Rapport 2023:19

Big Data Analytics and Firm Productivity
- A Literature Review

In the last decade, firms have made significant investments in acquiring the technology and human resources necessary to make full use of the capabilities of big data analytics (BDA). Through a systematic review of 94 published studies, this report analyzes the concept of BDA, explores its diverse impact on firm productivity, and offers practical insights for firms and policymakers to leverage the full potential of BDA.

Tillväxtanalys
Myndigheten för tillväxtpolitiska utvärderingar och analyser

Dnr: 2023/9
Myndigheten för tillväxtpolitiska utvärderingar och analyser
Studentplan 3, 831 40 Östersund
Telefon: 010 447 44 00
E-post: info@tillvaxtanalys.se
www.tillvaxtanalys.se
För ytterligare information kontakta: Ismail Ouraich
Telefon: 010-447 44 28
E-post: ismail.ouraich@tillvaxtanalys.se

Förord

Tillväxtanalys uppdrag är att utvärdera och analysera effekterna av statens insatser för en hållbar nationell och regional tillväxt. Vi ska också ge underlag och rekommendationer för utveckling, omprövning och effektivisering av politiken.

Den här rapporten ingår i Tillväxtanalys ramprojekt *Stordataanalys betydelse för svenska företags produktivitet*. Projektet syftar till att både sammanställa befintlig forskning om företags implementering och användning av stordataanalys (*big data analytics*; härefter BDA) och empiriskt undersöka dess effekter på produktivitet och andra variabler. Projektet belyser också vilka egenskaper företag som använder BDA har och hur svenska företags användning av BDA förhåller sig till företag i andra europeiska länder.

Tillväxtanalys har tidigare genomfört flera studier inom angränsande områden. Redan 2014 initierades flera projekt relaterade till att analysera digitaliseringen av det svenska näringslivet och dess bidrag till ekonomisk tillväxt och konkurrenskraft. Under 2023 avslutade Tillväxtanalys ett ramprojekt som undersökte Artificiell Intelligens (AI) betydelse för svenskt näringsliv och hur politiken kan utvecklas för att underlätta företagens AI-användning. Rapporterna har varit både litteratursammanställningar av befintlig forskning och empiriska studier baserade på mikrodata som bland annat analyserat näringslivets digitala mognad och sambandet mellan IT-användning och företagens produktivitet. Det här ramprojektet är en fortsättning på Tillväxtanalys arbete inom digitalisering och dess betydelse för svenskt näringslivs tillväxt och konkurrenskraft.

Den här rapporten är skriven av Patrick Mikalef, professor i datavetenskap och informationssystem vid institutionen för datavetenskap vid NTNU. Ismail Ouraich har varit projektledare och i projektet har också Petter Svärd medverkat. Ett särskilt tack riktas till projektets referensgrupp för många och värdefulla kommentarer. Referensgruppen har bestått av Irene Ek, Public Policy Manager vid Google Clouds Nordics; Erik Borälv, Vinnova; Daniel Gillblad, Föreståndare vid Chalmers AI Research Center; Fredrik Sand, Näringspolitik expert vid TechSverige och Susanne Stenberg, Senior forskare och rättslig expert vid RISE. Rapporten är skriven på engelska men med en utförlig svensk sammanfattning.

Östersund november 2023

Sverker Härd,
Generaldirektör, Tillväxtanalys

Innehållsförteckning

Sammanfattning

Stordata (eng. *big data*) utgör en allt viktigare insatsvara i företags strävan efter ökad produktivitet, effektivitet och konkurrenskraft. Genom teknologiska framsteg såsom ökad processorkraft och lagringskapacitet, snabbare och mer omfattande nätverk samt utveckling av sensorer som kan mäta och överföra data i realtid, har tillgången av stordata ökat markant det senaste decenniet. Men för att företag ska kunna kapitalisera på stordatans potential så krävs även en förmåga att genom olika verktyg och tekniker hantera och analysera stordata – *stordataanalys*, eller *big data analytics* på engelska, nedan kallat BDA. BDA möjliggör ett mer datadrivet beslutsfattande i företags strategiska och operationella verksamhet, vilket på olika sätt kan öka företagens produktivitet.

