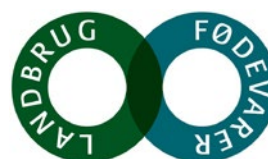






Danmarks  
Tekniske Universitet



Ministeriet for Fødevarer, Landbrug og Fiskeri  
Fødevarestyrelsen



# En fremtidssikret fødevareresektor

Dansk landbrug og fødevarerindustri er internationalt kendt og anerkendt. Knap 170.000 mennesker er beskæftiget i fødevareresektoren i Danmark, der i 2013 stod for en eksport på 156 mia. kr. til 180 lande. Fødevareresektoren har derfor stor betydning for Danmarks nationaløkonomi, valutabalancen og ikke mindst beskæftigelse. Det er vigtigt, at fødevareresektoren fastholder sin internationale position - og gerne styrker den yderligere. En af nøglerne til denne udvikling er Big Data.

Med denne rapport præsenterer Danmarks Tekniske Universitet (DTU) i samarbejde med brancheorganisationerne DI ITEK og Landbrug & Fødevarer en kortlægning af potentialet for innovation og vækst i fødevareresektoren ved at udnytte Big Data.

Hovedformålet med dette sektorudviklingsinitiativ er at accelerere udviklingen af nye løsninger inden for Big Data på fødevarerområdet. Sammen med virksomheder, brancheorganisationer og myndigheder har vi identificeret en lang række muligheder for Big Data i sektoren. Initiativet

har kortlagt, hvordan Big Data konkret kan udnyttes til at skabe bedre økonomi ved hjælp af optimeringen af driften, bedre fødevarerelitet, bedre fødevarerelikkerhed og øget sundhed for dyr.

Projektet er gennemført med deltagelse af DI ITEK, Landbrug & Fødevarer og førende eksperter fra DTU Compute, DTU Fødevarerelittuttet, DTU Veterinær-instituttet, DTU Systembiologi og DTU Management Engineering.

Arbejdet har kortlagt, hvad der skal til for at realisere potentialet i Big Data inden for forskellige områder af fødevareresektoren, bl.a. svineavl, kvægbrug og fjerkræavl. Rapporten indeholder et katalog af mulige forsknings- og udviklingsprojekter og et sæt anbefalinger, der kan accelerere udviklingen af den danske fødevareresektor. Vi er overbeviste om, at projekterne og en realisering af anbefalingerne samlet set vil bidrage til at fremtidssikre fødevareresektoren. Til glæde for både producenter og forbrugere.

God læselyst!

Niels Axel Nielsen  
Koncerndirektør, DTU

Adam Lebech  
Branchedirektør, DI ITEK

Jan Mousing  
Adm. dir., SEGES

Oktober 2015

## Redaktionsgruppe

Niels Axel Nielsen, koncerndirektør, DTU  
Bjarne Kjær Ersbøll, professor, DTU Compute (formand)  
Mark Bernhard Riis, Head of Innovation, DTU Compute  
Flemming Bager, afdelingschef, DTU Fødevareinstituttet  
Jørn Smedgaard, professor, DTU Fødevareinstituttet  
Nils Toft, professor, DTU Veterinærinstituttet  
Lovisa M. H. Suenson, innovationskonsulent, DTU Veterinærinstituttet  
Peter Jacobsen, lektor, DTU Management Engineering  
Peter Løngreen, Head of High-Performance Computing & IT, DTU Systembiologi  
Christian Graversen, chefkonsulent, DI ITEK  
Trine Vig Tamstorf, chefkonsulent, Landbrug & Fødevarer  
Morten Andersen, forskningsjournalist, City Pressekontor  
Jan E. Molzen, kontorchef, DTU  
Mads H. Odgaard, specialkonsulent, DTU (projektleder)

## Styregruppe:

Jan Madsen, professor, DTU Compute (formand)  
Christine Nellemann, institutdirektør, DTU Fødevareinstituttet  
Kristian Møller, institutdirektør, DTU Veterinærinstituttet  
Birthe Holst Jørgensen, viceinstitutdirektør, DTU Management Engineering  
Hanne Østergaard Jarmer, institutdirektør, Institut for Systembiologi

## Virksomheder:

Jesper Mogensen, udviklingschef og Søren Smedegaard, manager, Skov A/S  
Ole Kring, CEO, SMB A/S  
Urs Schuppli, Business Development Executive, Henrik Wieland, Associate Partner, Kim Escherich, Innovation Architect, Jakob Schuldt-Jensen, IT architect, Søren Peder Iversen, Delivery Project Executive, IBM A/S  
Morten Meldgaard, projektdirektør og Morten Danielsen, Department Manager, Chr. Hansen A/S  
Jesper Pagh, vicekoncerndirektør, Michael Agerly, svinefagdyrlæge, Jacob Dall, faglig chef, DLG  
Thomas Nejsum Madsen, direktør, TNM IT  
Jes Bjerregaard, adm. direktør, Jacob Roland Pedersen, senior manager, Danpo  
Kristian Kristensen, direktør og Thomas Holm, exportmanager, Danhatch  
Henrik Bunkenborg, chief consultant og Anders Langballe project manager, Lyngsoe Systems  
Vagner Bøge, direktør Ejerservice, Kurt T. Petersen, produktionsdirektør, Søren R Tinggaard, direktør, eksport,  
Karl Kristian Møller, chefanalytiker, Søren Rosenkrantz Riber, R&D Manager, Danish Crown  
Lene Vognsen, Senior Director Process & Technology, Arla Foods  
Christian Skøtt Maltesen, kundeindsigtsdirektør, Coop A/S  
Jan Cordtz, Big Data Architect og Jesper Døssing, Solutions Sale Big Data, Oracle  
Vibeke Christensen, afdelingsleder, Tine Hejgaard Sørensen, specialkonsulent, Johannes Frandsen, teamleder,  
Jørgen Nielsen, specialkonsulent, Jaap Boes, teamleder, SEGES/Kvæg  
Poul Bækbo, chefkonsulent, Henrik Thoning, teamleder, SEGES/Svin  
Vibeke Møgelmoose, afdelingsleder, Morten Andersen Linnet, afdelingsleder, Jan Dahl, chefkonsulent,  
Birthe Steenberg, sektorchef, Henrik Thoning, teamleder, Landbrug & Fødevarer  
Viktor Mayer-Schönberger, professor, Oxford Institute of Internet, Oxford University  
CEO og co-creator Birgitte Andersen, Big Innovation Centre  
Irene Lopez de Vallejo, Senior Partnerships Manager, Connected Digital Economy Catapult  
Chris Hankin, professor and director, Institute for Security Science and Technology, Dr David Stokes, Programme Coordinator (Cross Faculty) – Digital City Exchange, Yi-Ke Guo, professor, director, Data Sciences Institute, Imperial College of London

## Myndighed:

Annelise Fenger, kunde- og udviklingsdirektør, Kenny Larsen, sektionsleder, Anne Dyrhave Jensen, projektleder, Bente Grand, specialkonsulent, Fødevarestyrelsen

Henvendelser vedr. rapporten kan rettes til projektleder Mads H. Odgaard, DTU, maod@dtu.dk

# INDHOLD

<b>Resumé</b> .....	SIDE 6
<b>Værdien af Big Data i fødevarerektoren</b> .....	SIDE 12
<b>Udfordringer og løsninger</b> .....	SIDE 16
- <b>Kvæg</b> .....	SIDE 17
- <b>Svin</b> .....	SIDE 23
- <b>Fjerkræ</b> .....	SIDE 37
- <b>Salg af fødevarerdata</b> .....	SIDE 40
- <b>Forbrugersundhed</b> .....	SIDE 41
<b>Tværgående forskning i Big Data</b> .....	SIDE 44
<b>Konklusioner og anbefalinger</b> .....	SIDE 50

# Resumé



# Udnyt data fra jord til bord

Danske landmænd og virksomhederne i fødevarersekto- ren har gode forudsætninger for at drage nytte af den rivende udvikling inden for indsamling og bearbejdning af data:

- Danmark har en stærk fødevarer sektor. Det skyldes bl.a., at alle dele af værdikæden arbejder tæt sammen. Fra primærproducenterne, over forarbejdnings- industrien, agroindustrien til videns- og forsknings- miljøerne. Effektiv ressourceudnyttelse og fokus på optimering i hele værdikæden gør sektoren i stand til at konkurrere på verdensmarkedet.
- Danske fødevarer virksomheder har altid været gode til at opdyrke nye forretningsmodeller og finde nye in- novative veje til øget værdiskabelse. For eksempel gen- nem smartere måder at producere på, levere produk- terne på eller at indarbejde større værdi i produkterne, så de kan sælges med større fortjeneste.
- Dansk landbrug og hele værdikæden i fødevarersek- toren producerer store mængder af data. Det skyldes bl.a. et højt automationsniveau og myndighedernes krav til dokumentation af fødevarer kvaliteten, når de danske producenter leverer fødevarer til forbrugerne verden over.

Der er imidlertid et stort spring fra at råde over store mængder af data til at bruge dem aktivt i forretningsud- viklingen. Denne rapport viser, hvordan Big Data kan være ét af omdrejningspunkterne for den fremtidige udvikling i fødevarer sektoren.

## Kvæg, svin, fjerkræ og forbrugersundhed

Rapporten identificerer en række lovende forsknings- og udviklingsprojekter inden for kvæg- og svineproduktion, fjerkræbranchen og forbrugersundhed. Forslagene er til gavn for alle parter – for danske landmænd som for føde- vareindustri, forbrugere og samfundsøkonomi.

## Sektorudvikling

Sektorudviklingsprojekter er et af de værktøjer, som DTU bruger til at samarbejde med erhvervsliv og myndigheder. Målet med projekterne er at styrke teknologiintensive branchers konkurrenceevne ved at skabe overblik og handlingsplaner for anvendelsen af nye teknologier.

Metoden går ud på at:

- Kortlægge og analysere teknologianvendelsen i branchen.
- Identificere flaskehalse og udviklingsbehov både hos virksomheder, myndigheder og DTU.
- Bidrage til at løfte visionerne for teknologi og marked inden for en sektor.

Sektorudviklingsprojekterne bliver skabt i et forum bestående af repræsentanter for virksomheder, forskere fra DTU samt – hvor det er relevant – myndigheder.

## Interviewede organisationer

Skov A/S, SMB A/S, IBM A/S, Chr. Hansen A/S, DLG, Danpo, Danhatch, Danish Crown, Arla, Coop A/S, Oracle, Siemens, IBM, SEGES/Kvæg-Svin-Fjerkræ, Oxford University, Big Innovation Centre, Connected Digital Economy Catapult, Imperial College of London og Fødevarestyrelsen.

## Projektledelse

DI og Landbrug & Fødevarer har hele vejen arbejdet med at sikre den erhvervsmæssige værdi og forskningshøjde i de indsigter og projektmuligheder, der er resultatet af sektorudviklingsprojektet.

Arbejdet har været et samarbejde mellem fem DTU-institutter, hvor DTU Compute og Afdeling for Innovation og Sektorudvikling på DTU har været tovholdere i et tæt samarbejde med DTU Fødevarer- instituttet, Veterinærinstituttet, Management Engineering og Systembiologi.

## Indsatsen om at udnytte Big Data i fødevarerektoren bør organiseres i fire trin:

Trin 1: Forretningsudvikling

Trin 2: Udveksling og integration af data på tværs af værdikæden fra jord til bord

Trin 3: Tværgående Big Data værktøjer

I forbindelse med udarbejdelsen af rapporten har DTU opstillet en liste over de vigtigste Big Data værktøjer:

- a. Dataopsamlings teknologi
  - i. Via sensorer
  - ii. Eksisterende data / registre
- b. Forbehandling af data (tidligt stade)
  - i. Datareduktion
  - ii. Kun gemme relevante data
- c. Data lagring, teknologier til intelligent lagring
- d. Forbehandling af data (sent stade)
  - i. Dataheterogenitet og integrationsudfordringer
- e. Dataanalyse
  - i. Automatisk strukturering
  - ii. Dataanalytiske værktøjer og algoritmer
- f. Visualisering og beslutningstagen
- g. Datasikkerhed, privacy

Trin 4: Datasikkerhed. Håndtering af udfordringer om datasikkerhed og privacy, når man deler data

Arbejdet med rapporten har kortlagt, at indsatsen bør organiseres i fire hovedtrin (se boksen).

Første trin i processen har fokus på forretningsudvikling. Forudsætningen for, at et projekt vil lykkes i de senere led, er, at der er udsigt til forretningsmæssige gevinster. Eksempelvis udvikling af nye produkter, optimering af arbejdsgange eller mindsket ressourceforbrug.

Andet trin er integration af data på tværs af værdikæden fra jord til bord. Fra produktion af planter og dyr, forarbejdning til forbruger. Sektorudviklingsprojektet har demonstreret, at det er i krydsfeltet af data på tværs af de enkelte led i værdikæden, at de betydningsfulde indsigter ligger.

Tredje trin bør være udvikling og integration af de forskellige Big Data værktøjer fra den indledende dataopsamling, til lagring, analyse og anvendelsen af data. Data anvendes i store mængder, har stor varians, produceres meget hurtigt og har varierende kvalitet. Værktøjerne

skal bl.a. understøtte, at man kan besvare allerede kendte spørgsmål eller helt nye spørgsmål, som genereres i løbet af processen, samt at man genkender mønstre i data, der danner grundlag for erkendelse af nye indsigter, som ofte præsenteres i form af visualiseringer.

Endelig bør det fjerde og sidste trin angribe de udfordringer omkring datasikkerhed og privacy, som uvægerligt vil opstå i forbindelse med anvendelsen af Big Data. Hvordan sikrer man sig mod, at uvedkommende kan skaffe sig adgang til følsomme data? Og modsat: Hvordan sikrer man, at de relevante parter faktisk har adgang?

For en nærmere definition af datadrevet forretningsudvikling ved hjælp af Big Data og forskellen på datadrevet forretningsudvikling, Business Intelligence (BI) og datadrevet forretningsudvikling ved hjælp af Big Data henvises til Appendix 1. For nemheds skyld vil datadrevet forretningsudvikling ved hjælp af Big Data blive omtalt som Big Data i den resterende del af rapporten.





### Sådan kommer vi videre

Rapporten viser, at potentialet ved Big Data i fødevareresektoren er meget stort, men påpeger samtidig, at opgaven er kompleks. Hvis man for alvor skal opnå gevinst af Big Data, kræver det en samlet indsats fra en lang række aktører på tværs af virksomheder, videninstitutioner og myndigheder. Med henblik på at den danske fødevareresektor kan realisere potentialerne inden for Big Data fremkommer rapporten med fem anbefalinger.

#### Anbefaling 1: Høst de lavt hængende frugter

Rapporten viser, at den danske fødevareresektor producerer og anvender store mængder af data. Det gælder på besætningsniveau, i forarbejdningsledet, blandt producenter af udstyr, i detalledet og hos relevante myndigheder.

Rapportens hovedanbefaling er, at der igangsættes udviklingsprojekter inden for 11 specifikke områder identificeret i dette sektorudviklingsinitiativ. Det gælder problem-

felder inden for svin, kvæg, kyllinger og produktion af mælk. Inden for de identificerede områder er der store potentialer for anvendelse af Big Data, hvis der udvikles en række nye teknologiske løsninger.

Eksempler er:

- Projekter på besætningsniveau blandt svine-, kvæg-, mælk- og fjerkræproducenter. Det gælder bl.a. realtidsmonitorering af forbruget af foder og medicin.
- Projekter på slagterier og mejerier, hvor data opsamles på sti- og staldniveau. Desuden fremme af nye skanningsmetoder og robotudskæringer, udvikling af nye mælkeprodukter af meget ensartet kvalitet, samt tidlig indgriben over for trusler mod dyrenes sundhed og trivsel.

### **Anbefaling 2: Skab rammerne for et tættere samarbejde om data**

Realiseringen af potentialerne inden for Big Data forudsætter et betydeligt tættere samarbejde om data i hele værdikæden i fødevaresektoren. Rapporten anbefaler derfor, at der etableres et nationalt netværk eller klynge, hvor alle relevante aktører trækker i samme retning.

Rapporten peger på, at en afgørende faktor for at udvikle en klynge inden for området er, at der eksisterer en tillidsfuld ramme for at udlevere og dele data på tværs af individer, virksomheder, myndigheder og videninstitutioner. Rapporten anbefaler derfor, at de centrale aktører i klyngen i fællesskab arbejder for at skabe aftaleforhold om produktion, adgang til, lagring og brug af data, så de tillidsbårne samarbejdsrelationer opretholdes.

Klyngen kan bestå af virksomheder i fødevarerhvervet, IT-virksomheder, DTU, DI ITEK, DI Fødevarer, Landbrug & Fødevarer, SEGES og Fødevarestyrelsen.

### **Anbefaling 3: Styrk forskningen i tværgående Big Data værktøjer**

Anvendelsen af Big Data i primærproduktionen, i forarbejdningsindustrien og i agroindustrien kan på hver deres måde fremme udviklingen inden for kvæg- og svineproduktion og fjerkræbranchen. Men, sektorudviklingsprojektet demonstrerer, at det er i krydsfeltet af data på tværs af de enkelte led i værdikæden, at de betydningsfulde indsigter og dermed potentialet for forretningsudvikling ligger. Men, udvekslingen af forskellige former af data rummer en række store tekniske udfordringer. Det er derfor vigtigt, at en styrket dansk indsats inden for Big Data i fødevaresektoren sker i et samspil med en mere overordnet, generiske forskning i Big Data. Eksempelvis forskning i dataanalytiske værktøjer, teknologier til lagring af data og datasikkerhed. Kun på den måde opnår man, at løsningerne kan tale sammen, samt at smarte løsninger på ét område kan spredes til andre områder.

