>	<b>Group sum</b>		
<b>Udgifter</b>	<b>Group</b>	<b>Expense</b>	
A-Skat personale	Account	Expense	
Afgifter (varme, el, olie o.l.)	Account	Expense	
Afskrivninger	Account	Expense	
Alm. personaleløn	Account	Expense	
AM-bidrag	Account	Expense	
ATP-bidrag medarbejder	Account	Expense	
ATP-bidrag Virksomhed	Account	Expense	
Bankgebyrer	Account	Expense	
Brændstof	Account	Expense	
Diesel	Account	Expense	
Diverse udgifter	Account	Expense	
Finansielle udgifter	Account	Expense	
Forplejning	Account	Expense	
Forsikring	Account	Expense	
Fragt	Account	Expense	
Gebjør	Account	Expense	
It	Account	Expense	
Konsulenter og freelancere	Account	Expense	
Kontorartikler	Account	Expense	
Kørselsgodtgørelse	Account	Expense	
Leasing Ydelser BA 69 119	Account	Expense	
Leasing Ydelser BB 24 446	Account	Expense	
Lokaler og husleje	Account	Expense	
Personalepleje	Account	Expense	
Reklamer og markedsføring	Account	Expense	
Rep BA 69 119	Account	Expense	
Repræsentation	Account	Expense	
Restaurant	Account	Expense	
Telefon og internet	Account	Expense	
Transport, hotel og rejser	Account	Expense	
Vedligeholdelse og inventar	Account	Expense	
Værktøj	Account	Expense	
Website	Account	Expense	

  

	kr.	kr.
Nettomsætning	<a href="#">note</a>	
Autosum <input checked="" type="checkbox"/> Vis ekstra felter <input type="checkbox"/>		
Produktionsomkostninger	<a href="#">note</a>	
<b>Bruttoresultat</b>		
Autosum <input checked="" type="checkbox"/>		
Distributionsomkostninger	<a href="#">note</a>	
Administrationsomkostninger	<a href="#">note</a>	
Andre driftindtægter	<a href="#">note</a>	
Andre driftomkostninger	<a href="#">note</a>	
<b>Resultat af ordinær primær drift</b>		
Autosum <input checked="" type="checkbox"/> Vis ekstra felter <input type="checkbox"/>		
Andre finansielle indtægter fra tilknyttede virksomheder		
Andre finansielle indtægter	<a href="#">note</a>	
Øvrige finansielle omkostninger	<a href="#">note</a>	
<b>Ordinært resultat før skat</b>		
Vis ekstra felter <input type="checkbox"/>		
Skat af årets resultat	<a href="#">note</a>	
Autosum <input checked="" type="checkbox"/>		
<b>Årets resultat</b>		

<sup>2</sup> I får ikke en mere uddybende beskrivelse, da API'et i sig selv er præcist denne beskrivelse.

Når denne mapping er gjort, kan data hente så ofte ERST måtte ønske det.

Næste spørgsmål er, hvor meget data ERST ønsker. Det ville klart være hurtigst blot at hente summer, men som beskrevet under afsnit 12.3, vil implementering af denne proces, kombineret med en regelbaseret afklaring af tvivlsspørgsmål, kunne bruges til at oparbejde det datasæt, ERST skal bruge for at automatisere processen vha. ML. Ønsker man at oparbejde dette datasæt der er nødvendigt at hente mere end blot summer. I stedet skal man hente hele posteringsjournalen, med tilhørende bilag, for at ML kan se, forstå, og lære af konteksten. Det skal også bemærkes, at før dette vil give et korrekt resultat, skal der laves efterposterings i ERP systemet. Der er tale om de rettelser, der bliver lavet efter ERP brugeren har overleveret rådata til revisor, men inden aflagt regnskab. Der er typisk tale om nedskrivninger, periodiseringer, hensættelser (bla. Feriepenge). Det er denne efterbehandling, ERST ønsker at automatisere. Det er allerede muligt at gøre det regelbaseret, men som beskrevet i andre afsnit, ser vi ikke at der eksistere et datasæt, der vil være velegnet til at automatisere vha. ML. derfor er vores anbefaling jf. afsnittet om mulige tiltag at man implementere en regelbaseret approach og bruger den til at oparbejde det nødvendige datagrundlag.