Under det senaste decenniet har allt fler företag insett värdet i stordata och investerat i teknisk infrastruktur och humankapital för att öka sin BDA-kapacitet. I takt med att antalet företag som implementerar BDA ökar, har även antalet studier ökat för att undersöka hur företag använder BDA och dess påverkan på produktivitet samt andra viktiga indikatorer som marknadsandel, marknadsposition, effektivitetsvinster, finansiell utveckling, kundnöjdhet och innovation inom produkter och tjänster.

Den här rapporten är en systematisk litteraturgenomgång av 94 studier publicerade under de senaste tio åren. Rapporten sammanfattar den aktuella forskningen och är uppdelad i tre huvuddelar: Den första delen belyser den begreppsmässiga utvecklingen av *big data* (stordata) och *big data analytics* (stordataanalys; BDA) samt definierar och avgränsar begreppen och dess egenskaper från andra närliggande begrepp såsom artificiell intelligens. Den andra delen redogör för den empiriska evidensen för hur BDA påverkar företagens produktivitet och andra relevanta indikatorer, samt genom vilka kanaler denna påverkan uppstår. Studierna har baserats antingen på *objektiv* eller *subjektiv* självrapporterade data. Den tredje delen lyfter policyimplikationer för att underlätta investeringar och användning av BDA inom svenskt näringsliv utifrån ett individ-, företags- och branschperspektiv.

Tillväxtanalys tidigare studier inom digitaliseringsområdet

Tillväxtanalys har tidigare genomfört flera studier inom angränsande områden. Redan 2014 initierades flera projekt relaterade till att analysera digitaliseringen av det svenska näringslivet och dess bidrag till ekonomisk tillväxt och konkurrenskraft. Under 2023 avslutade Tillväxtanalys ett ramprojekt som undersökte Artificiell Intelligens (AI) betydelse för svenskt näringsliv och hur politiken kan utvecklas för att underlätta företagens AI-användning. Rapporterna har varit både litteratursammanställningar av befintlig forskning och empiriska studier baserade på mikrodata som bland annat analyserat näringslivets digitala mognad och sambandet mellan IT-användning och företagens produktivitet. Det här ramprojektet är en fortsättning på Tillväxtanalys arbete inom digitalisering och dess betydelse för svenskt näringslivs tillväxt och konkurrenskraft. Se även appendix C för en redogörelse för Tillväxtanalys tidigare rapporter relaterade till svenska företags digitalisering.

Vad kännetecknar big data och BDA?

I början av 2010-talet populariserades det datavetenskapliga begreppet *big data* för att beskriva data som var för stora och/eller för komplexa för att hanteras eller analyseras genom konventionella bearbetningsmetoder. Stordata i sig är en obearbetad råvara och för att vara lämplig för analys menar forskningen att den måste uppvisa hög volym, hastighet, variation och tillförlitlighet (eng. *the four "Vs": volume, velocity, variety och veracity*).

BDA syftar till att analysera och extrahera kunskap och information från stordata genom olika analytiska tekniker och metoder. Målet med BDA är att upptäcka samband, mönster, trender och annan användbar information för att möjliggöra bättre beslutsfattande. BDA bygger vidare på tidigare metoder såsom affärsanalys och datautvinning, men utgör ett nytt paradigm av digital transformation som innefattar nya teknologier, tekniker, färdigheter, processer och tankesätt. BDA är teknik- och metodoberoende där ingen särskild typ av verktyg eller metod dikterar vad som konstituerar BDA. Denna egenskap skiljer BDA från exempelvis artificiell intelligens (AI), som syftar till att skapa intelligenta system som kan lära sig av erfarenhet för att utföra uppgifter som traditionellt kräver mänsklig intelligens. Trots att BDA och AI tjänar olika syften kompletterar de varandra; BDA använder sig ofta av AI-tekniker, såsom maskininlärningsalgoritmer, för att upptäcka mönster och utvinna kunskap från stora datamängder. Å andra sidan tränas och utvecklas AI-modeller genom att använda omfattande datamängder via BDA.