I det lys bør der etableres et center for forskning i Big Data. Et center for Data Science skal imødekomme behovet for forskning i indsamling og håndtering af data, dataanalyse, datadeling, fortolkning af data samt forskning i datasikkerhed og privacy.

### **Anbefaling 4: Uddan fremtidens medarbejdere**

En væsentlig forudsætning for at udnytte Big Data er, at kompetencerne er til stede i den private og offentlige sektor. Enten i form af nyuddannede kandidater eller opkvalificering af eksisterende medarbejdere.

DI ITEK vurderer, at den samlede efterspørgsel efter it- og elektronikkandidater frem til 2020 viser en mangel i størrelsesordenen 6.000 kandidater. Efterspørgsel efter ingeniørmæssige kompetencer specifikt inden for Big Data vil stige i takt med, at flere virksomheder ser mulighederne.

De danske videninstitutioner bør fremadrettet udbyde nye uddannelser og efteruddannelser, der giver den studerende kompetencer inden for:

- Big data værktøjer
- Ledelse og organisering af Big Data
- Viden om fødevarekvalitet og fødevaresikkerhed og kobling til Big Data
- Privacy, etiske hensyn og klare aftaleforhold i forbindelse med datasikkerhed

### **Anbefaling 5: Styrk iværksætteri inden for Big Data**

Iværksætteri og vækstvirksomheder spiller en væsentlig rolle i at skabe teknologisk udvikling i en erhvervssektor. Der er derfor behov for at indtænke iværksætteri i en samlet indsats for at styrke Big Data i fødevaresektoren og at udnytte det innovations-øko-system, der allerede er på fx universiteter, i offentlige programmer og i virksomheder. DTU arbejder løbende for at fremme iværksætteri og innovation og vil her medvirke til at fremme spinoff-virksomheder – også inden for Big Data.

Et væsentligt element i en sådan indsats er at sætte fokus på studenterinnovation. Det kan enten være i form af erhvervsrettede aktiviteter i forbindelse med de studerendes uddannelse, fx aktiviteter, der sikrer flere studenterprojekter i virksomhederne. Eller det kan være ekstra-curriculære aktiviteter såsom Hackathons, der kan føre til, at studerende udvikler Big Data løsninger for virksomheder eller offentlige institutioner, og at de studerende etablerer start ups baseret på disse løsninger.

## Projektet om Big Data fra jord til bord

Det vil kræve strategiske tiltag fra erhvervsliv, forskningsinstitutioner og myndigheder at indføre Big Data optimalt på fødevarerområdet. Samtidig er det væsentligt, at offentlige og private aktører samler kræfterne. På den baggrund besluttede DTU sammen med DI og Landbrug & Fødevarer at gennemføre dette sektorudviklingsprojekt, der har kortlagt, hvordan Big Data kan bidrage til vækst i fødevarerhvervene.

I den indledende fase i første halvår af 2014 blev en række virksomheder på tværs af værdikæden i fødevarer- og IT-sektoren interviewet. Det drejede sig om sensorproducenter, IT-leverandører, ejere og udviklere af databaser inden for landbrug og fødevarer, producenter af foderstoffer, ingredienser og fødevarer samt detailvirksomheder. De indsigter og hypoteser, som interviewene gav anledning til, blev drøftet på en workshop. På baggrund af tilbagemeldingerne fra workshoppen blev der gennemført flere interviews med nye virksomheder i anden halvdel af 2014.

Indsigter og hypoteser blev afprøvet igen på en international Big Data konference arrangeret af DI, Landbrug & Fødevarer, OPALL projektet og DTU i november 2014 og på en workshop i februar 2015 arrangeret i samarbejde med Fødevarestyrelsen.

Resultatet er denne redegørelse for en række tværgående metodologiske udfordringer samt et antal forskning- og erhvervsmæssigt værdifulde projektmuligheder inden for Big Data for bæredygtig, ressourceeffektiv fødevarerproduktion, sundhed for dyr og fødevarer kvalitet og -sikkerhed.

# VÆRDIEN AF BIG DATA I FØDEVARESEKTOREN



# Milliarder at tjene på Big Data i fødevareresektoren



Big Data kan styrke produktiviteten og indtjeningsvevnen i virksomheder. Forskere ved MIT har dokumenteret, at virksomheder, der er i top tre i deres branche og bygger på datadrevne beslutninger er 5 % mere produktive og 6 % mere indtjenende end deres konkurrenter (Kilde: McAfee og Brynjolfsson (2012): ”Big Data: The Management Revolution”, Harvard Business Review).

Analysen fra OECD, konsulentfirmaet McKinsey Global Institute og andre internationale eksperter viser, at stort set alle sektorer i samfundet kan høste store gevinster med Big Data.

## Sådan skaber Big Data værdi

En landmand kan fx bruge data fra sensorer og andet udstyr til at overvåge sin besætning. På den måde kan han forebygge sygdom og mistrivsel hos dyrene. Det betyder højere produktivitet, fordi dødeligheden blandt dyrene falder, og landmanden opnår en højere tilvækst per døgn. Samtidig mindsker han de udgifter, der er forbundet med syge dyr. Sidst men ikke mindst leverer han produkter af højere kvalitet.

Et andet eksempel er afsætningen af produkter på eksportmarkederne. Forbrugernes smag ændrer sig til stadighed. Samtidig er der vækst i nye lande, hvor middelklassen vokser og begynder at efterspørge fødevarer af høj kvalitet. Big Data kan være vejen for fødevarerindustrien til at opfange de nyeste tendenser, så man kan levere netop de varer, som de nye grupper af kunder efterspørger – inden konkurrenterne gør det.

Et tredje eksempel er skanning af slagtekroppe. De danske svineslagterier har allerede indført CT-skanninger. Skanningerne bruges til at planlægge opskæringen af det enkelte dyr, så man opnår høj kvalitet og minimalt spild, selvom processen overvejende er automatisk.

I fremtiden kan man udvide konceptet. Gennem de senere mange år har branchen forsøgt at fremstille slagtesvin, der er så ens som muligt både med hensyn til størrelse, smag, fedtprocent og så videre. Ensartetheden har været nødvendig for at kunne standardisere opskæringen og den senere forarbejdning. Men takket være skanningerne kan man åbne for en meget større variation blandt dyrene. På den måde kan man ramme markedet med udskæringer og smagsvarianter, der er tiltænkt bestemte grupper af kunder.

## Tingenes internet

Man behøver ikke kigge langt efter et eksempel på, hvordan det kan gøres. Vestas Wind Power forsyner virksomhedens vindmøller med sensorer, som overvåger vind- og vejrforhold samt holder styr på møllens driftstilstand. På den måde kan virksomheden lave et mersalg, hvor man kan rådgive kunderne bedre og samtidig udføre forebyggende vedligeholdelse, som er langt billigere i forhold til nedbrud.

Eksemplet passer med en global tendens om, at virksomheder, som leverer fysiske produkter, henter en stadig større andel af deres omsætning ved at sælge afledte serviceydelser. Ifølge det seneste regnskab fra Vestas kommer ca. halvdelen af omsætningen fra salg af serviceydelser,

hvor Big Data er et vigtigt element. En række andre danske virksomheder, heraf flere inden for fødevarerområdet, er begyndt at følge samme spor. Det gælder bl.a. Man B&W, Chr. Hansen, Foss, Danfoss og DLG.

Samtidig illustrerer Vestas-eksemplet et andet voksende fænomen, nemlig at sensorer og kommunikationsudstyr i stadig stigende grad fungerer autonomt. Man behøver ikke at kravle op i vindmøllen for at tjekke dens tilstand. Det kan man gøre online hjemmefra. Samtidig er det stadig mere udbredt, at indlejret udstyr selv er i stand til at "ringe hjem", hvis der er tegn på problemer. Eventuelt etableres forbindelsen via internettet, og nogle systemer sættes op således, at udstyret selv kan korrigere for et aktuelt problem omgående efter dialog med en central server. Man kalder også fænomenet "Internet of Things".

Ifølge en rapport udarbejdet af GTS-institutterne for Styrelsen for Forskning og Innovation er netop fødevarerområdet et af de vigtigste anvendelsesområder for Internet of Things. Rapporten taler om "Smart landbrug" i lighed med andre anvendelsesområder som Smart Energi, Smart Helbred, Smarte Bygninger, Smart Transport, Smart Living og Smarte Byer (Kilde: Smarte produkter og Internet of Things. Styrelsen for Forskning og Innovation. 2015).



### Udenlandske konkurrenter er i gang

Flere steder i udlandet ser man allerede på mulighederne i Big Data inden for fødevarerområdet. Det gælder specielt i USA, hvor man generelt er længere fremme i alle erhvervssektorer lige fra IT over sundhed til produktionssektorer mv.

Gennem de seneste år har amerikanske Monsanto investeret milliarder af dollars i opkøb af dataanalysefirmaer inden for landbrug for at sikre systematisk viden om, hvordan man optimerer afkastet på fødevarerproduktionen. Mange andre amerikanske virksomheder er også langt med at afsløre grænser for, hvad der er muligt i spændet mellem teknologi, lovgivning og marked med hensyn til udnyttelse af Big Data. Det har skabt et betydeligt løft i innovation på området.

I det hele taget bliver konkurrenterne på de globale markeder stadigt dygtigere. Danske fødevarer virksomheder må derfor være klar til at opdyrke nye forretningsmodeller og finde veje til at sikre en øget værdiskabelse i produktionen samt produkterne eller tjenesteydelser. Det kan enten være gennem smartere måder at producere på, levere produkterne på eller at indarbejde større værdi i produkterne, så de kan sælges med større fortjeneste. Ligeledes er landbrugets produktion og udviklingen af det danske råvaregrundlag væsentligt for vækstmulighederne gennem hele værdikæden i fødevarerhvervet.

### Fra mange data til Big Data

Landbruget og resten af fødevarersektoren producerer allerede enorme mængder af data. Det skyldes bl.a. et højt automationsniveau og myndighedernes krav til eksempelvis dokumentation af fødevarer sikkerhed, når de danske producenter og salgsled leverer mad til forbrugere. En landmand har en stor viden om sine jordarealer og dyr. Under hvilke vejrforhold skabes de bedste vækstbetingelser for afgrøderne? Hvilken betydning har vind og frost, og hvornår skal man sprøjte for ukrudt og skadedyr for at sikre den bedste produktion? Det er en viden, der bygger på erfaringer og historiske data og opgørelser. Nye teknologier og metoder til at udnytte data anvendes derfor i stigende grad i landbruget. Det gælder fx pc, Ipad, internetteknologi og GPS. I moderne mejetærskere er der fx GPS trackingsystemer, der fortæller landmanden, hvilken rute der er optimal at køre på marken, når han skal høste afgrøderne.

Det er imidlertid et spring fra at have en meget stor mængde af data til at udnytte dem gennem Big Data. I den forbindelse er det en udfordring, at små og mellemstore virksomheder (SMV'er) udgør en meget stor andel af erhvervsstrukturen i Danmark. SMV'er har typisk svært ved at afsætte store ressourcer til udviklingsprojekter. Samtidig er der tendens til, at SMV'er er tilbøjelige til at betragte IT i meget høj grad som en driftsopgave frem for noget, der kan skabe ny forretningsudvikling.

### Et dansk bud på etisk brug af Big Data

Selvom USA rent teknologisk må regnes som førstelandet indenfor Big Data, står det klart, at nogle satsninger har haft deres bagsider. For eksempel er det kommet frem, at kreditkortoplysninger på millioner af amerikanske kunder er lækket på grund af dårlig datasikkerhed, og der har været uheldige tilfælde, hvor virksomheder har truet kritiske journalister med at offentliggøre private oplysninger om dem.

Derfor er der stadig mulighed for, at Danmark kan blive førende. Nemlig ved at skabe forretningsmodeller, der udnytter Big Data på en etisk forsvarlig måde, hvor brugernes privacy respekteres, og brugerne har kontrol med, hvad der sker med deres data. Blandt andet er EU meget optaget af tanken om et europæisk perspektiv på Big Data.

Danmark har gode forudsætninger for at være med til at skabe unikke forretningsmodeller, produkter og tjenester, hvor man ikke kun har en enkelt virksomheds bundlinje i tankerne, men også øje for fællesskabets interesser.

Netop fødevareområdet kunne være et godt afsæt for en sådan dansk model for Big Data. Senere kunne den spredes til andre områder som for eksempel sundheds-, energi- eller detailsektoren.

#### Udnyt arven fra andelsbevægelsen

Den nordiske samfundsmode og vores rødder i andelsbevægelsen udgør et solidt fundament for, at Danmark kan udnytte Big Data på fødevareområdet. Vi har generelt et højt tillidsniveau, som gør, at vi er villige til at udlevere data om os selv i tiltro til, at data ikke vil blive misbrugt. Vi har en høj demokratisk bevidsthed, som gør, at vi både som borgere og forbrugere ikke alene stiller krav, men også deltager i at udvikle robuste løsninger på samfundets problemer. Samtidig besidder vi en høj grad af åbenhed og samarbejdsevner.

I den sammenhæng er det interessant, at DTU også forsker i, hvordan man kan sætte almindelige brugere i stand til selv at udvikle deres IT-løsninger og reelt udføre basal programmering. Dette er faktisk muligt, så længe man holder opgaverne inden for afgrænsede områder og samtidig beskriver de handlinger, som skal foretages, i et sprog, der er tilpasset den pågældende målgruppe.

Dette understreger det vigtige i, at man kobler projekterne, som skal indføre Big Data på bestemte domænerområder i fødevarebranchen, med tværgående, ”generisk” forskning. Der er al mulig grund til at tro, at værktøjer, der eksempelvis kan hjælpe landmænd med at udføre deres egen programmering også kan bruges af andre faggrupper – og omvendt.

Sammenfattende er opgaven med at indføre Big Data i fødevarebranchen præget af stor kompleksitet. Så meget desto større vil belønningen være for de lande og virksomheder, som formår at løse den.

## Eksporten af fødevarer er vigtig for Danmark

Fødevaresektoren har stor betydning for Danmarks økonomi og beskæftigelse. Med en eksport på 155 milliarder kr. årligt (ca. en fjerdedel af den samlede danske vareeksport) og med 183.000 beskæftigede er det et af Danmarks vigtigste eksporterhverv. Fødevareeksporten er i vækst på alle markeder og især på de nye markeder i Asien, Mellemøsten og Afrika, hvor de årlige vækstrater de sidste 5-6 år har været på 2-3 %. Halvdelen af fødevareeksporten består af landbrugsvarer, blandt andet kød, korn, mælk, ost og skind. Heraf står svinekød alene for over 20 % af den samlede fødevareeksport. En fjerdedel udgøres af agroindustrielle levnedsmidler, hovedsageligt forarbejdede fødevarer, blandt andet øl, chokolade, fedtstoffer og enzymer. Den sidste fjerdedel fordeler sig på eksport af fisk og fiskeriprodukter samt maskiner mv. til anvendelse i agroindustrien.

## Et bidrag til globale løsninger

Udnyttelsen af Big Data i fødevaresektoren kan bidrage til at løse store udfordringer både i Danmark og globalt. Det gælder særligt inden for tre områder:

For det *første* kan sektoren bidrage til at skaffe fødevarer til en voksende befolkning. Verdens befolkning forventes at vokse fra de nuværende 7 milliarder mennesker til 8 milliarder i 2025 og mere end 9 milliarder i 2050. Samtidig vil milliarder af mennesker i verden blive løftet ud af fattigdom, hvilket vil få den globale middelklasse til at vokse fra knap 2 milliarder til næsten 5 milliarder allerede i 2030.

For det *andet* kan sektoren bidrage til en mere miljøvenlig produktion. En stigende fødevareproduktion lægger pres på verdens miljø, klima og ressourcer. Den udvikling betyder, at der er en global efterspørgsel på løsninger, som muliggør en mere ressourceeffektiv fødevareproduktion, og som belaster miljøet mindst muligt. Samtidig forventes stigende efterspørgsel efter løsninger, der kan nedbringe spild af ressourcer i produktionen i andre erhverv. Det gælder fx behovet for at sikre en effektiv udnyttelse af de biologiske rest- og biprodukter fra fødevareproduktionen til fremstilling af biobaserede produkter og løsninger. Det har blandt andet betydning for områder som energi, plastik, tekstiler, kemikalier, medicin og biotek.

For det *tredje* kan sektoren bidrage til at sikre bedre sundhed. Den stigende globale velstand i mange lande har medført en stigning i livsstile relaterede sygdomme, fx fedme og diabetes, i takt med at befolkningen både er blevet rigere og lever længere. Den voksende globale middelklasse kan på sigt øge efterspørgslen efter sundere og mere ernæringsrigtige fødevarer samt efter ny viden og knowhow inden for sundhed, kost og ernæring mv.