## **10. Indsendelse af opstillede årsrapporter**

### **10.1 Behandling af data**

Så snart at dokumenterne er genereret, er de klar til videre behandling. De kan her fortsætte i pipelinen, og blive afsendt til det eksisterende indberetningsværktøj. Her må antages at det er muligt at kommunikere disse filer, vha. et HTTP API, eller lignende. Når dokumenterne skal sendes, bør der fokuseres på en sikker transmission. Dette kan opnås ved brug af HTTPS, så der ikke kan opstå et MITM-angreb (Man in the middle), eller internt i et privat netværk, der er afgrænset for offentligheden, for yderligere forhøjet sikkerhed.

### **10.2 Kommunikation mellem systemer**

Når behandling og generering af dokumenter er afsluttet, ergo alt data er indhentet og rapporter er genereret, skal dokumenterne sendes videre til evt. videre- og/eller efterbehandling. Dette skal ske, som førnævnt, over en sikker forbindelse, så pågældende data ikke kan ses af uvedkommende, ved



dataoverførslen. Dette stiller, som før nævnt, et krav for at nuværende indberetnings-plattform kan modtage disse rapporter. Den data der skal sendes, er heraf bare rå-filer.

## **11. Sikkerhed**

### **11.1 Servere**

Opbevaring af store mængder af sensitivt data åbner op for overvejelser om sikkerhed. En relevant faktor her er hvor sensitivt den opbevarede data er. Bogføringsdata og årsrapporter er som udgangspunkt ikke persondata og de nødvendige overvejelser det ville medbringe er ikke relevante. Til gengæld er bogføringsdata sensitivt i form af det kan give dyb indsigt i en virksomhed og kan have betydelige konsekvenser hvis dataene bliver tilgængelig for konkurrenter eller andre der kan have gavn af industrispionage.

Det første niveau af sikkerhed der er nødvendigt at tage i betragtning er fysisk tilgængelighed. Givet fysisk adgang til en server er det væsentligt lettere at få uautoriseret adgang til den data der ligger på serveren. Derfor er det vigtigt at en servers fysiske beliggenhed er sikret. Normale foranstaltninger her er at placere servere i datacentre der er designet til formålet.

At have servere placeret i et datacenter gøres ved at købe servere af en hosting provider. Alternativet er at opsætte et datacenter selv men omkostningerne forbundet hermed er væsentlige og gøres typisk kun af meget store virksomheder, som for eksempel Facebook og Google.

### **11.2 Kryptering og opbevaring af data**

En måde at afhjælpe den ovennævnte risiko forbundet med fysisk adgang er anvendelse af kryptering. Ved at sørge for dataene kun kan læses givet den tilhørende nøgle sikres det at en kopi af dataene ikke kan læses uden også at have adgang til nøglen. Kryptering har dog den konsekvens at læsning og skrivning af data bliver langsommere. Der er teknologier der afhjælper dette men krypteret data vil være langsommere at arbejde med og kraftigere hardware må derfor bruges for at opveje hastighedstab. Nødvendigheden af kryptering er diskutabel. Det kan blive stillet som krav for at opnå en revisionserklæring og lignende, men ud over beskyttelse mod fysisk adgang er det begrænset hvad det beskytter imod. Selv hvis servere ligger i et datacenter kan denne beskyttelse kan da også være relevant, hvis det er nødvendigt at garantere at dem der har fysisk adgang (sikkerhedsvagter,

rengøringspersonale) ikke kan kopiere dataene.

Relevant er også at nævne hackerangreb. Dette udtryk dækker over en del forskellige typer angreb med forskellige formål. Overordnet er der tre forskellige formål: Nedlægning af service, tyveri af data, destruktion af data.

Forsøg på at nedlægge en service er typisk med et Distributed Denial of Service (DDoS) angreb. Her bruger angriberen et stort antal af klienter til samtidigt at tilgå en service. Formålet er at den betydelige ekstra mængde af trafik blokerer reelle anmodninger til serverne der har travlt med at behandle angriberens anmodninger. Det kan også få serveren til helt at gå ned. Denne type angreb er svært at beskytte i mod helt. Foranstaltninger kan dog tages i form af software designet til at opdage angreb og lukke ned for kilderne af det, eller manuelle processer til det samme. Konsekvensen af et succesfuldt angreb er at en service bliver utilgængelig, hvilket kan være betydningsfuldt, men intet data bliver mistet eller kommer i de forkerte hænder.