BDA kan öka företags produktivitet

Studier baserade på objektiva data har identifierat ett kausalt samband mellan företags investeringar i BDA inom vissa branscher, såsom tillverknings- eller teknologibranscherna, med årliga produktivitetsvinster på 3 till 7 procent. Överlag var investeringar fördelaktigare inom mer konkurrensutsatta och/eller teknikinriktade branscher. I icke-konkurrensutsatta branscher observerades inga signifikanta effekter av BDA-investeringar. En annan studie baserad på objektiva data indikerar att investeringar i BDA inom servicesektorn resulterar i förbättrad innovationsförmåga av tjänster och produkter, samtidigt som sannolikheten för att dessa produkter blir framgångsrika ökar. Utfallet var dock beroende av i vilken utsträckning dessa företag hade investerat i ändamålsenligt humankapital, genom rekrytering eller utbildning av anställda.

Ett flertal studier baserade på självrapporterade data, så som enkäter, indikerar att företag som implementerar BDA i sin verksamhet upplever en 10–25 procentig ökning av konkurrenskraften jämfört med deras främsta konkurrenter. Andra studier baserade på självrapporterade data har påvisat att investeringar i BDA har en positiv och signifikant påverkan på ökad innovation inom tjänster, processer och produkter. Slutligen visar forskningen att investeringar i BDA kan ge förbättringar av företagens operationella effektivitet, förutsatt att de integreras organiskt i företagens kärnverksamhet.

Även om vissa företags BDA-investeringar har lett till ökad produktivitet visar litteratursammanställningen i den här rapporten att långt ifrån alla investeringar i BDA är lönsamma. För en betydande del av företagen som investerat i BDA tar det lång tid innan lönsamhet kan uppnås. De flesta BDA-investeringarna kommer aldrig ens till produktionsfasen utan avstannar redan i prototypstadiet. Studierna visar också att en framgångsrik investering också är avhängig *på vilket sätt* den implementeras i företagets organisation och verksamhet. Andra studier visar att många företag där BDA skulle kunna vara fördelaktig inte investerat på grund av organisatoriska, teknologiska och ekonomiska begränsningar.

Produktivitetsvinster uppstår indirekt och över tid

Forskning visar att det tar ett till tre år innan företagens investeringar i BDA resulterar i ökad produktivitet. Fördröjningen beror på det omfattande förberedelsearbetet, experimenteringen och testningen som krävs innan BDA kan integreras effektivt i verksamheten. Det är värt att notera att fördröjningen inte finns på samma sätt när företag implementerar andra teknologier.

När företags användning av BDA leder till förbättringar i effektivitet och produktivitet, sker dessa i regel genom indirekta kanaler såsom förbättrade besluts- och verksamhetsprocesser. En sådan indirekt kanal är, enligt forskningen, att genom effektiv identifiering av förändrade omständigheter optimera beslutsprocesserna. Detta kan innebära att anpassa sig efter skiftande kundpreferenser för att nå nya eller behålla befintliga kunder, analysera konkurrenters ageranden och vilka konsekvenser det medför, eller utforska nya marknadsmöjligheter.

En annan indirekt kanal involverar granskning och optimering av interna verksamhetsprocesser. Studier har undersökt hur BDA kan generera värde genom användningsområden såsom s.k. förebyggande underhåll, optimering av leveranskedjor och effektivisering av produktionsprocesser. Inom branscher som tillverkning, logistik, bearbetning av råmaterial, olja och gas samt transport är värdeskapande genom denna indirekta kanal särskilt framträdande.

Produktivitetsvinster beror på bransch och användningsområde

Forskningen visar att förmågan hos BDA-investeringar att skapa värde är starkt kopplad till både den specifika bransch och de konkreta användningsområden där BDA används. För närvarande drar vissa branscher, som till exempel mer tekniskt inriktade branscher, större nytta och har en mer omfattande användning av BDA jämfört med andra branscher. Det är en trend som har observerats globalt och kopplas främst till branschernas höga konkurrenskraft, både nationellt och internationellt, samt deras förmåga att snabbt digitalisera sin verksamhet. Även större företag använder sig av BDA i högre utsträckning jämfört med små och medelstora företag.

Detta är, tillsammans med det faktum att produktivitetsökningarna i regel sker genom indirekta kanaler och med en fördröjning, en utmaning när det kommer till att adekvat mäta och jämföra effekterna av BDA i varierande kontexter. För att bättre fånga upp BDA:s olika kanaler för värdeskapande är det viktigt att utveckla och använda mått och indikatorer som sträcker sig bortom enbart finansiella aspekter och inkludera faktorer som innovationskraft, marknadsandelar och verksamhetseffektiviseringar.