# UDFORDRINGER OG LØSNINGER





# Udfordringer og løsninger for Big Data i fødevarerektoren

Kapitlet udgør et katalog af mulige projekter indenfor fem forskellige domæneområder (kvæg, svin, fjerkræ, salg af fødevarerdata og langsigtede perspektiver om forbrugersundhed)

For hver af projektmulighederne indenfor domæneområderne kvæg, svin og fjerkræ samt langsigtede perspektiver om forbrugersundhed er beskrevet:

- En vision
- Hvordan en indsats kan skabe værdi
- Hvilke konkrete Big Data værktøjer der er nødvendige
- Hvilke kompetencer det er nødvendigt at udvikle

- Hvilke privacy/etik og sikkerhedsmæssige udfordringer, der skal håndteres
- Hvilke teknologiske og organisatoriske udfordringer der skal håndteres
- Hvad der kan gå galt i løbet af det enkelte projektforsøg
- Hvilke partnere der er relevante

Gennemgangen af projektmulighederne demonstrerer, at forskellige typer af virksomheder inden for landbrug og fødevarerindustri samt myndigheder står med en række udfordringer og mulige løsninger, der har tværgående forskning som forudsætning.



## Udfordringer og muligheder for kvæg

Kvægbruget kan både udnytte Big Data i forbindelse med mælkeproduktion og i forbindelse med kødkvæg. Fælles for begge delområder er, at man har en unik mulighed internationalt set ved at udnytte Kvægdata-basen, som drives af SEGES. Kvægdata-basen indeholder variabler på

individniveau fx data om fødsel, reproduktion, sygdom, flytninger, død/slagtning, medicinering, data fra ydelseskontrol og mælkemængder. Desuden registrering af fedt, protein og celletal 11 gange pr år. Databasen er i princip umiddelbart tilgængelig.

# BIG DATA TIL KVÆG (I): "BIG MÆLK" VISIONEN

Den eksisterende kvægdatabase og IT-infrastruktur i dansk kvægproduktion skal udvides til at være rigtig Big Data, hvor sensorer i stalden, på mejeriet og på slagteriet leverer data til et datafællesskab med de eksisterende kilder. Dette vil styrke en stærkt integreret værdikæde fra jord til bord med en bæredygtig og ressourceeffektiv fødevareproduktion og øget sundhed for dyr og fødevarekvalitet og -sikkerhed som resultat. Det vil give mulighed for i højere grad at optimere på mælkekvaliteten og udvikle nye forretningsmuligheder og -koncepter.

## Big Data til kvæg (I): "Big Mælk"

For at holde sig konkurrencedygtigt har dansk kvægbrug behov for at øge produktionen af mælk. Samtidig er erhvervet pålagt en række miljøkrav. Det er kun muligt at øge produktionen uden at overskride miljøkravene ved at etablere en integreret værdikæde fra jord til bord.

SEGESs Kvægdatabase indeholder unikke historiske data på individniveau, som helt naturligt kan anvendes til Big Data inden for mælkeproduktion. Der genereres desuden en mængde andre data fra malkeudstyr og andre former for processtyr. Disse data bliver ikke overført til kvægdatabase, selvom de i princippet kan overføres. Det drejer sig blandt andet om mælkemængder, flowdata, målinger af temperatur og hjerterytme mv., brunstobservationer og bevægelsesmønstre. Området er stort, og der udvikles fra industrien på mange forskellige sensorer. Udfordringen er at få dem til at måle stabilt og være tilstrækkelig driftssikre.

Modellering af sådanne data forventes at kunne afdække særlige mønstre, som giver mulighed for at udvikle nye beslutningsværktøjer. Bl.a. kan landmanden på et tidligt stadie blive opmærksom på behov for intervention fx i forhold til ændringer i mælke kvaliteten, optimering af fodringsstrategi, forebyggelse af sundheds- og dyrevelfærdsproblemer i besætningen etc.

I takt med at besætningerne bliver større, vil den enkelte besætning levere større mængder mælk. Det betyder, at en enkelt besætning potentielt kan fylde en tankbil. Det betyder igen, at mælken kan håndteres særskilt. Det giver øgede muligheder for at anvende sensortechnologi til at sortere mælken og dermed udnytte den bedre mhp. forretningsudvikling af nye mælkeprodukter.

### Sådan skaber indsatsen værdi

Den øgede datamængde har værdi for både landmand, mejeri og slagteri. Detailviden om mælkekvalitet og kvantitet helt ned på det enkelte dyr giver mulighed for nye produkter af meget ensartet kvalitet. Fodringen af en ko kan baseres på den analyserede mælkekvalitet kombineret med online målinger eksempelvis fra sensorer i malkerobotterne. Sensorerne kan vise protein, fedt, somatisk celle tælling, dyrets kropstemperatur, hjerterytme, m.v. Mælken kan sorteres i separate tanke efter kvalitet til givne formål – fx særlige oste, yoghurt, etc. Staldene monitoreres med sensorer, kameraer samt mikrofoner, og køernes livskvalitet monitoreres. Opbevaringstanke på landbruget, mælketankbiler og tanke på mejerier monitoreres, ligesom mælkeproduktet under produktion, opbevaring, transport, salg monitoreres – hele vejen til kundens køleskab. Herved sikres sporbarheden.

Detailviden om det enkelte dyrs status og mulighederne for optimal intervention vil give en besparelse på kort sigt, og et yderligere kvalitetsløft og forbedret image for dansk landbrug på langt sigt. Kan der gribes ind på et tidligt stadie, så problemer med dyrenes

sundhed og trivsel forebygges, vil det give besparelser og øget produktivitet.

Detailviden om det enkelte dyr, foder, sundhed etc. kan have værdi for det enkelte slagteri. Fx kan kvalitetskontrollen planlægges ressourceeffektivt, så man bruger flest kræfter på kontrol i situationer, hvor der faktisk er risiko for problemer. Samtidig vil en forståelse af datasammenhænge øge produktiviteten, og antallet af driftsforstyrrelser som produktionsstop kan minimeres.

### Big Data værktøjer

Det kræver anvendelse af alle former for værktøjer inden for Big Data at realisere værdien.

### Kompetencebehov

Der er behov for kompetencer inden for sensorudvikling, logistiske problemer med data-flow og opbevaring. Dertil kommer et behov for forståelse af de vigtigste feedback mekanismer og af domæneviden om husdyr- og mejeriproduktion. Sidst men ikke mindst er der behov for kompetencer i strukturering af data (eventuelt også automatisk strukturering) samt i modellering og analyse. De nødvendige uddannelser i denne henseende dækkes i stort omfang af DTU's eksisterende udbud.

### Privacy/etik og sikkerhed

Det må sikres, at den enkelte landmand/besætning ikke "hænges ud", men derimod opnår fordelene ved, at pågældendes data indgår i et større hele. I øjeblikket må data i kvægdatabase bruges til "forskning, forsøg og udvikling" af "Dansk Kvæg", under forudsætning af at de er anonymiseret. Tredje part kan i visse tilfælde få adgang til data ved at indhente accept hos den pågældende landmand.

I det øjeblik data udvides med fx sensordata og data fra malkerobotter, vil der opstå behov for at justere på adgangen for de enkelte aktører og evt. af konkurrencehensyn at arbejde med anonymiserede data.

### Teknologiske og organisatoriske udfordringer

Der skal arbejdes med udvikling af sensorer, overførsel af data, sikkerhed for integritet, identificerbarhed af data, relevant modellering og visualisering/præsentation.

Det vil være nødvendigt at organisere en deling og evt. samproduktion af data på tværs af virksomheder (besætninger, mejerier, butikker), myndigheder (veterinær og fødevarerkontrol) og brugere. Her vil en udfordring blive at integrere data og skabe en fælles platform, hvilket vil indebære, at spørgsmål om ejerskab til eksisterende data og evt. nye data samt adgang til dem vil skulle håndteres.

### Hvad kan gå galt?

Kæden fra jord til bord er meget lang. Derfor kan det blive vanskeligt at optimere hele kæden på én gang. Dog er der mulighed for at arbejde med udvalgte led i kæden og arbejde sig frem skridt for skridt.

# BIG DATA TIL KVÆG (II): KØDKVÆG VISIONEN

Markederne i foregangslande som Storbritannien og Japan er på forkant med krav om stadig øget sporbarhed, mindsket medicinforbrug, og øget dyrevelfærd. Brugen af Big Data vil spille en væsentlig rolle i opnåelse af disse mål.

## Big Data til kvæg (II): Kødkvæg

SEGESs Kvægdatabase indeholder som nævnt i afsnittet ovenfor unikke historiske data på individniveau, som helt naturligt kan anvendes til Big Data for kødkvæg.

Modellering af sådanne data forventes at kunne afdække særlige mønstre, som giver mulighed for at udvikle nye beslutningsværktøjer. Bl.a. kan landmanden på et tidligt stadie blive opmærksom på behov for intervention fx i forhold til optimering af fodringsstrategi, forebyggelse af sundheds- og dyrevelfærdsproblemer i besætningen etc.

### Sådan skaber indsatsen værdi

Detailviden om det enkelte dyrs status og mulighederne for optimal intervention vil give en besparelse på kort sigt og et yderligere kvalitetsløft og forbedret image for dansk landbrug på langt sigt. Kan der gribes ind på et tidligt stadie, så problemer for dyrenes sundhed og trivsel forebygges, vil det give besparelser og øget produktivitet.

### Big Data værktøjer

Det kræver anvendelse af alle former for værktøjer inden for Big Data at realisere værdien.

### Kompetencebehov

Der er behov for kompetencer i sensorudvikling, logistiske problemer med data-flow og opbevaring, forståelse af de vigtigste feedback mekanismer og for kompetencer i strukturering af data – eventuelt automatisk strukturering.

### Privacy/etik og sikkerhed

Det vil være en yderligere konkurrencefordel at øge kvaliteten af dansk kvægbrug gennem forbedret sundhed. Dyreetisk er det også et positivt tiltag. Alligevel kan der være privacy issues. Derfor må det sikres, at den enkelte landmand ikke ”hænges ud”, men derimod opnår fordelene ved, at hans data indgår i et større hele. I øjeblikket må data i kvægdatabasen bruges til ”forskning, forsøg og udvikling” af SEGES, Kvæg under forudsætning af, at data er anonymiseret. Tredje part kan i visse tilfælde få adgang til data ved at indhente accept hos landmanden.



### Teknologiske og organisatoriske udfordringer

Det er muligt at oprette en første version på grundlag af de nuværende registreringer af data. Imidlertid er organiseringen en udfordring. En række virksomheder skal samarbejde på tværs af værdikæden – besætninger, dyrlæger, mejerier m.fl. – lige som fødevarekontrol, veterinærmyndigheder og brugere må involveres.

En version 2 vil især omfatte udvikling af nye og billigere sensorer til brug overalt i værdikæden på enkeltdyrsniveau. Overførslen af data må opgraderes gennem oprettelse af internet overalt i produktionsanlæggene. Både for gamle og nye data er det nødvendigt at sikre integritet, identificerbarhed, validering og kvalitetssikring. For at kunne udnytte data optimalt vil det være nødvendigt med løbende udvikling og forbedring af modellering og præsentation. Endelig skal relevant feedback fra modelleringen udnyttes til optimering.

### Hvad kan gå galt?

Den store længde af kæden udgør en risiko for, at man ikke kan opnå adgang til alle ønskede data. Der er dog mulighed for at sub-inddele kæden og optimere hvert led for sig. Imidlertid er det ønskeligt at optimere kæden som et hele på sigt.

### Partnere

Besætningsejere, SEGES, Kvæg, sensor-fabrikanter, Fødevarestyrelsen.



## Udfordringer og muligheder for svin

Svinebranchen er udfordret af ledig kapacitet. Med andre ord er produktionsapparatet beregnet til langt flere enheder svin, end der bliver produceret.

Erhvervet har derfor behov for at øge produktionen. Givet de gældende miljøkrav kan dette kun ske ved at etablere en bæredygtig og intensiv produktion med moderne stalde, teknologi og forbedringer i foder-effektiviteten.

Hertil kommer, at især slagteriernes lønkonkurrenceevne er under et stadigt hårdere pres fra slagterier i bl.a. Tyskland og Polen.

Svinebranchen skal derfor på en og samme tid lykkes med produktion af meget stor volumen af svin baseret på meget effektive produktionsprocesser og løntilbageholdenhed og samtidigt differentiere sig på højværdiprodukter ved hjælp af produktudvikling.

Ligesom i kvægerhvervet ønsker svinebranchen selv at gå videre og etablere en endnu mere integreret værdikæde fra jord til bord, satse endnu mere på økologi og på at producere velsmagende produkter.

Det er et udtalt og meget stort ønske, at svinebranchen bliver reguleret på en måde, som ikke hæmmer produktiviteten, men tværtimod fremmer vækst.

Det er en erklæret målsætning i svinebranchen at udnytte tilgængelige data til professionel, daglig og strategisk ledelse af den enkelte bedrift.

Branchen kan især udnytte Big Data inden for tre områder, nemlig kødkontrol, foderoptimering samt kortlægning af tendenser i forbrugerønsker til svinekød.

Med hensyn til optimering af foder findes der allerede et projekt ved navn GainMax i regi af DLG. Vi har derfor valgt at beskrive dels mulighederne for at udnytte Big Data i forbindelse med GainMax 1 og dels mulighederne i at oprette et nyt projekt, hvor man fra starten optimerer mulighederne for Big Data – her kaldet GainMax 2. Visionen for GainMax 2 er at inddrage en lang række faktorer ud over foderet for at optimere svinenes sundhed samt branchens produktivitet.



# BIG DATA TIL SVIN (I): KØDKONTROL OG KVALITETS- STYRING

## VISIONEN

Når projektet er gennemført, er der etableret CT-baseret fuldkropsskanning tidligt på slagtekæden (efter sortskraber) på alle større svineslagterier i Danmark. Resultatet:

1. Skanningsresultaterne kombineret med besætningsdata og oplysning om tidligere kødkontrolfund danner grundlag for opdeling af kroppene i de, der skal gennemgå standard kødkontrol, og de, der kontrolleres visuelt. Udgiften til kødkontrol er reduceret med 20 %.
2. Skanningen giver et totalbillede af kødprocent og spæktykkelser og bruges til at styre slagtebotter og opskæringsbotter i realtid. Udbyttet er forbedret med 3 %.



## Big Data til svin (I): Kødkontrol og kvalitetsstyring

Kødkontrol har i over 100 år været et essentielt element i fødevarerikket. Kontrollen har sikret, at kun kød fra sunde og raske slagtedyrr når frem til forbrugeren. Principperne for kødkontrol er stadig stort set uændrede, selv om de mulige sundhedsrisici er væsentligt anderledes i dag end for 100 år siden. Det betyder, at kødkontrollen er særdeles ressourcekrævende. I erkendelse heraf har ny EU-lovgivning gældende fra 1. juni 2014 gjort det muligt at basere kødkontrollen på visuel inspektion, med mindre andre oplysninger taler imod. At ændre kødkontrollen på danske virksomheder forudsætter imidlertid, at vigtige eksportmarkeder i 3. lande accepterer ændringerne og fortsat betragter kontrollen som sikker.

I 2014 er forbrugerrisici forbundet med kød især associeret med bakterier som Salmonella, der kan overføres fra dyr til mennesker – såkaldte zoonotiske bakterier – og med antibiotikaresistente bakterier. Ingen af de to typer af bakterier kan opfanges ved kødkontrol. Derfor står forbruget af ressourcer på kontrollen ikke mål med nytteværdien. F.eks. bruger kødkontrollen tid på registrering af brysthindear, som er uden sundhedsmæssig betydning for forbrugeren, men mere kan betragtes som en kosmetisk fejl ved produktet. En bedre udnyttelse af eksisterende data, offentlige såvel som private, kombineret med nye teknologiske muligheder giver imidlertid mulighed for at gøre kødkontrollen bedre og mere effektiv, samt i højere grad at integrere kødkontrol med slagteriernes kvalitetsstyring. En EFSA risikovurdering pointerer bl.a., at data fra produktionskæden – besætninger og slagterier – ikke udnyttes i tilstrækkelig grad i forbindelse med kødkontrol, hverken i forbindelse med mikrobiologiske eller toksikologiske risici eller til vurdering af risiko for smitsomme husdyrsygdomme eller dyrevelfærd.



Det grundlæggende princip i den nye EU-regulering er at fokusere intensiteten af kødkontrollen ud fra viden om den besætning, som slagtekroppen stammer fra. Det betyder, at mange sygdomsproblemer er knyttet til bestemte besætninger. EU's hygiejneforordning har som et af sine krav, at information om fødevarekæden er tilgængelige i forbindelse med slagting.

Relevante besætningsoplysninger er:

- Viden om forekomst af Salmonella
- Viden om medicinforbrug, herunder årsager til hvorfor der er brugt antibiotika
- Viden om kødkontrolfund, sidst der blev slagtet svin fra besætningen
- Leverandørbesætningens samhandel med andre besætninger
- Viden om, hvorvidt grisene har været på friland eller udelukkende er opdrættet indendørs

På dette grundlag vil det være muligt at definere svin fra besætninger med lav/ingen Salmonellaforekomst, lavt antibiotikaforbrug og "uden bemærkninger" ved sidste kødkontrol som udgørende en lav risiko i forhold til fødevarerikket og husdyrsygdomme. Det kan igen danne grundlag for at vurdere, om slagtekroppen kan undersøges visuelt, eller om det skal ske ved palpation og incision. Et velfungerende system bør på langt sigt kunne føre til, at kødkontrollen kan udføres på en stikprøve af slagtekroppene i stedet for på alle. Det vil imidlertid kræve ændring af den gældende EU-lovgivning.