Tyveri af data og destruktion af data vil typisk have de samme fremgangsmåder men hvor tyveri dog kan være sværere at udføre. En fremgangsmåde her er den ovennævnte fysiske tilgang hvor en angriber opnår fysisk adgang til en server og kopierer dataene. Her kan kryptering hjælpe mod tyveri og backup hjælpe mod destruktion. Udefrakommende uautoriseret adgang, effektivt igennem Internettet, er væsentligt sværere hvis et system har en blot fornuftig mængde af sikkerhed. Med fornuftig menes her at der er beskyttet mod alle kendte typer angreb, hvilket i praksis ikke er indviklet. Det er dog langt fra alle servere der har en fornuftig mængde af sikkerhed, så det er bestemt muligt at efterlade åbninger for angreb hvis man ikke er bevidst om sikkerhed.

Succesfulde angreb på sikre servere vil oftest være igennem en person der laver en manuel fejl. En server kan være nok så sikker, men hvis en autoriseret person giver sin adgang væk til en angriber er al sikkerhed ligegyldigt. Logging af hvem der har tilgået data hvornår kan hjælpe med at finde ud af hvilken adgang der blev brugt til at udføre angrebet, men først efter angrebet er sket. De eneste mulige præventive beskyttelse her er begrænsning af adgang, manuelle processer og uddannelse. Begrænsning af adgang indebærer at så få som muligt har adgang til data og dem der har adgang kun har netop den adgang de har brug for. Manuelle processer kan for eksempel være hyppigt skift at kodeord og flere faktor autentisering. Uddannelse kan for eksempel inkludere hvordan kodeord opbevares og hvilke e-mails der ikke skal åbnes. I tilfældet af e-mails kan et anti spam filter også godt

hjælpe, men det kan ikke bruges alene.

## **12. Mulige tiltag**

Som det er nu, er teknologien for automatisering af årsrapporterne til stede. Dog mener vi ikke at den tilgængelige data er tilstrækkelig, eller af høj nok kvalitet til at projektet vil blive en succes. Med andre ord: forudsætningerne for udnyttelse af teknologien ikke til stede.

Vores klare anbefaling er derfor at man først fokuserer på at få forudsætningerne på plads, for udnyttelse af machine learning. Nedenfor er en række tiltag der vil øge muligheden for projektets succes.

### **12.1 Øge kvaliteten af den allerede tilgængelige data**

Som nævnt tidligere i rapporten, er der i dag store problemer med forkerte branchekoder. Dette udgør i sig selv, et problem for projektet. Billy har gennem sit arbejde med automatisering af bogføring, ved hjælp af machine learning, gjort en del erfaringer omkring dette. Eksempelvis er en af de mest beskrivende faktorer, for bogføring af et bilag, sælgers branchekode, da det typisk indikerer hvilken type varer der kan være tale om. Modsat er købers branchekode også vigtig, da den siger en del om hvordan køber har tænkt sig at bruge varen. Der er eks. stor forskel på hvordan et køb af kaffebønner skal bogføres hos henholdsvis en kaffebar og et konsulent hus.

Vi anbefaler derfor at der indledningsvis fokuseres på at sikre kvaliteten af virksomheds data i de offentlige systemer.

### **12.2 Implementere en standardisere konto plan**

Implementeringen af en fælles referenceramme er en nødvendighed for at ML vil kunne kategorisere og standardisere de data der kommer ind fra virksomhederne.

Årsrapporten er i sidste ende den mest forsimplede udgave af en fælles kontoplan. Korrekt afrapportering vil kræve en del flere detaljer for at kunne håndtere alle scenarier, samtidig er stort set alle kontoplaner i virksomheder meget mere detaljeret, hvilket nødvendiggøre en mapping fra

virksomhedens interne kontoplan, til den offentlige standard. Dette åbner igen op for spørgsmålet om hvem der skal stå for denne mapping?

Vil ERST selv stå for denne mapping? Det vil muligvis blive ret omkostningstungt, og meget tidskrævende, samtidig vil det være endnu en situation hvor man kan pådrage sig ansvar for generering af regnskaber er ikke giver et retvisende billede.

Skal data, der kommer fra ERP systemerne, være beriget med information om hvor i den standardiserede kontoplan, en given konto hører hjemme? Man kunne gøre det til virksomhedsejernes ansvar at mappe deres kontoplan mod den offentlige standard. Det vil lægge byrden med at gøre dette muligt over på ERP systemerne. Det er vores opfattelse, at cloud baserede ERP systemer hurtigt (inden for få måneder), vil kunne sikre at alle nye brugeres kontoplan som udgangspunkt er mappet. Udfordringen er at få alle eksisterende kontoplaner opdateret.