Kompletterande investeringar i humankapital och organisationskapacitet är viktiga

En viktig slutsats från forskningen är att värdeskapande genom BDA kräver utöver konventionella investeringar i teknisk infrastruktur även kompletterande investeringar i humankapital och organisationskapacitet – även kallad ”BDA-kapacitet”.

Gällande humankapital visar flera studier att företag som investerar i utbildning eller rekrytering av personal med rätt kompetens och färdigheter får högre lönsamhet från sina BDA-investeringar. Behovet är särskilt påtagligt inom den tekniska implementeringen, där dataingenjörer, dataarkitekter och datavetare krävs för att effektivt samla in, bearbeta och analysera data. Trots detta utgör utbildning eller rekrytering av personal med rätt kompetenser en betydande utmaning för många företag. I synnerhet små och medelstora företag har svårighet att rekrytera eller utbilda individer med rätt kompetenser, ofta till följd av begränsade resurser och färre anställda jämfört med större företag.

Gällande företags organisationskapacitet menar forskningen att det är viktigt att stordata behandlas som en central insatsresurs och att datadrivet beslutstagande utgör en kärnkompetens som genomsyrar hela organisationen, snarare än att enbart betraktas som en teknisk uppgift för IT-avdelningen. Detta handlar inte enbart om att ha tillgång till rätt tekniska verktyg, utan också om att forma en företagskultur där integrering och analys av data ses som en integrerad del av beslutsfattandet på alla nivåer. Studier visar att de företag som betraktar BDA som en central prioritering och som säkerställer transparens och tillgång till kritiska datakällor samt ger utrymme för experimentering, får större framgång för sina investeringar.

Policyimplikationer

Sammanställningen av tidigare empirisk forskning visar att BDA kan leda till betydande produktivitetsvinster för företag, under vissa förutsättningar. Även om vissa branscher och typer av företag hittills haft större nytta av BDA än andra (t.ex. tillverknings- och teknologibranscherna), finns det en del utmaningar som de flesta företag möter när de implementerar och använder BDA. Dessa utmaningar återkommer ofta i studier och utgör viktig kunskap i syfte att ta fram implikationer för beslutsfattande. I den här rapporten presenteras de i tre kategorier av företags förmåga att använda BDA: kunskaper och färdigheter, nödvändiga villkor på företagsnivå och underlättande faktorer på branschnivå.

Kunskaper och färdigheter hos enskilda personer

Ett av de främsta hindren för många företag som planerar att implementera BDA är svårigheten att hitta medarbetare med rätt kompetens. Flera empiriska studier har specifikt fokuserat på frågan om vilka kunskaper och färdigheter som behövs i stordataåldern, och även lyft fram några områden där det för närvarande finns stora brister. En studie undersökte till exempel de brister som finns i norska datavetenskapsstudenters utbildning i förhållande till branschens behov och konstaterade att företagen behöver anställda med tekniska, administrativa och mjuka kompetenser relaterade till datavetenskap. Samtidigt menade författarna att de akademiska institutionerna inte riktigt erbjöd utbildningar som tillgodoság kunskap kring användning av industrispecifika verktyg och tekniker. Även samarbetsinriktade och ämnesöverskridande färdigheter som möjliggör arbete med analys av stora datamängder erbjöds i otillräcklig utsträckning. Studien betonade hur viktiga företagen ansåg att dessa färdigheter var för att skapa konkurrenskraftig implementering och användning BDA. Vidare framgick vikten av att företagen har tillräckligt med resurser och finansiering för att utbilda sin personal genom flexibla inlärningsmiljöer och utbildning på arbetsplatsen.

Sammantaget menade forskningen att det finns ett behov av incitament för att främja vidareutbildning i datavetenskapliga discipliner som företagen efterfrågar, samt incitament att dessa genomförs i nära samspel med branschens krav och behov. Detta har även lyfts fram i Tillväxtanalys ramrapport *Hur omformar AI näringslivet och hur kan politiken utvecklas?* från 2023, som föreslår att staten kan underlätta investeringar i AI genom att bidra till att ge bättre samverkan mellan lärosätena och näringslivet avseende kompetensförsörjning.