Generelt kan produktionsstyringen optimeres ud fra et bedre datagrundlag. En forbedret CT-skanning vil give en bedre robotstyring og dermed optimere den samlede produktionsstyring. Sammenholdes CT-skanningen med øvrige data kan styring af de enkelte robotter og dermed den samlede produktionsstyring yderligere optimeres. Forebyggelse af forurening af slagtekroppen i forbindelse med udtagning af organer er helt central for opretholdelse af høj slagtehygiejne, herunder minimering af risiko for overførsel af Salmonella til kødet.

Optimeret produktionsstyring vil samtidig give en forbedret udskæring og dermed en bedre kødkvalitet.

Til at støtte fokuseringen af udgifterne til kødkontrol etableres CT-skanning af slagtekroppene før udtagning af organer. Skanningen har flere formål:

- Påvisning af sygdomsprocesser, der kræver kødkontrollens bevågenhed, herunder påvisning af læsioner, der vanskeliggør automatiseret udtagning af organer og dermed øget risiko for forurening af slagtekroppen
- Analyse af slagtekroppens fordeling af kød og fedt med henblik på forbedring af udbyttet ved opskæring og med oplysning om dens størrelse til prædiktion af sygdomsmæssige forandringer

### Sådan skaber indsatsen værdi

- Forbedret kødkontrol gennem øget inddragelse af informationer om besætning
- Økonomisk besparelse på virksomhedernes udgifter til kødkontrol
- Forbedret grundlag for brug af slagte- og opskæringsrobotter på virksomhederne og dermed forbedret udbytte og færre slagtedefekter

Der er således en gevinst for alle parter i kæden, svineproducent, slagteri, kødkontrol og forbruger.

### Big Data værktøjer

Tættere integration og bedre udnyttelse af eksisterende store mængder data fra besætninger, slagterier og den offentlige kødkontrol vil gøre det muligt at udnytte data bedre og samtidig kombinere disse data med billedanalyser på grundlag af CT-skanning af alle slagtekroppe. Den væsentligste metodologiske udfordring vil formentlig bestå i at udvikle algoritmer, der med høj grad af sikkerhed kan detektere sygdomsdefekter, der er relevante for kødkontrollen, eller som kan gøre det vanskeligt for slagterobotter at udtage organer. Men heller ikke samkøring af eksisterende data og udvikling af et beslutningsstøtteværktøj til brug for dels kødkontrol, dels slagteriet er en triviell udfordring. Endelig vil der være tekniske udfordringer i forbindelse med CT-skanning og med identifikation af slagtekroppe og hovedudskæringer.



### Teknologiske og organisatoriske udfordringer og privacy/etik og sikkerhed

Identifikation af slagtekroppens oprindelsesbesætning: Slagtesvinene mærkes med besætningens leverandørnummer, når de køres til slagteriet. Der anvendes en tatovørhammer, og tatoveringen kan kun aflæses af en trænet operatør efter sortskraber. For øjeblikket er denne operatør placeret sidst på slagtekæden, ved klassificeringen. Etablering af maskinlæsbar besætnings-ID vil være hensigtsmæssig. Det vil ligeledes forøge udbyttet af projektet, hvis der kan etableres et system til at påføre slagtekroppens ID flere steder på kroppen, så hovedparteringer er identificeret, fx ved hjælp af en kode påført med inkjet printer.

Samkøring af relevante besætningsdata og udvikling af beslutningsalgoritmer, der sorterer slagtekroppene: Data er tilgængelige i offentlige registre, men de skal samkøres i realtid, og relevante konklusioner skal udtrages og dels præsenteres for veterinærkontrollen, dels indgå i slagteriets processtyring. Selve samkøringen af data – især oplysninger om medicinforbrug i besætningerne

– vil kunne give anledning til etiske overvejelser, fordi besætningsejere vil kunne have den holdning, at de ikke vedkommer slagteriet. Det skal derfor overvejes nærmere, hvorledes dette kan håndteres.

Færdigudvikling af CT-skanning til brug i realtid ved slagtebåndshastighed: Danish Meat Research Institute (DMRI) arbejder på at udvikle en CT-skanner til brug på hele slagtekroppen. Der er fremdeles udfordringer i at få skanningen til at fungere på hele slagtekroppen med organer inden for kommerciel slagtehastighed.

#### Hvad kan gå galt?

Der vil være risiko for, at CT-skanning af slagtekroppe med organer viser sig ikke at kunne implementeres praktisk eller viser sig at blive for omkostningstung. Konsekvensen kan i så fald blive, at kategorisering af slagtekroppe i høj/lav risiko gennemføres alene på grundlag af eksisterende data. Produktivitetsgevinst for slagteriet bliver dermed mindre.

Et andet scenarie kunne være, at algoritmer til billedanalyse viser sig ikke at kunne påvise relevante sygdomsmæssige forandringer i slagtekroppen med tilstrækkelig stor sikkerhed. Konsekvensen kan blive, at kategorisering af slagtekroppe i høj/lav risiko gennemføres alene på grundlag af eksisterende data - evt. kombineret med forudsigelse på grundlag af oplysninger om forhold mellem kød- og fedtprocent samt vægt.

Desuden er der risiko for, at nogle besætningsejere ikke vil ønske, at slagteriet får kendskab til medicinforbrug i besætningerne. Konsekvensen kan blive, at beslutningsstøtteværktøjet udformes, så disse oplysninger udelukkende er til rådighed for kødkontrolpersonale.

I øjeblikket kendes i bedste fald kun dyrene på sti-niveau. Dvs. i grupper. Det vil være ønskværdigt med identifikation af enkeltdyr. På grund af det meget store antal dyr, der er tale om, er det naturligvis noget af en udfordring, primært af logistiske grunde. En mulighed kan være at udføre et pilotprojekt blandt interesserede besætningsejere, som kan bruges til at afdække fordele og ulemper, og dermed skaffe et mere reelt beslutningsgrundlag.



# BIG DATA FOR SVIN (II): OPTIMERING AF FODER – GAINMAX 1 VISIONEN

GainMax 1 har udviklet sig fra at være et pilotprojekt drevet af en sønderjysk dyrlæge-ildsjæl til at være et produkt, som viser potentialet ved udnyttelse af Big Data. Med udgangspunkt i det nuværende system vil det forholdsvis enkelt kunne lade sig gøre at tilføje ny og mere avanceret funktionalitet trinvist i dialog med DLG og interesserede landmænd.

## Big Data for svin (II): Optimering af foder - GainMax 1

GainMax drives af DLG. Grundlaget for systemet er, at den enkelte landmand vejer alle grisene i en eller flere repræsentative sti(er) ugentligt. Herudover registreres fodermængde og foderblandning, energiindhold mm.

På sin pc kan landmanden følge tilvæksten som kurveforløb. Tilvæksten bliver relateret til fodertype, forbrug etc. Herudover kan landmanden sammenligne med oversigter over det gennemsnitlige forløb på andre bedrifter. Der skabes på denne måde en opmærksomhed på udviklingen i tilvækst, og hvordan tilvæksten hænger sammen med foderforhold og andre forhold i stalden. Dette har bidraget til interessen for at indgå i erfa-grupper og udveksle viden mellem landmænd.

Det nuværende system ønskes udvidet med et analysemodul og visualiseringsmodul, som kan knytte usikkerheder på kurveforløbene, så man kan følge, hvornår et udsving er betydningsløst (ikke signifikant), og hvornår det er statistisk signifikant – og en eventuelt handling må foretages. Det er vigtigt, at den nye information fremstår så intuitiv som muligt.

Forskningshøjden i projektet kan øges gennem metoder fra statistisk forsøgsplanlægning, hvor (små) ændringer i fodringsstrategi bliver implementeret på mange gårde. Herved kan genereres værdifuld information om optimale fodringsstrategier. Dette kan suppleres med metoder, hvor optimeringen angår den enkelte landmand. Da ændringerne i foderstrategi er små, er den økonomiske risiko ved at indgå i sådanne forsøg begrænset for den enkelte landmand og for DLG.

### Sådan skaber indsatsen værdi

Løbende detailviden om besætningernes tilvækst giver værdifuld information om, hvordan netop denne landmand bør fodre sin besætning, samt potentielt også andre driftstiltag som fx ændring af lys, ventilation eller temperatur. Allerede den præsentation af data, der foregår i dag, har vist sit værd som værktøj til at skabe fokus om eventuelle problemer med tilvæksten og eventuelle årsager til disse. Desuden skaber programmet loyalitet hos landmændene i forhold til DLG.

### Big Data værktøjer

Udvidelse af det nuværende system med modeller og analyser af allerede eksisterende data vil utvivlsomt forbedre produktionen og formentlig også resultere i bedre dyrevelfærd og mindre sygdom.

### Kompetencebehov

Der er behov for kompetencer indenfor statistik og data-analyse samt intuitiv visualisering af resultater. Det er nødvendigt med forståelse af de vigtigste feedback mekanismer og nødvendigt med indgående viden om husdyrhold.

### Privacy/etik og sikkerhed

En udvidelse af GainMax vil skabe en yderligere konkurrencefordel gennem øget produktivitet og kvalitet i dansk svineproduktion. Dyreetisk vil det også være et positivt tiltag. Men samtidig er der udfordringer i forhold til privacy. Det må sikres, at den enkelte landmand ikke ”hænges ud”, men derimod opnår fordelene ved, at pågældendes data indgår i et større hele.

### Teknologiske og organisatoriske udfordringer

Umiddelbart er projektet tilgængeligt baseret på nuværende data. Projektet kan forbedres ved validering og kvalitetssikring af nuværende og kommende data. Metoder fra statistisk proceskontrol vil være oplagt at introducere.

Her benyttes statistisk baserede metoder til at vurdere, om et afvigende datapunkt er ”en enlig svale”, der ikke skal reageres på, eller om der er tale om en blivende ændring i produktionen, som kræver et indgreb.

Det er naturligvis vigtigt, at nye muligheder fremstår så brugervenlige og intuitive som muligt. Ellers er der stor risiko for, at de ikke bliver brugt. Den nødvendige organisering på tværs af virksomheder (gårde, dyrlæger, DLG) er allerede på plads. Gårdejerne er bekendte med produktet og vil derfor hurtigt kunne tage de nye muligheder i brug.

### Hvad kan gå galt?

Umiddelbart er der ingen risici. Projektet kan begynde med det samme. Hvis der skal søges forskningsmidler, kan det være et problem at opnå tilstrækkelig forskningshøjde. Dette kan muligvis håndteres gennem forsøgsplanlægningen. Alternativt kan det nuværende system/plattform formentlig udvides gradvist/trinvist med nye mere og mere avancerede analyse og visualiseringsmetoder.

### Partnere

Gårdejer(e), DLG, Danske svine-dyrlæger og SEGES, Svin.

# BIG DATA FOR SVIN (III): OPTIMERING AF PRODUKTIVITET – GAINMAX 2

## VISIONEN

Et system, hvor data fra forskellige datakilder gennem hele værdikæden fra avl til produktion af slagtesvin indrapporteres i realtid, bearbejdes og præsenteres overskueligt, vil tillade forbedret styring af dyrenes sundhed og en højere produktivitet.

Resultatet bliver en bedre, hurtigere og mere effektiv styring af tilvækst og sundhed samt en øget økonomisk gevinst med baggrund i mindskede omkostninger pga. sundere dyr i besætninger.

Et fælles og koordineret datasystem vil føre til:

- En forbedret og løftet sporbarhed igennem hele kæden af produktionen
- En bedre besætningspleje
- Optimerede sygdomsforebyggende tiltag med hurtigere og mere målrettet sygdomsforebyggelse og behandling i form af vaccinationer, antibiotika mv.
- Økonomiske gevinster, reduceret dødlighed og forbedret dyrevelfærd
- Forhøjet overordnet standard og produktivitet af dansk svineproduktion

## Big Data for svin (III): Optimering af produktivitet - GainMax 2

Vi mangler i dag reeltids data, som kan bruges til at overvåge sundhed og produktivitet i svinebesætninger. Det begrænser mulighederne for at udnytte data til at styre dyrenes sundhed og svineavlens produktivitet optimalt.

Samtidig får vi færre og færre svineproducenter og større og større besætninger. For at fastholde konkurrenceevne og rentabilitet er landmanden nødt til at opretholde en så høj produktivitet som muligt. Derfor søges en optimal daglig tilvækst, hvilket forudsætter en god standard for dyresundhed. Som eksempel kan nævnes, at tilvæksten hos slagtesvin er steget fra 833 g dagligt i 2002 til 895 g dagligt i 2010. Samtidig bliver kravene på dyrevelfærds-siden gradvist skærpede.

For på en gang at kunne opnå øget tilvækst, nedsat foderforbrug og et forbedret sundheds- og velfærdsniveau, bruger landmanden i dag forskellige systemer. De fleste driftssystemer, der i dag benyttes i svinebesætninger, bruger uspecifikke målinger som vand-, foderforbrug og vægt. GainMax, udviklet af DLG, er et sådant eksempel. Ugentlige målinger af vægtforøgelse og foderforbrug giver landmanden et billede af svinebesætningens tilvækst i forhold til fodring (inklusive fodermængde, foderblandning, energiindhold mv.).

For større svinebesætninger er det obligatorisk at have en rådgivningsaftale med praktiserende dyrlæge. Som minimum skal dyrlægen besøge besætningen en gang hvert år. Dyrlægen skal rapportere fund af visse sygdomme og medvirke til bekæmpelse af smitsomme sygdomme. For at sikre et højt sundhedsniveau har branchen udarbejdet særlige retningslinjer. I Specifik Patogen Fri (SPF)-besætninger stilles der høje krav om ekstern smittebeskyttelse og overvågning af specifikke sygdomme. Omkring 73 % af alle søer og 38 % af alle slagtesvin har i dag SPF-status. Dyrlæger, apoteker og fodermøller registrerer blandt andet forbrug af lægemidler, sera og vacciner i besætninger, som indrapporteres i databasen VETSTAT (Veterinary Medicine Statistics). Disse oplysninger registreres sammen med information om dyreart, aldersgruppe, besætningsidentitet (CHR-nummer) og diagnose. På baggrund af de indsamlede data overvåges forbruget af lægemidler i de forskellige besætninger. VETSTAT data er imidlertid historiske og kan ikke umiddelbart anvendes i et system baseret på reeltids data.

Der er planer om systematisk sundhedsovervågning baseret på effektive målinger af infektionsprofiler for svin på sti- eller sektionniveau. Laboratorieanalyser vedrørende sygdomme er i en rivende udvikling, og anvendelse af nye metoder og materialer vil i løbet af få år gøre det muligt at implementere løbende undersøgelser så tæt på realtid, at det vil kunne bidrage til 1) Tidligere erkendelse af

infektiøse sygdomme 2) Vurdering af hvilke sygdomme, der er mest betydende og 3) Løbende tilpasning af indsats vedr. forebyggelse og behandling. Ved samkøring med data vedrørende fodring og tilvækst vil sådanne laboratoriedata bidrage til at vurdere, hvordan virusinfektioner (PRRS, PCV2, influenza), luftvejsbakterier (Mykoplasma hyopneumoniae, Actinobacillus pleuropneumoniae) og diarreabakterier (E. coli, Lawsonia intracellularis, Brachyspira spp.) indgår i et dynamisk samspil. Dermed kan forebyggelse og behandling optimeres.

### Sådan skaber indsatsen værdi

Det er vigtigt at bibeholde Danmarks ry for at levere svinekød af høj kvalitet og sikkerhed, der er produceret med en høj grad af dyrevelfærd. Derfor vil et beslutningsstøttesystem, der kan støtte op om at bevare højest mulige standarder for vores svineproduktion, have stor værdi. Ved at koble sundhedsdata og produktionsdata på sti- eller staldniveau vil produktiviteten kunne forbedres og sundheden overvåges, så der hurtigere kan reageres på problemer og sygdomme. Det vil kunne bidrage til et mere målrettet og omkostningseffektivt forbrug af foder samt valg af vaccinationer og antibiotikabehandlinger, herunder potentielt en reduktion af sidstnævnte. En stærkere og mere detaljeret kobling mellem fodring, tilvækst og sundhed vil give landmanden et mere robust værktøj til at styre sin besætning.

Der er stor enighed i branchen om, at dødeligheden i svinebesætninger bør sænkes. Ambitionen er at mindske den gennemsnitlige dødelighed med 20 pct. i 2020. Den økonomiske værdi af at sænke dødeligheden vurderes at være ca. 6 kr. pr. gris pr. procentpoint, som dødeligheden reduceres i slagtesvinebesætninger (ca. 30 kg til slag), og 3-4 kr. pr. gris.

Der ligger et stort potentiale i at koble forskellige data sammen for at få et bedre overblik, bedre muligheder for sundhedsstyring og optimering af tilvækst. Rent praktisk vil det give god mening at bruge det nuværende GainMax-system som forløber og modelsystem, og udvide dette med data for medicinforbrug, laboratoriedata, slagtedata, flyttedata og data fra destruktionsanstalt. Et system med data tæt på realtid vil gøre det muligt at handle hurtigt på eventuelle mistanker eller afvigelser.