### **12.3 Igangsætte systematisk indsamling af manglende data**

Som nævnt igennem denne rapport, mangler der et fyldestgørende datagrundlag for at anvende ML til generering af årsrapporter. Det er nødvendigt med en ubrudt tråd fra bogføring af et bilag til et færdigt regnskab. Denne data ligger os bekendt ikke nogen steder i dag. Billy har data fra bilag til råbalance. Caseware har fra råbalance til årsrapport/skatteregnskab, hvilket ville svært at få et godt ML produkt ud af, da det mangler konteksten (bilagene og posteringsjournalen).

CrediWire vil på sigt ligge inde med et datagrundlag, der med lidt tilretning ville kunne bruges. Et alternativ til at opsøge dataene, kunne være at begynde at skabe den selv. Der er allerede muligheder for regelbaseret generering af årsrapporter, og ERST vil forholdsvist nemt selv kunne sætte en sådan løsning op. Vores tanke er at gøre det muligt for virksomhedsejere i selskabsklasse B, at eksportere deres data fra deres ERP system, som beskrevet tidligere, for derefter at gå gennem processen med at besvare nødvendige spørgsmål, for at systemet kan generere en årsrapport. Der findes i dag allerede fungerende løsninger, der kunne implementeres.

Det er et forslag at tilbyde det som en service, der er frivillig at benytte sig af, men bruge servicen det til at indsamle dataene. Et setup hvor ERST henter alt rå data, og bilag, for derefter at genere årsrapporten, vil netop danne det datasæt der er nødvendigt for senere at anvende ML.

## 12.4 Implementer infrastruktur der gør det muligt at hente data

ERST kunne starte med mindre datatunge projekter, der kræver samme infrastruktur. Her tænker vi på at man kunne lave indberetninger gennem API, eks. af moms, lønsum, fradrag, mm. Det kan ske gennem samme API'er som det endelige ML projekt skal bruge, men de kræver ikke et datagrundlag for at komme i gang. Mange brugere ville kunne se fordelene af at disse indberetninger sker automatisk, og vil med sikkerhed give ERST adgang<sup>3</sup>.

## 13. Perspektiveringer

Det er vores forståelse at projektet har det primære mål er at spare dansk erhvervsliv for den store administrative byrde, det er at lave årsrapporter. Endemålet er derfor ikke nødvendigvis at lave et system der har højere grad af troværdighed. Dog bør kvaliteten af rapporterne ikke forringes for meget som resultat at projektet. På sigt ser vi gode muligheder for at opnå både besparelser, og samtidig opnå en højere kvalitet af regnskaber.

Digitalisering af virksomheders økonomistyring og ikke mindst rapportering er stadig i sin meget tidlige begyndelse. Der er ikke endegyldige sandheder, da det samme bilag kan behandles forskelligt af to mere eller mindre ens virksomheder, men være behandlet korrekt i begge tilfælde. Samtidig er der massive lovmæssige krav til dokumentation, periodiske kontroller samt komplekse regler der skal håndteres. Dette gør det mere en almindeligt svært at programmere sig ud af.

### 13.1 Machine learning

Denne rapport har fokuseret på at anvende ML til udarbejdelse af årsrapporter. Vi vil her gøre opmærksom på at der er en række andre områder hvor ML også kunne implementeres, og for nogle, ville det også være lettere.

#### 13.1.1 Offentlig godkendt konterings motor

Har ERST først fået bygget en infrastruktur hvor de, gennem API adgange, kan hente alle posteringer med tilhørende bilag fra alle større ERP systemer (på samme vis som CrediWire har gjort over de

---

<sup>3</sup> Når vi siger med sikkerhed, bunder det i af Billy har brugere der ønsker en sådan løsning, og ikke kan forstå at Billy ikke kan udvikle den selv.

sidste 18 måneder), vil de ligge inde med en meget omfattende data, der vil være ualmindeligt godt egnet til at lære en computer at bogføre (det samme Billy forsøger med Billys egen data).

Er dataene først tilrettelagt vil en stor del af træningen muligvis kunne gøres unsupervised, hvilket vil stille en del mindre krav til resurser. ERST kan udvikle en offentlig bogføringsstandard for virksomheder, man kan vælge at bruge. Vælger en virksomhed at anvende denne, ville ERST vide hvilken grad af sikkerhed der var i det pågældende regnskab, og mindre virksomheder kan helt lade være at forholde sig til bogholderi.