Nödvändiga villkor på företagsnivå

Ett annat viktigt resultat som framkommer i litteraturgenomgången är att stora företag använder BDA i större utsträckning än små och medelstora företag. Studier visar att sambandet har att göra med den höga kostnad som är förknippad med investeringar i BDA, vilket utgör ett hinder för många små och medelstora företag. I synnerhet kräver tekniska lösningar för datalagring och databehandling betydande investeringar, som särskilt många små och medelstora företag tvekar att göra på grund av på förhand oklara fördelar. Framför allt företag inom icke-teknikorienterade branscher släpar efter vid investeringar och användning av BDA. Studier lyfter vikten av incitament för att tillhandahålla teknisk infrastruktur särskilt till små och medelstora företag men även till branscher som kan dra nytta av BDA, vilket kan bidra till ökad konkurrenskraft och produktivitet.

Studier lyfter att särskilt förvaltning och ägande av data som svåröverskådliga områden vid implementering av BDA, vilket följer bland annat av regelverken kring den allmänna dataskyddsförordningen (GDPR)[1] och EU:s Datalag[2]. Eftersom många tjänsteleverantörer erbjuder molnbaserade BDA-lösningar menar studier att det är otydligt för företag hur personuppgifter eller känsliga uppgifter bör hanteras för att överensstämma med dessa förordningar. Detta är särskilt utmanande för företag med mindre erfarenhet av teknikbaserade initiativ, såsom BDA, där det fanns en brist på teknisk och rättslig kunskap. Liknande slutsatser lyfts fram från Tillväxtanalys rapport *Drivers of AI adoption – A literature review* från 2021. För att minska detta hinder för företagen skulle det vara fördelaktigt att fastställa nationella riktlinjer och lättillgängliga ramar med viktig information om hur företagens dataförvaltningen bör utföras och vad som behöver tänkas på vid val av tjänsteleverantör.

Ett viktigt resultat från litteraturstudien är att BDA leder till värdeskapande för företagen på ett indirekt sätt samt att detta medför fördröjningseffekter. Det är därför två faktorer

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679
[2] https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113

som måste beaktas när man utvärderar företags värdeskapande av BDA. Den första är att resultatförbättringar kan vara ganska svåra att mäta och kräver en bredare uppsättning indikatorer för att korrekt mäta effekterna av BDA, som också är beroende av bransch och användningsområde. Den andra är att BDA-investeringar kräver avsevärd tid för att mogna och skapa värde för företagen. Detta understryker sammantaget behovet av incitament som ger företagen tillräckligt med tid att utveckla BDA till en central organisatorisk förmåga. Att effekterna uppstår med lång fördröjning visar också på vikten av att resultat följs upp med en tillräckligt lång tidshorisont.

Underlättande faktorer på branschnivå

Som tidigare nämnts är vissa branscher mer benägna att implementera BDA, samtidigt som andra branscher släpar efter betydligt. Detta fenomen har identifierats i flera länder och kopplas till branschernas konkurrenskraft och snabbhet när det gäller att driva på digitaliseringen av verksamheten. Men även andra faktorer än branschdynamik leder till att vissa företag investerar i BDA eller befinner sig i ett mycket tidigt skede av implementeringen. Litteraturgenomgången pekar på att finansiella incitament och stimulansåtgärder för att digitalisera verksamheten kan utgöra viktiga drivkrafter för företag att investera i BDA. Dessa möjliggör för företagen att ta det första steget och skaffa nödvändig infrastruktur för att utnyttja sina data samt bygga upp den tekniska infrastruktur och det humankapital som krävs för att på bästa sätt använda BDA och öka produktiviteten.

Dessa åtgärder har framhållits som viktiga för att öka företagens användning av ny digital teknik och bli mer konkurrenskraftiga på den globala marknaden. Att skapa incitament för teknisk infrastruktur, särskilt för små och medelstora företag, är avgörande med tanke på deras betydelse i den europeiska näringslivsstrukturen. Att utveckla särskilt anpassade finansiella incitament och stimulansåtgärder är därför avgörande för att underlätta deras investeringar i teknik som BDA. Många av dessa små och medelstora företag kanske ännu inte fullt ut inser värdet av teknik som BDA, vilket understryker behovet av en politik som underlättar införandet av och experimenterandet med digitala verktyg. Dessutom bidrar utbildnings- och fortbildningsprogram till att ge små och medelstora företag möjlighet att förbättra sina anställdas kompetens, vilket gör dem bättre rustade att dra nytta av tekniken. Sammantaget bör därför politiska åtgärder som stöder finansieringen av teknisk infrastruktur kompletteras med andra åtgärder som syftar till att förbättra kompetensen hos anställda i olika befattningar för effektiv användning av och värdeskapande med BDA.