### Big Data værktøjer

Den beskrevne løsning er en langsigtet vision, der vil kræve forpligtelse, engagement og en vedvarende indsats. Et fuldt integreret datasystem har mange fordele, herunder øget sporbarhed, øget forståelse af betydningen af sundhedsstatus for dødelighed og daglig tilvækst samt produktionsøkonomi.

Et forbedret overblik og en forbedret strategi på besætningsniveau vil kunne optimere de velfærdsmæssige og produktionsøkonomiske målsætninger. Det kan ske gennem hurtigere og målrettede sygdomsforebyggende tiltag, bedre produktionsresultater i den enkelte besætning, og med en positiv indflydelse på grisenes velfærd. Samtidig kan synkronisering af eksisterende og fremtidige data fra forskellige datakilder og forskellige aktører forbedre produktionen, resultere i lavere omkostninger og bedre dyrevelfærd, mindre sygdom og medicinforbrug.

### **Teknologiske og organisatoriske udfordringer og privacy/etik og sikkerhed**

De teknologiske udfordringer omfatter kompetencer indenfor datafangst, omkostninger til etablering/vedligeholdelse af hardware på bedrifterne, dataanalyse, synkronisering af forskellige datakilder samt præsentation af resultater på en praktisk brugbar måde.

Systemet vil kun anvendes af landmænd, der deltager frivilligt. Desuden må det sikres, at den enkelte landmand/besætning ikke "hænges ud", men derimod opnår fordelene ved, at hans data indgår i et større hele. Deling af data mellem besætninger er ikke relevant, men der kan laves centrale sammenligninger. Det vil være muligt at holde besætninger fri fra dette.

Ejerforhold, sikkerhed og brugerfrihed på data skal sikres, og overensstemmelse mellem de berørte parter skal opnås. Nogle af de nødvendige datakilder eksisterer allerede i dag, men det vil være en udfordring at koble dem sammen på en brugbar og værdifuld måde. Udvikling af et koncept til generering af sundhedsdata i realtid vil være af stor betydning for at optimere sundhedsstyringen i besætningerne. Sidst men ikke mindst bør validering og kvalitetssikring af data indgå i projektet.

### **Hvad kan gå galt?**

For at opnå optimal anvendelse af de koordinerede og synkroniserede data, skal dataene være rapporteret i realtid. De rigtige og relevante partnere er nødt til at identificere og forpligte sig til projektet, ligesom de rigtige og kritisk relevante datakilder er nødt til at indgå for at få maksimal udbytte og for at kunne realisere visionen.

### **Partnere**

Agrosoft (eneste udbyder af software til produktionsstyring i dag), praktiserende dyrlæger, Dianova og DLG.





# BIG DATA FOR SVIN (IV): FORBRUGERNES PRÆFERENCER VISIONEN

Hidtil har vurderinger af potentialet for vækst i dansk svineproduktion hovedsageligt været baseret på gennemsnitsbetragtninger. Med andre ord har målet været en hurtig og ensartet gennemsnitlig tilvækst for et givet foderbudget.

Der ligger imidlertid også et vækstpotentiale i at udnytte den naturlige varians i produktionen. Forbrugere i forskellige lande ønsker ikke nødvendigvis det samme standardprodukt, men ønsker måske særlige vægtklasser af dyr, særlige fedt/kød fordelinger, særlige smage osv.

## Big Data for svin (IV): Forbrugernes præferencer

Mange ressourcer i dansk svineproduktion er hidtil brugt på at udvikle standardiserede grise på de forskellige bedrifter. Grise leveres til slagtning med en bestemt vægt indenfor snævre tolerancer. Ligeledes sigtes efter et bestemt forhold mellem kød og fedt. Herved forenkles den videre proces med transport, slagtning, partering og fremstilling af slutprodukt.

Imidlertid har de forskellige markeder og grupper af kunder hver deres præferencer, som desuden forandrer sig over tid. Tidlig viden om udviklingen i præferencer kan derfor udnyttes til at målrette svineproduktionen.

Danske svineproducenter har god erfaring med at omstille produktionen hurtigt. For eksempel blev kød/fedt procenten for nogle år siden ændret mod en anden target-værdi. Dette blev opnået efter ganske få hold svin var produceret.

Traditionelt er forbrugerpræferencer blevet vurderet på basis af spørgeskemaundersøgelser og smagspanelbedømmelser. Disse instrumenter er omkostningstunge og besværlige. Derudover er det ikke sikkert, at den stikprøve, man rammer, er repræsentativ. Spørgsmålet er derfor: Er det muligt at basere sig på internetdata og data fra sociale medier, så man opnår en mere dækkende vurdering af forbrugernes nutidige præferencer og den fremtidige udvikling? Hvordan kan vi ved hjælp af tilgængelige internetdata tidligst muligt identificere og forudsige nye trends af relevans for afsætningen af svinekød? Konkret vil der skulle svares på følgende spørgsmål:

- Er det en voksende eller aftagende trend? Hvor langt er et produkt eksempelvis nået på en såkaldt adoptionskurve, hvor man først rammer forbrugere, der altid er med på de nyeste trends, senere får hovedparten af forbrugerne med, og til sidst efterfølgerne.
- Hvordan fordeler trenden sig i forhold til alder, køn, geografi, socialklasse mv.?
- Hvilken effekt har tendensen på afsætningen af svinekød nu og i fremtiden?

Herudover er det vigtigt at få svar på: Hvordan kan vi ved hjælp af tilgængelige internetdata identificere forretningsmuligheder med udgangspunkt i de nye trends med hensyn til:

- Nye produkter
- Nye segmenter
- Nye afsætnings- og kommunikationskanaler

Eksempler på aktuelle trends, som kunne være relevante at kunne monitorere og agere på, er:

- Opfattelsen af kvalitet, sundhed, prestige
- Sundhedsbølgen - fitness, protein som kosttilskud og sundhedsfremmende fødevarer
- Tendensen med færre hovedmåltider og flere snacks - mere "on-to-go", sund "fast food", flere single-husholdninger

Desuden er det værdifuldt at identificere parametre for kvalitet på de forskellige markeder.

Følgende cases er umiddelbart relevante, men udelukker ikke andre i fremtiden:

- Slagtekroppe med ornelugt. Hvis disse kan identificeres tidligt i processen, kan slagtekroppen udnyttes i produkter, hvor ornelugten ikke har betydning
- Markedets holdninger til dyrevelfærd. Hvordan er sammenhængen mellem ønsket om et bestemt niveau af dyrevelfærd og pris-følsomheden?
- Hvad er markedets holdninger til sundhedsegenskaberne af kødprodukter?
- Hvad er markedets holdninger til sporbarhed og fødevarer sikkerhed?
- Hvad er markedets holdninger til økologi?

### Sådan skaber indsatsen værdi

Identifikation af markeder og segmenter vil give et grundlag for øget differentiering af produkterne og dermed øget værditilvækst af råvaregrundlaget. Som følge heraf bliver det muligt at opnå højere afsætningspris.

Herudover vil forbrugerpræferencer udledt fra tilgængelige internetdata kunne bruges som input og supplement til forbrugerundersøgelser på basis af fokusgrupper og produkttest.

### Big Data værktøjer

Der skal skabes adgang til data fra internettet og sociale medier fx omkring prævalensen af positive/negative ytringer om svinekød og svineproduktion. Disse skal korreleres til markedssegmenter, herunder geografi, samt produktkategorier og faktorer som oprindelse, smag, mørhed og fedtindhold.

Man kan også forestille sig, at data skabes aktivt gennem brug af ad Words eller aktiv udnyttelse af sociale medier og debatfora, fx gennem brug af "battles" og provokative

udsagn, som så monitoreres for de reaktioner, der kommer fra brugerne.

Yderligere kan man forestille sig at identificere og udnytte tendenser i markedssegmenter (geografiske med mere), som traditionelt er ”foran” de danske. Endelig vil man kunne udnytte afsætningsdata fra slagteri og detailed (fx AC Nielsen) samt forbrugerpaneler (fx GFK) evt. i kombination med data fra internettet til at lave nye analyser og vurderinger af forbrugerpræferencer.

### **Kompetencebehov**

Det vurderes, at de nødvendige kompetencer allerede er til stede, eller kan udvikles baseret på nuværende kompetencer parterne i mellem.

### **Privacy/etik og sikkerhed**

Dette punkt er meget vigtigt. Ud over den privacy/etik og sikkerhed, der er nødvendig overfor branchen er der nu kommet en yderligere dimension, nemlig forbrugersiden. Det er helt afgørende at få forholdet til internetdata og forbrugeroplysninger afklaret. I modsat fald risikerer man negative ”virale” forhold, som kan medføre ubodelig skade for branchen.

### **Teknologiske og organisatoriske udfordringer**

Feltet spænder hele vejen fra jord til bord og involverer derfor hele spektret af interessenter. Det vil være

nødvendigt med robuste måder at indsamle data om forbrugernes præferencer og koble dem til viden og data om produktion af svin. Produktionsdata vil omfatte sammensætning og mængde af foder, tidspunkter for ændring af fodersammensætningen samt nøje monitorering af dyrenes udvikling. Eventuelt kan det komme på tale at udvikle sensorer, der kan indsamle indikatorer for den endelige smag og tekstur af produkterne. I sagens natur er der organisatoriske udfordringer, da et stort antal af leddene i værdikæden er omfattet.

Teknologisk er opgaven også stor. Der skal udvikles sensorer, større datamængder end sædvanligt og nye typer af data skal håndteres, og endelig er ”feedback-sløjferne” med den høstede information mere mangfoldig end i andre anvendelser af Big Data på fødevareområdet. Til gengæld er potentialet enormt.

### **Hvad kan gå galt?**

De mulige risici er nævnt under ”Privacy/etik og sikkerhed” samt ”Teknologiske og organisatoriske udfordringer”.

### **Partnere**

Danish Crown/Tican, Danish Meat Research Institute. Svineproducenter, SEGES, svin.

## Udfordringer og muligheder for fjerkræ

Fjerkræproduktionen er et eksempel på succesfuld anvendelse af systematisk indsamlede store mængder af data. Indsatsen udspringer af den såkaldte Salmonella-handlingsplan. Fra at branchen næsten måtte opgive produktionen i Danmark, er den nu kommet på fode igen, er blevet rentabel og har fået et bedre omdømme.

SEGES administrerer en database for Kvalitetssikring i Kyllingeproduktionen (KIK). Basen indeholder unikke historiske data for hver kyllingebatch. Disse data vil helt naturligt kunne anvendes til Big Data projekter.



# STYRING AF PRODUKTIVITET I DEN ENKELTE KYLLINGEFLOK VISIONEN

Målet med projektet er gennem en kortlægning og analyse af forsyningskæden fra kyllingefarm til køledisk at opnå følgende:

- Øget sporbarhed og dokumentation af dyrevelfærd
- Etablering af mulighed for realtidsstyring af produktiviteten af slagtekyllingeflokke
- Bedre sporbarhed af faktorer som kan påvirke fødevareresikkerhed
- Bedre sporbarhed af hygiejniske tiltag i en specifik besætning
- Bedre forståelse for hvorledes de enkelte data påvirker hinanden. Dataværdien for alle led tydeliggøres.  
Dvs., hvilke data kan og skal deles, og hvad er risici og fordele ved at gøre dette
- Muligheder for en optimeret markedsstrategi og forbedring af branchens renommé

## Styring af produktivitet i den enkelte kyllingeflok

Produktionen af kyllinger starter med daggamle kyllinger fra rugerierne og ender ca. 40 dage senere, hvor kyllingerne afhentes til slagtning. Produktionen af rugeæg til slagtekyllinger sker i særlige stalde, hvor der går både høner og haner. Hver hane "betjener" cirka 10 høner. Hønerne begynder at lægge æg, når de er ca. 22 uger gamle, og bliver ved til de er ca. 62 uger gamle. Hver høne lægger 150-160 æg i den 40 uger lange æglægningsperiode.

Forældredyrene bliver vaccineret mod forskellige alvorlige fjerkræsygdomme. Det mindsker risikoen for sygdomme, og desuden videregiver hønerne immunstoffer til deres afkom.

I 2013 blev der herhjemme slagtet 106 mio. kyllinger. Samme år blev der solgt 50,55 mio. kg æg på hjemmemarkedet, fordelt på 26,82 mio. kg buræg, 10,65 mio. kg skrabeæg, 2,73 mio. kg æg fra fritgående høns og 10,35 mio. kg økologiske æg.

Dansk Slagtefjerkræ har arbejdet dygtigt og systematisk med fødevarerikkerhed, sporbarhed og kvalitet. Konceptet Kvalitetssikring i Kyllingeproduktionen (KIK) har samlet branchen om et fælles mål: At skabe verdens bedste og mest veldokumenterede kyllinger.

Branchen er karakteriseret ved en meget snæver fortjenstmarginal. Derfor kan sygdomsudbrud, akut opståede velfærdsproblemer eller et parti foder med en utilsigtet uheldig sammensætning få store økonomiske konsekvenser for dels besætningsejeren, dels for slagteriet.

KIK databasen indeholder i forvejen en lang række produktivitetsmål, foruden data der giver høj grad af sporbarhed. Ofte findes oplysningerne imidlertid ikke i realtid, idet de i mange tilfælde først indtastes, når en flok er færdigproduceret og sendt til slagtning. Dette begrænser selskabet muligheden for aktiv styring af produktiviteten i den enkelte flok.

### Sådan skaber indsatsen værdi

Selvom den vellykkede indsats mod Salmonella har styrket forbrugernes tillid, er slagtekyllingeblandens image stadig ikke helt i top. Dette er i nogen grad ufortjent og baseret på fortidens synder. Men det gør det vanskeligt at få maksimalt markedsrettet udbytte af forhold, som at fjerkrækød er meget ernæringsrigtig, og at produktionen har et lavt CO<sub>2</sub>-aftryk. Et bredere kendskab til og bedre brug af data i forsyningskæden vil kunne bidrage til at skabe tryghed hos den enkelte forbruger.

Værdien af dette projekt er, at en samlet og tværgående analyse af data i realtid kan gøre det muligt at overvåge og styre produktiviteten i den enkelte kyllingeflok, fx gennem integration af data hos foderstofleverandøren. Med andre ord kan denne tilgang forbedre hele forsyningskæden i stedet for de individuelle led på en sådan måde, at alle led (f.eks. avlere, slagterier m.m.) får øget udbytte. Desuden vil det sikre større effektivitet, mindre spild og hurtig handling overfor sygdom eller dårlig trivsel hos dyr. Dette vil gavne ikke kun landmændene, rugerier, slagterier og transportfirmaer, men også detailhandlen og i sidste ende forbrugerne samt samfundet generelt i form af en øget fokus på effektivitet og velfærd gennem forsyningskæden (f.eks. brug af færre ressourcer).

Værdiskabelsen ligger dels i forbedret effektivitet i alle led af forsyningskæden, dels i øget omsætning i mange led i kæden; økonomisk besparelse på virksomhedernes udgifter; forbedret grundlag for brug af sygdomsanalyser pga. muligheden for at finde sammenhænge på tværs af forsyningskæden og øget kvalitetsstyring og kontrolmuligheder. Der er således gevinster for alle parter i kæden - avlere, slagterier og forbrugere.

### Big Data værktøjer

Modeller og analyse af ovenstående data vil forbedre produktionen og formentlig også resultere i bedre dyrevelfærd og mindre sygdom. Ved at kende den totale forsyningskæde og de data, der er forbundet dermed, kan der lettere udvikles nye forretningsstrategier. Man vil fx bedre kunne sammenholde forbrugernes ønsker og krav med opdrætternes muligheder. Derved kan der opstilles en bedre markedsstrategi og muligheder for udvikling af nye markeder og produkter.

Teknologiske og organisatoriske udfordringer og privacy/etik og sikkerhed

De enkelte databaser og registre skal sammenkøres, så de passer med KIK-databasen, da den er en central enhed.

### Hvad kan gå galt?

Kyllingeproduktionen er langt i forløbet, når det drejer sig om at udnytte Big Data. Da branchen tidligere har været stærkt presset i forbindelse med Salmonella-handlingsplanen, synes der også at være en rimelig afklaring omkring problematikken med privacy. Det skal dog ikke underkendes, at der er et stykke vej, før alle ønskede data og flow af data er på plads.



## Salg af fødevaredata

Når en virksomhed eller branche sætter gang i Big Data projekter, viser der sig ofte muligheder, der rækker ud over det, man normalt beskæftiger sig med. En virksomhed i fødevarerbranchen genererer måske data med det formål at optimere sin egen drift, men disse data kan udmærket have værdi for virksomheder i andre brancher. Vel at mærke i så høj grad, at det kan være penge værd. Det er der efterhånden en del eksempler på.

Et af de mest kendte eksempler på dette er den danske vindmølleproducent Vestas, som ved hjælp af sensorer i sine møller producerer data på vindprofiler. Vindprofilerne har værdi for vindprojektudviklere og andre vindmølleproducenter i kraft af, at de sandsynliggør om vindmøller stillet op på en lokalitet er lønsomme eller ej. Vestas sælger derfor disse data til vindprojektudviklere og andre vindmølleproducenter.

Også i fødevarerhvervet er der mulighed for at sælge egne data med fortjeneste. Her skal nævnes blot to eksempler:

- Værdiskabende dataanalyse og leverandører: Visionen er at sælge adgang til data koblet med biologiske (mm.) sammenhænge i landbruget og sælge analyser baseret på disse data

Konkret kunne SEGES sælge markinformation til Danish Agro. SEGES ville tjene penge på salg af data og tilhø-

rende konsulenttimer og faglig ekspertise forbundet med at rådgive Danish Agro om anvendelse af dataene.