### **13.1.2 Identifikation af snyd**

Ved at opsamle bl.a. rå bogføringsdata, sammenkoble det med virksomheder, og klassificere eksisterende regnskaber, som man ved at der har fusket i, vil man over tid kunne identificere senere regnskaber, og bede disse dokumentere deres bogføring. Dette vil over tid, kunne give et væsentligt bedre resultat på svindel-identificering, og føre relevante virksomheder i søgelyset, modsat de mere eller mindre tilfældige stikprøver der i dag udføres. Dette vil i starten kunne fungere som et supplement for stikprøverne, men vil på sigt kunne erstatte denne fuldstændigt, med en langt højere succes-rate. Ser vi på eksisterende virksomheder der forsøger at snyde på forskellige parametre, vil vi også se mange af de samme mønstre, som en maskine vil kunne lære af, og på den måde kategorisere potentiel svindel.

## **13.2 Kan ERST overhovedet levere en løsning der for alle?**

Machine Learning, er baseret på at lade computeren lære computeren "lære" reglerne for årsafslutning, ved at gennemgå store mængder eksempler på hvordan dette tidligere er været gjort. Der er altså tale om en sandsynliggørelse, Machine learning-tilgangen kræver ikke, at computeren kender reglerne, men derimod at computeren ser en række løsningseksempler. dvs. en form for sandsynliggørelse gennem sammenligning med lignende eksempler.

Her ser vi et potentielt problem. Var man en privat virksomhed, kunne man definere et bestemt segment man ønsker at ramme. Segmentet kunne vælges ud fra hvor komplekse det er og hvor tilgængeligt data er. Derefter kan man lave en løsning kun til dette segment.

Kan ERST, som offentlig instans, tillade sig at lave sådanne begrænsninger?

Hvis ERST ikke mener at de kan tillade sig dette, vil det stille meget store krav til det datasæt der skal indhentes. Dette skyldes at der er meget store variationer i skøn og regnskabsmæssig behandling af ens bilag. Det samme bilag, kan af ens virksomheder, behandles forskelligt, men i begge tilfælde være korrekt behandlet. Dette gør at computeren skal se rigtigt mange eksempler for at få afklaret alle scenarier og udfald.

Vi står tilbage med en tvivl, om hvorvidt det vil være muligt at skaffe et datasæt der vil gøre en computer i stand til at træffe korrekte beslutninger i alle cases. Det er vores overbevisning at det vil være svært at favne alle tænkelige scenarier, da ML som udgangspunkt er baseret på gennemgang af tidligere eksempler. Skal ML være i stand til at træffe beslutninger ud fra virksomhedens kontekst og kunne lave egentlige vurderinger af hensigtsmæssig behandling af en postering, vil det øge kravet til datasættets kompleksitet betydeligt.

### **13.3 Tredobbelt bogholderi**

I traditionel dobbelt bogholderi, indtastes hver transaktion to gange - den ene som en debet og en som en kredit. I dobbelt bogholderi, er det ikke muligt at have et aktiv, der kommer fra "ingenting", man kan heller ikke have en forpligtelse, der kommer ud af ingenting. De løbende ændringer i balance posterne kommer fra resultatopgørelsen, hvor ændringer i virksomheden økonomiske tilstand registreres og dokumenteres med et bilag. Er der balance i debit/kredit, kan en financen-manager med alt sandsynlighed stole på tallene, men for at øge troværdigheden i forhold til omverden, gennemføres der omkostningstunge uafhængige revisioner af virksomhedens finansielle informationer.

Når der tales om tredobbelt bogholderi, refereres der til at en økonomisk begivenhed ikke kan komme ud af ingen ting, og derfor må den pågældende begivenhed, have den modsatte postering i en anden økonomiskentitet(virksomhed)<sup>4</sup>. Valideringen af en transaktion og posteringerne i et regnskab, kommer hermed fra en anden virksomhed der tilkendegiver at de skylder, eller har betalt for en given transaktion. Dette vil skabe perfekt gennemsigtighed i erhvervslivet, og stort set gøre det umuligt at

---

<sup>4</sup> Se bilag 1.5 for eks

snyde. Deloitte kalder dette "a self auditing system" - [Blockchain\\_A game-changer in accounting](#)

Tredobbelt bogholderi kræver at der er en decentral offentlig database over alle transaktioner. Denne database skal være unhackerable, og uforgængelig. Dette har tidligere været uopnåeligt, men med Block Chain er det muligt.

Blockchain teknologi er netop en decentraliseret database over alle transaktioner. I tredobbelt bogholderi vil det være selve transaktionen i den decentrale database er der dokumentationen for posteringer i et virksomhedsregnskab, i stedet for et bilag. I et sådant system vil det offentlige have fuld adgang til alle informationer. Her vil det være nemt at vurdere hver eneste transaktion, og lige som man vil have alt tænkelig information i forhold til hvordan en transaktion bør behandles.