Mer data och analys behövs

En viktig aspekt som framkom i litteratursammanställningen är den begränsade tillgången till mikrodata på nationell nivå, samt bristen på jämförelser mellan olika länder. Därutöver framkom att den empiriska litteraturen domineras av studier som använder sig *antingen* av subjektiva eller objektiva datakällor. Ett förslag när det gäller datainsamling vore att undersöka möjligheten till att tillvarata de olika fördelarna med både subjektiva respektive objektiva datakällor. Tillgången till pålitliga mikrodata över tid är en central förutsättning för förbättrade analyser, i syfte att undersöka kausala samband mellan användning av BDA och företags produktivitet. Behovet av bra data över tid som förutsättning för välgrundad analys lyftes även i Tillväxtanalys rapport från 2023 *En kartläggning av AI-användning och produktivitet bland svenska företag*.

För närvarande finns det inte heller, vad vi känner till, någon vetenskapligt granskad (eng. peer-reviewed) empirisk studie som använder svenska mikrodata för att undersöka hur just BDA påverkar företags produktivitet. Dock undersöker Tillväxtanalys studie ovan sambandet mellan AI-användning i svenska företag och produktivitet. Som en naturlig fortsättning kommer Tillväxtanalys undersöka sambandet, inom ramen av detta ramprojekt, mellan svenska företags BDA-användning och produktivitet. Den rapporten kommer i likhet med den nämnda AI-studien vara baserad på SCB:s IT-undersökningar och registerdata på företagsnivå.

Summary

Over the last decade, a growing body of research has examined the adoption and impact of big data analytics (hereafter BDA) for organizations. The research has been somewhat fragmented and has frequently fallen short of delivering clear and tangible findings that would be beneficial for businesses and policymakers. This has been attributed to the conceptual opacity surrounding the term BDA, and the fact that research has used and examined *different* performance indicators and mechanisms. This study aims, through applying a systematic literature review of 94 published studies over the past ten years, to provide a conceptual analysis of the concept of BDA, examine the impact that BDA has on firm productivity through different mechanisms and provide actionable insight for firms and policymakers.

BDA is a distinct component of the digital transformation process

BDA not only represents and defines data by specific attributes but also encompasses the underlying technologies supporting data management, analysis, and decision processes. Furthermore, it influences the organizational structure of firms and the skill sets of individuals involved in the process. BDA represents a novel paradigm of digital transformation, requiring investments in technological infrastructure, employee training and education in data handling, analytical techniques, and data-driven management and operational decision-making. The research community is also discussing the distinction between BDA and Artificial Intelligence (AI). BDA aims to make sense of data for predictions and decision-making, whereas AI develops applications that emulate human-like intelligence and behavior.

BDA has a positive effect on firm productivity

The analysis of past studies highlights that investing in BDA can produce annual gains of 3-7 percent in firm productivity. The effects of BDA are more pronounced in industries that are highly competitive or more technologically oriented. In noncompetitive industries, no significant effects are observed. Other studies show that there is a positive and significant effect between the adoption of BDA and innovation growth and that firms that leverage BDA in their operations realized a 10-25% performance improvement in comparison with key competitors. Gains are mostly found in more technologically oriented industries, such as the service sector, the manufacturing industry, and IT/technology industries.

Productivity gains from BDA are often realized indirectly

Another important finding in from the empirical studies emphasize that the value of BDA to firm productivity is often realized indirectly through improvements in organizations' operational or competitive strategies. These are realized through different "mechanisms" such as enhanced decision processes, improved operational efficiencies, more accurate forecasting, and cost reduction. The productivity gains are contingent on how efficient firms are in utilizing these. By effectively monitoring and optimizing internal operations, sensing changing customer beliefs and requirements, and perceiving

emerging opportunities and competitive actions, BDA can be a strategic tool for firms to make data-driven decisions. Thus, it is important to understand the organizational shifts that BDA enables and how those indirectly impact productivity and value-generation.