- Udvikling af beslutningsstøtteværktøjer til den finansielle sektor og fødevarerindustrien: Visionen er at tilbyde landbrugsrelaterede virksomheder værktøjer til beslutningsstøtte

SEGES har allerede mulighed for at udvikle beslutningsstøtteværktøjer til bl.a. forsikringsbranchen.

Et eksempel er kvægdatabasen, der giver mulighed for at forudsige den enkelte mælkeproducents risiko for at levere mælk med rester af antibiotika. På baggrund af denne viden kan forsikringsbranchen udarbejde differentierede forsikringspræmier baseret på forskellige risikoprofiler. Det betyder, at mælkeproducenter med en lav risiko kan få lavere forsikringspræmie end mælkeproducenter, hvor der på baggrund af data i kvægdatabasen kan forudses en langt højere risiko.

Tilsvarende kunne der være mulighed for at udarbejde værktøjer til beslutningsstøtte for forsikringselskaber, der ønsker at udbyde forsikringer for kornpriser, eller beslutningsstøtteværktøjer baseret på økonomidatabasen i forbindelse med realkreditvurderinger.



## Forbrugersundhed som eksempel på fremtidens muligheder

Det foregående kapitel præsenterede en række projekter, hvor forudsætningerne for at opnå vækst og højere kvalitet i en række delområder inden for fødevarerektoren allerede er til stede. Samtidig er det en kendsgerning, at feltet Big Data udvikler sig med rivende hast. Derfor er der stor sandsynlighed for, at der i fremtiden vil åbne sig flere muligheder. Nogle af dem vil formentlig have en helt anden karakter end de projekter, som allerede er skitseret.

Forbrugersundhed er et af de områder, som kunne åbne mulighed for radikalt nyskabende brug af Big Data. Et stigende antal danskere får konstateret kost- og livstilsrelaterede sygdomme som diabetes og hjertekarsygdomme. Hertil kommer, at ca. 30 % af alle kræfttilfælde har sammenhæng med kosten. Dette fører til et betydeligt antal dødsfald årligt, hvilket medvirker til, at gennemsnitsalderen i Danmark er lavere end i de lande, vi normalt sammenligner os med. Dette er en betydelig udgift for samfundet både i form af direkte sundhedsudgifter og tab af "gode leveår".

Den store udfordring er, at spørgsmålet om sund kost og sund levevis ikke har et simpelt svar. Vi kender ikke alle faktorer og deres samspil, og vi ved ikke, hvad der er sund kost for den enkelte. Samtidig er området præget af mange holdninger og "skoler", der på mere eller mindre oplyst grundlag hævder at have løsningen.

Desuden er det ikke kun sammensætningen af vores kost, og hvor meget vi spiser af den, der har betydning. Også den måde vi fremstiller og behandler fødevarerne spiller ind. Det gælder såvel i landbrugets produktion og i fødevarerindustrien som i detailhandlen og i hjemmet. Dertil kommer, at man i hele kæden anvender et stort antal stoffer, mens andre stoffer overføres til fødevarer fra emballage og andre materialer, der kommer i kontakt med fødevarerne eller på anden måde finder vej ind i fødevarerne som forureninger.

Derfor er hele fødevarekæden fra landbrugsproduktionen via fødevarerindustrien og detailhandlen til bordet hos den enkelte borger en del af udfordringen. Desuden er sundhedsvæsnets og andre dele af samfundet vigtige interessenter.

### Sådan skaber indsatsen værdi

Forbrugersundhed er et område af stor samfundsbetydning, hvor Big Data kan skabe værdi. Herunder er nævnt eksempler på mulige områder:

#### *Bedre oplysning*

En dataplatform til bedre oplysning om kost og sundhed kan understøtte personlig kostplanlægning. Vi lever i en tid, hvor krops- og sundhedsbevidsthed fylder stadig mere. Man taler om "Quantified self" i form af sensorer, der overvåger vores sundhed på forskellig måde. Desuden



må personlig kostplanlægning baseret på kortlægning af den enkeltes komplette arvemasse – genom – forventes at få efterspørgslen på data og modeller til personlig vejledning til at stige.

Dataplatformen vil skabe baggrund for nye former for Big Data inden for oplysning om kost og sundhed baseret på solide data. Det kan være helt nye forretninger eller nye forretningsområder i eksisterende forretninger. Eksempelvis vil der være store sundhedspotentialer for den enkelte forbruger i at få at vide, ”Hvor sund er din indkøbskurv?” baseret på bon-data og upload af relevante personlige oplysninger.

### *Forståelse af sammenhæng mellem kost og sygdomme*

Diabetes-2 er i voldsom vækst i store dele af den industrialiserede verden. Det er bemærkelsesværdigt, at mens mange andre sygdomme går tilbage i takt med stigende velstand i et samfund, ser forekomsten af diabetes-2 ud til netop at vokse, når velstanden bliver større. Men hvilke mekanismer, der fører til diabetes-2, er stadig ikke klart.

Der er dog beskrevet et meget stort antal indikatorer i litteraturen: Kostens sammensætning, det samlede energiindtag, fysisk aktivitet, et stort antal forskellige kemiske stoffer i kost og fra miljø samt betydning af tarmfloraen.

Det samme gælder for de ca. 30 % af alle kræfttilfælde, der tilskrives kosten.

Udfordringen er, at disse effekter afhænger af et stort antal faktorer, lige som det spiller ind i hvilke kombinationer faktorerne forekommer, og over hvor lang tid man udsættes.

Endelig er der meget stor forskel på den individuelle følsomhed. Et eksempel på et stof er acrylamid, der mistænkes for at fremkalde brystkræft ved indtag igennem længere tid, eller fluorerede stoffer hvor data tyder på, at de undertrykker immunforsvaret. Området er genstand for en meget stor forskningsindsats, men ofte fokuseres på relativt få faktorer ad gangen.

Et fremtidigt scenarie kunne være en model baseret på kunstig intelligens, som kan skabe forståelse af sammenhænge mellem kost og sundhedseffekter. Modellen skal kunne trække data fra både den videnskabelige litteratur og alle tilgængelige danske registre, herunder for genetiske data. Modellen kan udvides med forståelse af kemiske stoffers betydning for væsentlige sundhedsudfordringer.

### *Befolkningens kostvaner*

Hvad er omkostningen for samfundet af befolkningens kostvaner? Et estimat af de omkostninger, der er for



samfundet, ved at befolkningen lever et usundt liv kan bruges som grundlag for at prioritere ressourcerne. Hvis det kan prissættes, hvad det koster samfundet, at en del af befolkningen har usunde kostvaner, vil det være muligt at prioritere samfundets økonomiske ressourcer til forebyggelse. Dette kan være igennem lettelse af afgifter, bedre prissætning eller støtte til sundhedsfremmende kost og bedre processer.

#### **Big Data værktøjer og kompetencebehov**

Der er behov for en flerstrengt strategi, som forudsætter kompetencer i at trække data ud af videnskabelig litteratur – såkaldt data-mining. Det kan for eksempel være oplysninger om sammenhænge mellem diabetes-2 og kostfaktorer. Desuden kan der anvendes indsamling af helbredsoplysninger (diagnoser) og de associerede helbredsoplysninger, adfærdsundersøgelser osv.

Der skal også opbygges en åben dataplatform, hvor kostdata kan hentes. Det gælder både de nuværende i fødevarerdatabanken, men også andre relaterede data om alt fra forureninger til helbreds faktorer i et valideret og dokumenteret format. Sidst men ikke mindst er det nødvendigt at udvikle kompetencer i udvikling af modeller til håndtering af data på tværs af domæner, herunder specielt håndtering af metadata.

#### **Privacy/etik og sikkerhed**

Der vil være en række sikkerhedsudfordringer, men nogle af disse kan sandsynligvis løses ved ”noget for noget”. Med andre ord får brugeren nyttig vejledning til gengæld for at dele sine personlige data. Adgang til sundhedsoplysninger fra læger o.a. vil helt sikkert være en udfordring, men skal mest bruges i forbindelse med den detaljerede data-mining for at kortlægge faktorer. De fleste data om fødevarers indhold af stoffer er allerede offentlige, men ikke i databaseformat.

#### **Teknologiske og organisatoriske udfordringer**

Samlet kræver indsatsen indsigt på tværs af mange områder: Medicin og sundhed, landbrugspraksis, industrielle processer, detailhandel, økonomi, sociologi, psykologi og meget mere.

Der findes i dag meget omfattede data om kost, fødevarer og sundhed i mange forskellige domæner, men disse er ikke relateret og anvendes kun i begrænset omfang i sammenhæng. Der er eksempler på meta-studier, der prøver at samle al information til belysning af et område, men selv disse studier er ofte ”farvede”, fordi de er sat i gang med en bestemt tilgang.

Samtidig giver de mange forskellige domæner omfattende og meget komplekse udfordringer med at samle og behandle data på tværs, ligesom etiske forhold spiller ind.

#### **Hvad kan gå galt?**

Der er en risiko for, at man ikke vil lykkes med at afdekke sammenhænge mellem kost og sundhed, der meningsfuldt kan bruges som udgangspunkt for vejledning. En anden risiko knytter sig til, at der er meget holdning, grænsende til ”religion”, på området. Det er kritisk, at der sættes en standard for dataindsamlingen, så der bruges valide data og dermed gives seriøse svar fra de anvendte applikationer.

#### **Partnere**

Sundhedsvæsenet som helhed, diabetes foreninger, virksomheder indenfor monitorering, sundhedsindustrien, detailhandel (der ønsker at oplyse om sundhed) og App-udviklere.

# TVÆRGÅENDE FORSKNING I BIG DATA



# TVÆRGÅENDE FORSKNING I BIG DATA

Kunsten at gå fra data til løsninger kan kun realiseres i et samspil med overordnet forskning i Big Data.

Det er i krydsfeltet på tværs af de enkelte led i værdikæden, at de virkelig værdifulde indsigter og dermed potentialet for at skabe forretning i sektoren findes. Dette kapitel undersøger potentialet i den tværgående anvendelse af Big Data i fødevarersektoren – og hvad der skal til at for at realisere gevinsten.

Et af eksemplerne i det foregående kapitel var GainMax, der anvendes af svineproducenter. GainMax er et godt eksempel på et produkt, der viser potentialet ved udnyttelse af Big Data på besætningsniveau. Grundlaget for systemet er, at den enkelte landmand vejer grisene ugentligt. Herudover registreres bl.a. fodermængde og foderblanding. Landmanden kan via sin PC følge tilvæksten hos grisene, der bliver relateret til fodertype og fodermængde. De mange data giver landmanden et redskab til at fremme tilvæksten på stiniveau ved at kunne skrue op og ned for fodermængderne og ændre fodertyperne på præcist de rigtige tidspunkter.

Men, hvis det fulde potentiale i Big Data inden for svineproduktion skal udfoldes forudsætter det dog en bredere, tværgående anvendelse af de mange data, der bl.a. genereres i GainMax. Ideelt set skal foderproducenterne, udstyrsproducenterne (fx af foderautomater), landmænd, slagterier og detailledet dele og udveksle data om dyrenes foderbehov, vækstkurver og salg. Dette vil samlet set kunne bidrage væsentligt til udviklingen af nye foderblandinger, nyt effektivt udstyr og i sidste ende forbedret produktion, dyrevelfærd og omsætning.

De mange mulige projekter, som er præsenteret i rapporten, kan på hver deres måde fremme Big Data inden for kvæg- og svineproduktion, fjerkræbranchen og forbrugersundhed. Imidlertid er det vigtigt, at indsatsen bliver koordineret, og at den sker i et samspil med en mere overordnet, tværgående forskning i Big Data. Kun på den måde opnår man, at løsningerne kan tale sammen, samt at smarte løsninger på ét område kan spredes til andre områder.

Indsatsen for at sikre en tværgående anvendelse af data kan opdeles i fire trin.

1. **Forretningsudvikling.** Anvendelsen af tværgående Big Data skal give forretningsmæssig mening. Projekternes mål og ønskede effekt skal defineres inden man går i gang. Den ønskede effekt kan eksempelvis være øget produktivitet, øget dyresundhed eller lavere miljøbelastning.
2. **Udveksling af data.** Der skal udvikles en teknisk platform, hvor aktørerne kan udveksle data på tværs af værdikæden. Det kan være et system, der sikrer, at mælkeproducenter kan dele viden på tværs af produktionskæden – fra foderstofproducenter, producenter af malkeudstyr, landmænd, mejerier, forretninger til kunderne.
3. **Tværgående Big Data værktøjer.** Næste skridt er at udvælge de rette Big Data værktøjer. Det kan være etablering af sensorer i en svinebesætning, algoritmer til behandling af data eller metoder til visualisering af resultaterne. Ofte er nogle af værktøjerne allerede tilgængelige på markedet, mens andre skal udvikles specifikt til projektet.
4. **Datasikkerhed.** Arbejdet med rapporten har afklaret, at datasikkerhed og privacy er så vigtige temaer inden for fødevarerområdet, at de udgør et selvstændigt, centralt problemfelt. Big Data rummer ofte følsomme data (person-, dyr- eller virksomheds-oplysninger), og i takt med behovet for at dele data opstår et stigende behov for at etablere klare aftaleforhold om brug af data, datasikkerhed og dataetik.

## Trin 1: Forretningsudvikling

Når forskere fra DTU Veterinærinstituttet sætter et videokamera op i en stald og lader det køre i tre måneder, koster det bagefter et helt årsværk at analysere de mange data. Eksemplet illustrerer, at brugen af Big Data kræver omtanke. Vi skal ikke generere masser af data, blot fordi vi kan. Hver gang vi samler data ind, medfører det omkostninger. Det kan være til udstyr, til lagring af data, til energiforbrug og løn.



Men, som rapporten tydeligt viser, så kan de mange data have stor værdi – både for landmanden og for mejeristen. Viden om mælkekvalitet og kvantitet helt ned på det enkelte dyr giver producenterne helt nye muligheder. Landmanden kan tilpasse fordringen af den enkelte ko, så kvaliteten og kvantiteten øges. Fodringen af en ko kan f.eks. baseres på den analyserede mælkekvalitet kombineret med online målinger eksempelvis fra sensorer i malkerobotterne. Mejeristen har mulighed for at udvikle nye produkter af meget ensartet kvalitet. Mælken kan eksempelvis på sigt sorteres i separate tanke efter kvalitet til givne formål – fx til produktion af ost eller yoghurt.

Indsamlingen og brugen af data skal således give forretningsmæssig mening. Der skal være et fornuftigt forhold

mellem de gevinster, som branchen og virksomhederne får ud af det, og de ressourcer, man sætter ind.

### Trin 2: Udveksling og integration af data

Andet trin er udveksling og integration af data på tværs af værdikæden fra jord til bord. Fra produktion af planter og dyr, forarbejdning til forbruger. Sektorudviklingsprojektet har demonstreret, at det er i krydsfeltet af data på tværs af de enkelte led i værdikæden, at de betydningsfulde indsigter ligger. Og dermed potentialet for forretningsudvikling. Men, udvekslingen af forskellige former af data rummer en række tekniske udfordringer.

I fødevarersektoren er data værdikæden organiseret over en række aktører fra primær- og sekundær erhvervene samt fra detailsektoren og forbrugerne (se Figur 1).



Figur 1. Værdikæden for fødevarerproduktionen med involverede aktører og eksempler på Big Data projekter.

På denne baggrund kan en øget forskningsindsats have til formål at fremme videnskabelige og teknologiske midler til at styre, analysere, visualisere, og udtrække nyttige oplysninger fra store, forskelligartede, distribuerede og heterogene datasæt, således at man:

- Fremskynder udviklingen af videnskabelige opdagelser og innovation
- Fremmer udviklingen af nye dataanalytiske værktøjer og algoritmer
- Letter skalerbar, tilgængelig og bæredygtig datainfrastruktur
- Øger forståelsen for menneskelige og sociale processer og interaktioner ift. øget og bedre brug af Big Data

**Trin 3: Tværgående Big Data værktøjer**

Værktøjskassen til Big Data er stor. Den rummer hele spektret af teknologier og systemer til opsamling af data, behandling af data, lagring, analyse til visualisering og beslutningsstøtte. Tabel 1 viser de væsentligste værktøjer og teknologier.