## Bilag 1.1

```
{
  "accountNumber": 1030,
  "accountType": "profitAndLoss",
  "balance": -484384.62,
  "blockDirectEntries": false,
  "debitCredit": "credit",
  "name": "Salg - DK - Licenser",
  "vatAccount": {
    "vatCode": "U25",
    "self": "https://restapi.e-conomic.com/vat-accounts/U25"
  },
  "accountingYears": "https://restapi.e-conomic.com/accounts/1030/accounting-years",
  "self": "https://restapi.e-conomic.com/accounts/1030"
}
```

## Bilag 1.2

```
{
  "account": {
    "accountNumber": 1010,
    "self": "https://restapi.e-conomic.com/accounts/1010"
  },
  "amount": -4900,
  "amountInBaseCurrency": -4900,
  "currency": "DKK",
  "date": "2013-10-18",
  "entryNumber": 60643,
  "text": "Kursus i Google Analytics Advanced, den 8/10-2013",
  "entryType": "customerInvoice",
  "vatAccount": {
    "vatCode": "U25",
    "self": "https://restapi.e-conomic.com/vat-accounts/U25"
  },
  "voucherNumber": 10578,
  "self": "https://restapi.e-conomic.com/entries/60643"
}
```

## Bilag 1.3

```

{
  "id": "imWwHxFCsrWaYxhKVbl5PA",
  "organizationId": "zqws9iOGTZciYEMCcHfbYQ",
  "name": "Afreignet moms",
  "accountNo": 7210,
  "description": "Momsafregningers beløb opbevares her efter du har indberettet til Skat, men inden du har betalt beløbet.",
  "groupId": "pvqlekSRTWku3axVKowvag",
  "nature": {
    "id": "liability",
    "reportType": "balanceSheet",
    "name": "Liabilities",
    "normalBalance": "credit"
  },
  "systemRole": "salesTaxPayable",
  "currencyId": "DKK",
  "taxRateId": null,
  "isPaymentEnabled": false,
  "isBankAccount": false,
  "isArchived": false,
  "bankId": null,
  "bankName": null,
  "bankRoutingNo": null,
  "bankAccountNo": null,
  "bankSwift": null,
  "bankIban": null
}

```

#### Bilag 1.4

```

{
  "id": "9U9CeOAZTfSkbXrGfFkCXg",
  "organizationId": "zqws9iOGTZciYEMCcHfbYQ",
  "transactionId": "fslyKSZORm2hF0Hb3qzq2w",
  "entryDate": "2015-12-24",
  "text": "",
  "accountId": "HwYVp3YRQLimaLUza5o0MQ",
  "amount": 12.5,
  "side": "credit",
  "currencyId": "DKK",
  "salesTaxReturnId": null,
  "isVoided": false,
  "isBankMatched": false,
  "priority": 3
}

```

#### Bilag 1.5

**Virksomhed A**

**Offentlig transaktionsliste**

**Virksomhed B**

**Transaktion 1**

Virksomhed A sælger skruer for 500 kr til virksomhed B, so lægges på vare lager

Resultatopgørelse	
Omkostning	Indtægt
D	K

250 | 500

Køber		Sælger	
Virk B	500	Virk A	500

Resultatopgørelse	
Udgift	Indtægt
D	K

Aktiver	
Kasse	
D	K

500

Passiver (EK)	
Egenkapital	
D	K

250

Aktiver	
Varelager	
D	K

250

Aktiver	
Kasse	
D	K

500

Passiver (EK)	
Egenkapital	
D	K

Aktiver	
Varelager	
D	K

500

**Transaktion 2**

Virksomhed B sælger tømmerarbejde med medmateriale forbrug til Virksomhed A

Resultatopgørelse	
Omkostning	Indtægt
D	K

1.500

Køber		Sælger	
Virk A	1.500	Virk B	1.500

Resultatopgørelse	
Udgift	Indtægt
D	K

1.500

Materialer 300  
Timer 700

Aktiver	
Kasse	
D	K

1.500 | 0

Passiver (EK)	
Egenkapital	
D	K

1.500 | 250

Aktiver	
Varelager	
D	K

Aktiver	
Kasse	
D	K

700 | 1.500

Passiver (EK)	
Egenkapital	
D	K

500

Aktiver	
Varelager	
D	K

300