The effects of BDA are contingent upon industry and application uses

Studies suggest that the malleability of BDA depends on both internal and external factors. BDA applications can be applied to a wide range of industries, and their ability to create value depends on both the specific industry in which they are implemented and the processes they are intended to automate or enhance. For instance, certain industries that use physical devices to generate and collect data can more closely monitor and optimize processes. In addition, firms that operate in the same industry can have vastly different uses for their BDA investments, resulting in differentiating productivity gains.

Lag effects in realizing productivity gains from BDA investments

Following the adoption of BDA projects, according to studies it typically takes at least one year, with an average of two years, before most firms begin to witness tangible improvements in terms of performance gains. This delay differs from the adoption of other technologies that these firms commonly embrace. The lag effects are attributed to the multifaceted requirements of BDA, which necessitate substantial preparatory work, experimentation, and testing before its applications can be effectively integrated into operations.

Investments in complimentary skills and organizational structure increase BDA-adoption

One of the most important findings from the study of academic literature, is that generating value through BDA requires more than just investing in technological infrastructure. It also involves building organizational capability, often referred to as BDA capability. This includes necessary investments in data management and technology infrastructure, ensuring that employees at all levels have the required knowledge and skills, and establishing an organizational structure and strategic direction that recognizes data as a fundamental resource. When companies view BDA as a central strategic initiative that permeates the entire organization, it promotes transparency and access to critical data sources, which ensures that projects can leverage the technologies optimally.

Policy implications

To unlock the full potential of productivity growth and value creation, three complementary drivers of BDA adoption should be taken into account: individuals' knowledge and skills, organizational barriers at the firm level, and facilitators at the industry level. Individuals' skillset in working with BDA is in high demand and frequently constitutes a lack of competence among employees in small and medium enterprises (SMEs). This requires more on-the-job training opportunities, a reevaluation of educational curricula to foster cross-disciplinary skills, and incentives to develop practical experience through industry-related projects. Many firms, particularly SMEs, encounter substantial barriers when attempting to adopt BDA. These barriers primarily revolve around insufficient funding for technological infrastructure, as well as a shortage of expertise and available time for experimentation. At the industry level, some firms operating in less dynamic industries have been slower to adopt BDA, primarily due to the difficulty of identifying the tangible benefits of implementing such technologies. It has been demonstrated that stimulus and support incentives can be instrumental in boosting the adoption of BDA.

More data and analyses in specific contexts are needed

Despite the recent increase in the number of empirical studies on the value of BDA, there remains a significant gap in terms of context-specific insights into the productivity effects of BDA and access to objective data sources. Currently, there is no large-scale study using data from Swedish firms that provides a comprehensive understanding of how these technologies affect specific sectors or through which key performance indicators they influence them. For a better understanding of how policy measures can boost BDA adoption and, in turn, enhance the productivity of Swedish firms, future studies should prioritize analyzing these effects while considering investments, not only in technology, but also in other complementary resources like human capital and educational programs. As more data becomes accessible, it would be beneficial for more studies to assess the value of BDA adoption systematically, evaluating its impact on different aspects of productivity and the underlying mechanisms.

1. Introduction

The last decade has seen an increased focus on the importance of data as a key asset of organizations. The combination of different technological developments around that type, including rapid growth in processing power, data storage capacity, faster and more far-reaching networks, as well as the development of sensors and devices that can capture and transmit data of different types in real time, gave rise to a new phenomenon in the field of computer science. The term big data emerged as a notion describing a new paradigm of managing and analyzing datasets that are too large or complex to be dealt with by traditional data-processing application software (McAfee et al., 2012a). Academic articles and popular press have argued that by analyzing such vast amounts of data through appropriate techniques, organizations can generate important insight that can inform operational and strategic decision-making. Recognizing the significance of big data and analytics, organizations have been investing massively in acquiring the necessary technological and human resources in order to effectively leverage the power of big data (Mikalef et al., 2018c). In fact, the prevailing view over recent years is that organizations that have not already invested in big data analytics[3] (BDA) will lose their competitive advantage and struggle to keep up with those that have already opted for data-driven decision-making (Wamba and Mishra, 2017).