Tabel 1. Oversigt over Big Data teknologier og værktøjer

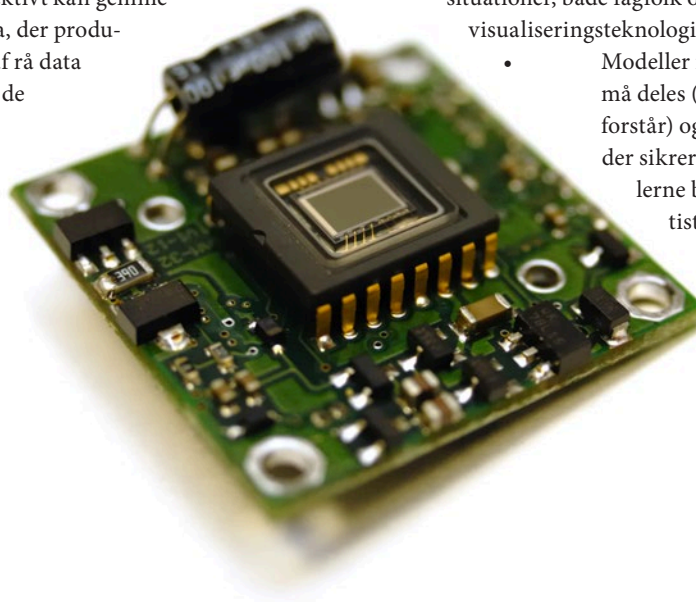
- a. Dataopsamlingsteknologi
  - Via sensorer
  - Eksisterende data / registre
- b. Forbehandling af data (tidligt stade)
  - Datareduktion
  - Kun gemme relevante data
- c. Data lagring, teknologier til intelligent lagring
- d. Forbehandling af data (sent stade)
  - Dataheterogenitet og integrationsudfordringer
- e. Dataanalyse
  - Automatisk strukturering
  - Dataanalytiske værktøjer og algoritmer
- f. Visualisering og beslutningstagen

I det følgende gennemgås nogle eksempler på vigtige forskningsudfordringer som er nødvendige at løfte, hvis udvekslingen og integrationen af Big Data skal lykkes.

**Opsamling og håndtering af data**

På en lang række områder indsamles data i stigende grad kontinuerligt. Den indsamlede datamængde stiger eksponentielt med tiden, og datalagring er derfor ofte distribueret. Behovet for automatisk opsamling (real-time og online, herunder i respekt for privacy samt offentlige krav og regler), udgør en stor udfordring. Udfordringen forstærkes yderligere af at analyse, filtrering, indeksering, komprimering og fortolkning af data derfor også skal foregå distribueret.

Den centrale teknologi til datalagring har udviklet sig hurtigt, og der er med tiden etableret store data warehouses, der effektivt kan gemme en stor del af de data, der produceres. Men lagring af rå data er blot en lille del af de større spørgsmål om at skabe og vedligeholde datalagre. Da datamængden vokser hurtigere end computerkraften er effektiv udnyttelse af den ligeledes voksende



computerkraft og nye maskiner (fx GPU'er) kun ét aspekt i at tackle disse udfordringer. Lige så vigtig er udviklingen af intelligente metoder til at analysere meget store og distribuerede datasæt.

**Dataanalyse og præsentation**

Der er behov for at udvikle stadig mere sofistikerede metoder til bl.a. data-mining og intelligent mønstergenkendelse (machine learning), der giver os mulighed for at udtrække flere nyttige indsigter fra mange datakilder. Fremtrædende nylige eksempler omfatter bl.a. søgemaskiner, automatiseret oversættelse, systemer for anbefalinger til kunder og afsløring af svindel med kreditkort. Vigtige forskningsområder omfatter:

- Systemer: Konstruktion af IT-systemer, der kan udføre komplekse behandlinger af data i meget store skalaer.
- Algoritmer: Udvikle mere sofistikerede machine learning teknikker samt teknikker til at analysere store datamængder.
- Programmering: Beregningsmodeller og sprog velegnet til at udtrykke dataanalyse-algoritmer i forhold til for eksempel store parallelle systemer (hvor data ligger spredt, og beregninger foretages flere steder parallelt) eller modelleringssprog til at formulere algoritmer for dataanalyse på et højt abstraktionsniveau uanset den fysiske fordeling af data.
- Cross-media information udvinding: Udvikle teknologi, der forstår tale, billeder, video og ustruktureret data; oversætte tale og tekst til andre sprog.
- Fusioneret information: Analyser, der kombinerer flere datakilder i flere forskellige former.

**Systemer og tekniske platforme**

Der er behov for at udvikle systemer og tekniske platforme, hvor interessenterne kan udveksle data på tværs af værdikæden. Nogle af udfordringerne er:

- Præsentation af data og interaktion med data og deres visualisering: Her er der behov for at udvikle teknologier, hvor komplekse data bliver forståelige og anvendelige for et bredt udsnit af personer i beslutningssituationer, både fagfolk og ikke-fagfolk, for eksempel visualiseringsteknologier og brugervenlige apps.
  - Modeller for at beskrive, hvordan data må deles (på en måde som brugerne forstår) og algoritmer og mekanismer, der sikrer (kontrollerer), at modellerne bliver fulgt. Nuværende statistisk baserede anonymiserede metoder giver ingen reelle garantier for databeskyttelse.



#### Trin 4: Datasikkerhed

Arbejdet med rapporten har afklaret, at datasikkerhed og privacy er så vigtige temaer inden for fødevarerområdet, at de udgør et selvstændigt, centralt problemfelt.

Data er en værdifuld ressource for de involverede interessenter. Derfor er der behov for særlige forholdsregler for at beskytte data (og tilhørende programmer) mod ulovlig indtrængen, herunder cyber-angreb. Det er væsentligt at etablere denne sikkerhed for at give interessenterne den tryghed, der gør, at de vil indgå i projekter om Big Data. De datasikkerhedsmæssige udfordringer beror dels på, at vi har at gøre med systemer med mange komponenter, som agerer uafhængigt af hinanden, og dels på at systemerne, takket være deres kommercielle og strategiske betydning, vil være oplagte mål for cyber-angreb.

Den vigtige pointe er, at i offentlighedens og privatpersoners øjne giver et ubegrænset kendskab til private data ubegrænset mulighed for at manipulere med individet. Omtrent det samme princip kan formodes at gælde i virksomheders og brancheorganisationers øjne. Her er det følsomme informationer om virksomhedens produktionsprocesser, produkter og forretningsmodeller, der i nogle tilfælde ikke ønskes offentlig opmærksomhed om. Konsekvensen er, at følelserne så at sige "siddet uden på tøjet". Samlet set skaber det en risiko for, at det fulde potentiale inden for Big Data ikke realiseres, fordi de negative konsekvenser ved misbrug af data er store. Både for privatpersoner og virksomheder.

Skal det fulde potentiale udnyttes, kræver det derfor klare aftaleforhold om produktion, adgang til, lagring og brug af data om mennesker og dyr og forretningskritiske virksomhedsdata ud fra et hensyn til respekten for den enkelte persons intimsfære og virksomhedernes arbejdsfelt.

Eksempler på relevante forskningsudfordringer inden for datasikkerhed er:

- Adgangskontrol. Formulering af politikker for, hvem der har adgang til hvilke data og håndhævelse af politikkerne.
- Informationsflow. Hvem kan påvirke data, og hvem kan udnytte data? Ofte tager man udgangspunkt i, hvem der ejer data, og i forholdet mellem de forskellige interessenter. Også her spiller formulering og håndhævelse af politikker en væsentlig rolle.
- Fortrolighed. Data, der er private for det enkelte individ, skal beskyttes samtidig med, at data om vedkommende kan indgå i statistiske beregninger og analyse.
- Troværdighed. Hvordan sikres, at data er troværdige? Emnet knytter sig dels til dataintegritet – altså sikkerhed for, at ingen har manipuleret med data – dels til, hvor data kommer fra, herunder kildens ry.
- Tilgængelighed. Hvordan sikres, at data altid er tilgængelige for de, der har retmæssig adgang til dem? Et centralt emne er, hvordan man sikrer sig mod cyber-angreb, der har til hensigt at blokere adgangen til data.



- Forretningsgange. Typisk vil data blive udvekslet mellem forskellige interessenter. Det er vigtigt at sikre sig, at data ikke kompromitteres undervejs. Her er de anvendte kommunikationsprotokoller, krypteringsnøgler og algoritmer centrale.
- Angrebsscenarier. Hvor er de svage led? Hvor vil det med andre ord være lettest at gennemføre et cyberangreb? Ud fra denne forståelse kan man studere forskellige forsvarsmekanismer og dermed styrke sit beredskab.

**Big Data og universiteter**

Der knytter sig en række særlige forhold til brug af Big Data i forskningsprojekter udført af universiteter.

Konkret betyder udfordringerne om datasikkerhed og privacy, at et universitet bedst vil kunne udnytte forskningspotentiallet i Big Data såfremt:

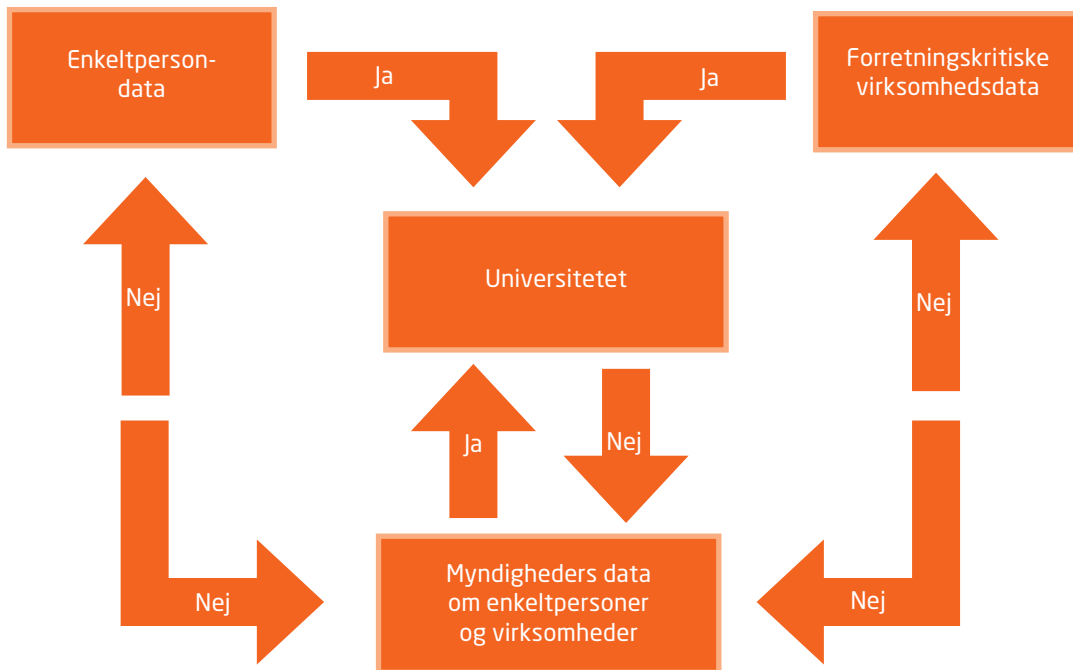
- Det tillidsbårne forhold mellem et universitet og de involverede parter om brug af data udbygges og konsolideres
- Der etableres klare aftaleforhold på et universitet om produktion, adgang til, lagring og brug af data (da klare aftaleforhold er en forudsætning for, at de tillidsbårne forhold kan fastholdes)
- Klare aftaleforhold omfatter både de rent tekniske aspekter af datasikkerhed og dataetik og de aftaleforhold, der skal være indgået og overholdes mellem de involverede parter.

Figuren nedenfor viser de nødvendige aftaleforhold mellem de involverede parter om deling og samproduktion af data på tværs af personer, virksomheder, myndigheder og universiteter. En pil med ”Ja” angiver, at der kan deles data mellem parterne. En pil med ”Nej” angiver det modsatte.

Aftalerne om, hvorvidt data deles eller ej, er det ene aspekt af klare aftaleforhold. Det andet er, om der er tale om kilde-identificerbare eller ikke-kilde identificerbare data.

Det enkelte universitet kan overveje at sikre sig, at alle data, der kommunikeres om til eksterne parter efter forskningen på universitetet er gennemført, er ikke-kilde identificerbare. Derved kan universitetet garantere eksterne parter, at disse data er anonymiserede. I figuren er dette princip illustreret ved boksen omkring ”Universitet”. Alle data inden for boksen er kilde identificerbare, hvorved kilden er kendt. Alle data uden for den røde linje er ikke-kilde identificerbare, hvorved data er anonymiserede.

Pointen er, at såfremt aftaleforholdene overholdes og de rent tekniske aspekter af datasikkerhed og dataetik om kilde-identificerbare og ikke-kilde identificerbare data tages hånd om af alle parterne, kan universitetet bedrive forskning under princippet om klare aftaleforhold.



Klare aftaleforhold: Organisering af deling af data i forbindelse med forskningsprojekter.

# KONKLUSIONER OG ANBEFALINGER



# Et nyt grundlag for anvendelsen af Big Data i fødevarerektoren

Potentialet er stort, men ingen aktør inden for dansk landbrug og fødevarereindustri kan løfte opgaven alene.

Den danske fødevarerebranche er verdenskendt for fødevarer af høj kvalitet og med høj fødevareresikkerhed, der kan sælges til en konkurrencedygtig pris. Desuden har Danmark ry for bæredygtig og ressourceeffektiv fødevarereproduktion, god dyrevelfærd og velfungerende logistik.

Fremtidens udfordringer for erhvervet er imidlertid store. Det kræver en målrettet indsats at fastholde produktionen og antallet af arbejdspladser til gavn for det danske velfærdssamfund.

En af udfordringerne er at producere mere med mindre. Det vil sige, at en øget fødevarereproduktion skal gå hånd i hånd med et lavere ressourceforbrug og en mindre miljøbelastning.

Samtidig skal konkurrenceevnen styrkes af en mere effektiv fødevarereproduktion gennem hele værdikæden. Herunder skal omkostningsniveauet mindskes gennem fortsat procesoptimering.

Endelig er der behov for at udvikle nye innovative fødevarer samt nye intelligente og effektive måder til at øge fødevarerekvaliteten og fødevareresikkerheden.

Alt dette forudsætter både teknologisk og organisatorisk innovation. Big Data er en væsentlig del af løsningen.

## Unik chance for Danmark

Danmark har unikke betingelser for at lykkes med Big Data i fødevarereerhvervet. Det skyldes, at fødevarereerhvervet og dets brancheorganisationer allerede har erfaring med forretningsudvikling baseret på store datamængder, og at kvaliteten af data på flere former for fødevarer er meget høj i forhold til i konkurrerende lande.

Der er milliarder at hente, hvis det danske fødevarereerhverv bliver bedre til at bruge disse data. En række teknologiske udfordringer skal løftes, såfremt fødevarereerhvervet skal lykkes med at realisere denne gevinst:

- Der skal skabes integration på tværs af Big Data værktøjerne lige fra sensorer, der opsamler data, til værktøjer, der skaber indsigter og kan visualisere indsigterne til brug for beslutningsstøtte – hele vejen fra jord til bord
- De enkelte Big Data værktøjer skal integreres med de specifikke forhold, der knytter sig til hver form for dyr, planter og foderstof, og inden for hvert af disse domæneområder skal der ske integration af data på tværs af værdikæden for produktionen. Dvs. fra produktion af planter/foder, dyr, forarbejdning til forbruger
- Der skal ske en reorganisering af arbejdet med data internt i den enkelte virksomhed og mellem den enkelte virksomhed, eksterne private virksomheder og offentlige samarbejdspartnere
- Deling og produktion af data med eksterne private og offentlige samarbejdspartnere skal ske under klare aftaleforhold - dvs. i tekniske og aftalebårne samarbejdsrelationer baseret på tillid og funderet i gensidig accept af parternes hensyn om privacy og etik om data
- IT- og management-sektorerne skal arbejde samordnet, så fødevarereerhvervet kan omsætte Big Data hardware og software i nye og bedre arbejdsprocesser, der skaber værdi.





## Anbefaling 1:

# Høst de lavt hængende frugter

Rapportens hovedanbefaling er at sætte projekter i gang inden for de 11 problemfelter, som er identificeret i sektorudviklingsprojektet i samarbejde mellem DTU, virksomheder, brancheorganisationer og myndigheder. Skal dette ske, er det nødvendigt, at både universiteter, virksomheder, brancheforeninger, myndigheder samt det offentlige forskningsbevillingssystem i fællesskab understøtter viden- og teknologiudviklingen inden for Big Data på fødevarerområdet.

Konkret handler det om projekter, der adresserer problemfelter inden for produktionen af svin, kvæg, kyllinger og mælk på tværs af værdikæden. Projekterne vil inddrage besætningsniveauet, forarbejdningsleddet, producenter af udstyr, detalilledet og relevante myndigheder, herunder Fødevarestyrelsen.

Eksempler på problemfelter er:

- Besætningsniveau. DTU, DI og L&F anbefaler, at der igangsættes aktiviteter, der adresserer de udfordringer og store muligheder, der er for anvendelse af Big Data på besætningsniveau, både blandt svine-, kvæg-, mælk og fjerkræproducenter. Det gælder bl.a. realtidsmonitorering af forbruget af foder og medicin. Initiativet kan ske i et partnerskab mellem besætningsejere, DLG, SEGES, DI ITEK, Fødevarestyrelsen og DTU inden for rammerne af et demonstrationsprogram, fx Grønt Udviklings- og Demonstrationsprogram (GUDP).
- Forarbejdningsleddet (slagterier og mejerier): Opsamling og bearbejdning af sundhedsdata og produktionsdata på sti- og staldniveau, fremme af nye skanningsmetoder og robotudskæringer, nye mælkeprodukter af meget ensartet kvalitet, optimal og tidlig indsats om trusler mod dyrenes sundhed og trivsel, hvilket resulterer i besparelser og øget produktivitet. Mulige partnere er: Besætningsejere, slagterier, mejerier, SEGES, DI ITEK, sensorfabrikanter, Fødevarestyrelsen og DTU.

Ligeledes er det væsentligt at inddrage relevante myndigheder. Rapporten peger på en række områder, hvor Big Data kan fremme myndighedernes monitorering og tilsyn. Det gælder bl.a. en mere effektiv og på sigt billigere kødkontrol. Det vil forudsætte, at rammebetingelserne for tilsynet ændres på EU-niveau. I det hele taget er det afgørende at tage højde for eksisterende og kommende fødevareregulering i EU og Danmark.

Det er afgørende, at de nye projekter gør Big Data håndgribeligt for dansk erhvervsliv. Målet er at gøre potentialet klart for den enkelte virksomhed. Herunder er målet at udvikle værktøjer, løsninger og uddannelsesforløb, der kan skabe sikkerhed for reelle afkast af investeringerne. Et væsentligt element vil være at sætte fokus på, hvordan danske virksomheder kan skabe en konkurrencemæssig fordel ved at indarbejde IT-sikkerhed og privacy i deres Big Data-løsninger. En væsentlig erfaring fra projektet er, at Big Data også kan handle om at eksperimentere med data, man ikke på forhånd ved, om der kan skabes værdi af. Projekterne, der igangsættes, kan derfor fungere som ”lærende værksteder”, hvor virksomheder og myndigheder med lav risiko kan få overblik og kan eksperimentere, afprøve og løfte deres brug af data ved hjælp af Big Data værktøjer. Samtidig kan sådanne værksteder give virksomhederne en tillidsbaseret ramme at dele data eller samproducere data i såvel internt i egen organisation og eksternt i samarbejde med andre aktører, eksempelvis andre virksomheder eller offentlige myndigheder.

## Anbefaling 2:

# Skab rammerne for et tættere samarbejde om deling af data

Rapporten peger på, at et værdiskabende arbejde med Big Data forudsætter et tæt samarbejde mellem videninstitutioner, erhvervsliv på tværs af værdikæder, organisationer og myndigheder, idet der er en lang række aspekter inden for forskning, udvikling, tests, regulering og organisering, som er sammenhængende og griber ind i hinanden. Samtidigt har arbejdet med sektorudviklingsprojektet synliggjort, at der er tegn på udvikling af en klynge af aktører, der på forskellig vis har relation til fødevarer og Big Data.

Derfor er der gode grunde til at udvikle en national klynge, hvor alle relevante aktører trækker i samme retning. DTU vil gerne bidrage til at udvikle klyngen. Klyngen kan eksempelvis bestå af virksomheder i fødevarerhvervet, IT-virksomheder, DTU, DI ITEK, DI Fødevarer, Landbrug & Fødevarer, SEGES, Region Midtjylland, Region Sjælland, Region Hovedstaden og Fødevestyrelsen. Klyngen bør tænkes ind i international sammenhæng, bl.a. i forhold til

det offentlig-private partnerskab om Big Data under EU's rammeprogram for forskning og innovation Horizon2020.

Rapporten peger på, at en afgørende faktor for at udvikle en klynge inden for området er, at der eksisterer en tillidsfuld ramme for at udlevere og dele data på tværs af individer, virksomheder, myndigheder og videninstitutioner. Rapporten anbefaler derfor, at de centrale aktører i klyngen i fællesskab arbejder for at skabe klare aftaleforhold om produktion, adgang til, lagring og brug af data, så de tillidsbårne samarbejdsrelationer kan opretholdes. Det indebærer bl.a., at det sikres, at universiteter på en og samme tid kan samarbejde med erhvervslivet om brug af erhvervslivets data eller/og deling af data med erhvervslivet og kan yde forskningsbaseret rådgivning til myndigheder, der har ansvar for, at EU-regulering og dansk lovgivning håndhæves over for erhvervslivet.

## Anbefaling 3:

# Styrk forskningen i tværgående Big Data værktøjer

Rapporten dokumenterer et stort behov for forskning på tværs af Big Data værktøjer. Rapporten dokumenterer også, at der er behov for forskning i, hvordan Big Data værktøjer i forskellige sektorer integreres. I det lys bør der etableres et center for forskning i Big Data.

Et center for Data Science skal imødekomme behovet for forskning i indsamling og håndtering af data, dataanalyse, datadeling, fortolkning af data samt forskning i datasikkerhed, privacy og klare aftaleforhold. Denne viden skal gøre os bedre i stand til at skabe Big Data værktøjer på tværs. Centret bør både sætte fokus på forskning, der på kort sigt har værdi for erhvervslivet og samfundet, og på forskning, der kan bidrage til langsigtet kompetenceopbygning inden for området. Samtidig skal centret have et tværfagligt perspektiv, der har fokus på tilknyttede forskningsområder som Internet of Things og Smart manufacturing.

Centret kan placeres på DTU. Institutterne DTU Compute og DTU Systembiologi kan bidrage med den tværgående forskning, mens relevante DTU faginstitutioner (bl.a. DTU Fødevareinstituttet og DTU Veterinærinstituttet)

kan bidrage med forskning inden for specifikke områder. Centret kan organiseres med andre videninstitutioner med relevante kompetencer, der supplerer DTU's.

Centrets arbejde skal bygge på en finansieringsstrategi, der udarbejdes på tværs af relevante institutter på DTU. Strategien skal have sigte på at finde finansiering til aktiviteter via Innovationsfonden, Horizon2020 og private fonde. Et væsentligt led i centret er at udnytte DTU's deltagelse i Big Data Value Association (BDVA), hvor DTU er "full member". BDVA er et public-private partnership, som over de næste 5-6 år støttes med 500 mio. euro af EU gennem Horizon 2020. Dertil kommer minimum et tilsvarende beløb fra industrien. Målet er, bl.a. med udgangspunkt i sektorudviklingsprojektet, at skabe indflydelse på EU-dagsordenen ift. Big Data og fødevarer, ikke mindst på kommende calls i Horizon 2020. Et konkret mål er at gøre Big Data inden for fødevarer til et europæisk fyrtårnsprojekt inden for rammerne af BDVA.

Det vil ligeledes være en oplagt opgave for centret at bidrage til at følge op på Danmarks Vækstråds anbefalinger angående Big Data.

## Anbefaling 4:

# Uddan fremtidens medarbejdere

En væsentlig forudsætning for at udnytte Big Data er, at kompetencerne er til stede i den private og offentlige sektor. Enten i form af nyuddannede kandidater eller opkvalificering af eksisterende medarbejdere.

DI ITEK vurderer, at den samlede efterspørgsel efter it- og elektronikkandidater frem til 2020 viser en mangel i størrelsesordenen 3.000 kandidater. Efterspørgsel efter ingeniørmæssige kompetencer specifikt inden for Big Data vil stige i takt med, at flere virksomheder ser mulighederne.

De danske videninstitutioner bør fremadrettet udbyde nye uddannelser og efteruddannelser, der giver den studerende kompetencer inden for:

- Big Data værktøjer
- Ledelse og organisering af Big Data

- Viden om fødevarekvalitet og fødevarerikkerhed og kobling til Big Data
- Privacy, etiske hensyn og klare aftaleforhold i forbindelse med datasikkerhed

DTU er allerede – ligesom andre universiteter – i gang med at opruste på uddannelsessiden inden for Big Data. DTU har i 2013 etableret en uddannelse for cybersikkerhed og i 2014 etableret en uddannelse i Data Science. Derudover har DTU som et af de første universiteter i Norden fået mulighed for at tilbyde de studerende og forskere adgang til at udnytte IBM's Big Data værktøjer, IBM Bluemix og Watson.

Det er en udvikling, DTU vil fortsætte fremover for at imødekomme samfundets behov for nye kompetencer inden for anvendelse af Big Data.

## Anbefaling 5:

# Styrk iværksætteri inden for Big Data

Iværksætteri og vækstvirksomheder spiller en væsentlig rolle i at skabe teknologisk udvikling i en erhvervssektor. Der etableres et stigende antal iværksættervirksomheder med afsæt i Big Data, og flere mindre virksomheder ser vækstmuligheder i Big Data og relaterede tendenser såsom Internet of Things. Erfaringen fra rapporten er, at det ikke alene er store etablerede virksomheder, der kan drive udviklingen inden for Big Data, men at der i høj grad også er et potentiale for mindre virksomheder og iværksættere for at drive udviklingen. Der er derfor behov for at tænke iværksætteri ind i en samlet indsats for at styrke Big Data i fødevaresektoren og at udnytte det innovations-øko-system, der allerede er på fx universiteter, i offentlige programmer og i virksomheder.

DTU arbejder løbende for at fremme iværksætteri og innovation og skal her medvirke til at fremme spinoff-virksomheder – også inden for Big Data. Der er dog et særligt behov for etablering af virksomheder med et forretningsgrundlag, der går på tværs af de mange bestanddele, som er nødvendige for optimal udnyttelse af Big Data. Typisk

kræver det samarbejde med virksomheder inden for fødevareerhvervet og i IT-sektoren.

Et væsentligt element i en sådan indsats er at sætte fokus på studenterinnovation. Det kan enten være i form af erhvervsrettede aktiviteter i forbindelse med de studerendes uddannelse, fx aktiviteter, der sikrer flere studenterprojekter i virksomhederne.

Eller det kan være ekstra-curriculære aktiviteter såsom Hackathons, der kan føre til, at studerende udvikler Big Data løsninger for virksomheder eller offentlige institutioner, og at de studerende etablerer start ups baseret på disse løsninger. DTU og DTU Skylab er netop i samarbejde med en række virksomheder ved at etablere initiativet ”Open Innovation – X”, der skal skabe rammer for at virksomhederne kan komme tættere på de studerende og kan udnytte de studerendes innovationskraft inden for fx Big data, Internet of things mv.

## Potentialet for Danmark

Dialogen med virksomheder, brancheorganisationer og myndigheder i løbet af sektorudviklingsprojektet har vist:

- Fødevareerhvervet er internationalt meget konkurrencedygtigt og i fuld færd med at dygtiggøre sig inden for de parametre, som bliver afgørende for at udvikle konkurrenceevnen både på kort og langt sigt.
- Fødevaremyndighederne har en unik mulighed for at blive en positiv medspiller, særligt i etableringen af rammerne for deling og samproduktion af data under klare aftaleforhold.
- Big Data er et vigtigt og brugbart redskab til forretningsudvikling.
- Fødevareerhvervet og relevante myndigheder har unikke historiske data, som rummer et stort erhvervsmæssigt potentiale.

- Relevante uddannelser af kandidater i Big Data er allerede etableret.
- Der er store produktivetsgevinster og udsigt til vækst i eksportomsætningen ved at satse på Big Data i fødevareerhvervet.
- En styrkelse af fødevaresektoren gennem Big Data vil få positiv betydning for beskæftigelsen i Danmark. Den øgede beskæftigelse vil i høj grad komme landdistrikterne til gode og vil tilgodese et bredt spekter af medarbejdere fra ufaglærte over faglærte til højt specialiserede medarbejdere.

Sammenfattende kan det konkluderes, at der er et endog meget stort potentiale for Danmark i at udnytte mulighederne i Big Data i fødevaresektoren.



## Matcher anbefalinger fra vækstteam

Den tidligere regerings Vækstteam for Fødevarer anbefalede i 2013:

- Udvikling af værktøjer til at dokumentere og underbygge fortællingen om en bæredygtig og ressource-effektiv fødevarersektor
- Regulering af fødevarersektoren på dens miljøbelastning og ikke på, hvor meget der produceres
- Fokus på produktions- og procesteknologi i fødevarersektoren for at styrke virksomhedernes værdiskabelse og ressourceeffektivitet
- Initiativ om offentlig-privat partnerskab inden for fødevarersektoren for at skabe de bedste vilkår for radikal innovation, fx inden for udvikling af nye kølesystemer, forarbejdning af biomasse til nye materialer, produkter med øgede sundhedsfremmende kvaliteter og bæredygtige emballageløsninger

Det er vækstteamets forventning, at tiltagene kan øge eksporten fra fødevarersektoren med mindst 30 mia. kr. og skabe mindst 10.000 nye arbejdspladser.

Udnyttelsen af Big Data i fødevarersektoren vil bidrage til at fremme alle de fire anbefalinger fra vækstteamet.

## Ny kandidatuddannelse i Big Data

Rækken af Big Data værktøjer er lang. Ligeledes kræves viden inden for specifikke fagområder - i dette tilfælde fødevarerområdet - så man kan udnytte Big Data optimalt.

Ifølge en analyse fra Erhvervsstyrelsen fra 2013 mangler de virksomheder, der ønsker at styrke deres forretning gennem Big Data, kvalificerede medarbejdere. Det kræver handling fra uddannelsessystemet både ift. udbud af de rette erhvervsrettede uddannelser, universitetsuddannelser og efter- og videreuddannelser.

I lyset af det behov har DTU i 2014 lanceret et nyt uddannelsesforløb, hvor kandidatstuderende får mulighed for at specialisere sig inden for en lang række centrale områder inden for Big Data, fx data opsamling, avanceret statistik, visualisering og datasikkerhed. Uddannelsesspecialiseringen er blevet etableret i tæt samarbejde med erhvervslivet for at sikre uddannelsens relevans og kvalitet, og uddannelsen ligger godt i forlængelse af anbefalingerne fra regeringens Vækstteam for IKT og digital vækst om en styrkelse af de digitale kompetencer.

# Appendix



## Hvad mener vi med Big Data?

Der findes ingen alment accepteret definition på, hvornår man går fra almindelig udnyttelse af data til Big Data.

Baseret på erfaringerne sektorudviklingsprojektet og international litteratur er rapportens definition, at:

- Man med udgangspunkt i meget store mængder data udvikler algoritmer, der kan skabe mønstergenkendelse i dataene og deraf afledte indsigter, som støtter udviklingen af ny forretningsmæssig værdi
- Analysen mhp. mønstergenkendelsen kan eksempelvis foretages ved hjælp af værktøjer som induktiv statistik og nonlinear systemidentifikation
- Dataene kan være interne og/eller eksterne, af stor varians, produceres meget hurtigt og være af varierende kvalitet
- De opnåede indsigter præsenteres ofte gennem visualiseringer, eventuelt i form af applikationer
- Man enten besvarer allerede kendte spørgsmål eller besvarer helt nye spørgsmål, som genereres i løbet af processen, og som man i nogle tilfælde end ikke i udgangspunktet vidste, at man gerne ville besvare.
- Mht. selve dataene er det nye, at der er tale om meget store mængder
- Omkostningsniveauet forbundet med at udnytte de meget store mængder data er meget lavere indtil videre. De lave omkostninger skyldes blandt andet, at det er blevet markant billigere at lagre data
- Analysen er ny i den forstand, at de platforme, værktøjer og software, man anvender, kan analysere de meget store mængder data hurtigt og billigt på en måde, som standard IT-teknologier ikke kan

Med til en definition af Big Data hører også en afgrænsning fra andre velkendte tilgange til udvikling af forretningsmæssig værdi ved hjælp af data. Især er det vigtigt at skelne datadrevet forretningsudvikling, Business Intelligence (BI) og Big Data fra hinanden.

### Datadrevet forretningsudvikling vs. Big Data:

- Datadrevet forretningsudvikling handler om at bruge data og dataanalyse i selve forretningsudviklingen. Data udgør selve grundlaget for, hvordan man markedsfører sig, produktudvikler og servicerer kunderne. Big Data udvider dette. Her er der tale om alle de karakteristika, der er opstillet umiddelbart ovenfor.

### Business Intelligence (BI) vs. Big Data:

- I BI anvender man data alene mhp. at besvare allerede kendte spørgsmål. I Big Data kan der være tale om enten at besvare allerede kendte spørgsmål eller besvare helt nye spørgsmål, som genereres i løbet af processen.
- I BI anvendes typisk deskriptiv statistik mhp. at identificere trends og måle på allerede fastlagt måleparametre. I Big Data anvendes også mere avancerede former for værktøjer som fx induktiv statistik og nonlinear systemidentifikation mhp. at skabe mønstergenkendelse i data og skabe deraf afledte indsigter, som støtter udviklingen af ny forretningsmæssig værdi
- I BI indsamler man data mhp. at stille kendte spørgsmål og rapportere. I Big Data er der tale om egentlig forretningsudvikling.
- I BI foregår indsamlingen fortrinsvist fra interne datakilder, og den kan ofte ske ved at hente data ved hjælp af online analytiske processer. I Big Data indsamles og produceres data fra både interne og eksterne kilder og kan ske ved hjælp af mange forskellige værktøjer.

Konsekvenserne af fremkomsten af Big Data er, at data, teknologi, forretningsmodeller, funktionsområdet IT-afdelingen og dataejere (offentlig som privat) skifter karakter:

- FRA passive og ustrukturerede data TIL aktive og strukturerede data, der kan anvendes i virksomheders innovation af produkter og tjenesteydelser samt udvikling af forretningsmodeller.
- FRA teknologi som motoren i innovation TIL teknologi som værdiskabende element i produkter og tjenesteydelser og i forretningsmodeller
- FRA IT-afdelingen primært som driftsressource TIL IT-afdelingen som udviklingsressource på lige fod med udviklings-, salgs- og marketingafdelingen
- FRA data, som myndigheder selv anvender til at regulere og oplyse privatpersoner, virksomheder og andre myndigheder med, TIL data, som myndighederne stiller til rådighed for private virksomheder, regioner og kommuner

$$f(x+\Delta x) = \sum_{i=0}^{\infty} \frac{(\Delta x)^i}{i!} f^{(i)}(x)$$