While there have been some exemplary cases of use and value generation from such technologies (Raguseo, 2018), the literature points out that there are still a considerable number of organizations that have not managed to harness their potential or are still unable or unwilling to adopt BDA due to organizational, technological, and financial restraints (Bag et al., 2021). Similarly, not all organizations that have adopted BDA have managed to realize performance gains, with a sizeable chunk of organizations still struggling to break even from their investments (Côrte-Real et al., 2019). Several studies have shown that many BDA projects fail to make it to production, with a vast majority only reaching a prototype phase (Mikalef et al., 2021). In addition, the value that is realized for BDA projects is heavily dependent on how well they are aligned with an organization's strategy and competitive positioning. As a result, studies have noted that many BDA projects fail to materialize concrete performance results (Akter et al., 2016). With over a decade of academic research in the domain of BDA and business value, it is possible to extract some useful insight into what studies have found in order to guide future initiatives as well as policymaking efforts. The extracted knowledge can help us understand not only how BDA should be setup within organizations, but also what type of performance gains can be realized and through what mechanisms.

In this report, a review of the literature over the last decade is performed to provide an overview of what we currently know concerning BDA adoption from organizations, as well as the value that they can deliver. The first part of the report provides some background information concerning the notion of BDA, distinguishing the notion from other ones that have been prominent in the past, as well as drawing the line with new emerging technologies such as artificial intelligence (AI). An overview of definitions

[3] Henceforth, big data analytics is abbreviated as BDA.

concerning what BDA entail is presented, and key characteristics of these technologies are highlighted. The second part presents a review of empirical evidence of BDA and its potential impacts on firms' productivity and performance. By analyzing the relevant empirical studies, an overview of the effects of BDA is developed, along with the important enablers and inhibitors. Finally, the report concludes with a discussion of the implications of the findings with regard to policymaking for facilitating the adoption of BDA, as well as some important aspects that need to be considered by firms and policy makers to facilitate value realization from such investments.

1.1 Method

To provide a comprehensive and complete overview of the relevant research that has been done on BDA and its effect on firm productivity, this report has followed a systematic review approach. This approach ensures that all the latest academic research and empirical studies on this topic are included in the synthesis of findings. As such, the method that was followed builds on a series of steps defining appropriate keywords and then developing appropriate search strings to identify articles in academic literature databases and meta-search engines. The key words included a technology-focused part, which, for example, included terms such as "big data analytics" and an organizational, or business value part, such as "firm productivity". Different combinations of keywords were combined to create search strings, and for each search string results were pooled in a reference manager software application.

The databases that were used to identify relevant studies included, among others: Scopus, the ACM digital library, IEEE Xplore, as well as publisher databases (e.g., Springer, Willey) and meta-search engines (e.g., Google Scholar). In this way, it was ensured that all important academic studies were identified and included in the pool of papers. More information concerning the details of the selection, filtering, and assessment of articles can be found in Appendix A, which also includes a flow chart of the sequence of activities performed to reach the final sample of 94 studies. The rigorous process of assessing articles based on relevance and quality ensured that the sample of papers represents the most important work on BDA and firm productivity.

The final step in the methodology to extract the main findings from the articles identified involved categorizing them and organizing their content. In this process we followed a structured approach, which distinguished between some important thematical areas. The first step was to provide a coherent and complete definition of BDA and identify how this notion differs from others that have emerged over recent years. This step was important in ensuring that there was consistency in the empirical findings that were synthesized in the following steps. The second step included an analysis of the effects of BDA on firm productivity. In addition, this synthesis examined factors that were mentioned in the academic literature concerning the challenges organizations faced in leveraging BDA for value creation, as well as value-generating mechanisms. The third and final step of the process concerned the identification of any recommendations the studies made on a practical and policy-making level.

Through this approach, we were able to ensure that the most important work that has been undertaken over the past decade was included in the analysis. Moreover, the reporting and synthesis of findings enabled a complete understanding of how BDA is used in contemporary organizations, as well as the challenges and potential value it can deliver.

2. An introduction to big data analytics

<u>Section summary</u>

- Big data analytics represents a distinct technological shift that not only concerns the data types that are leveraged, but also the supporting technologies, people's skills, and the processes of managing BDA projects.
- Big data, with the defining characteristics of the four Vs - volume, velocity, variety and veracity - represent the raw input, whereas big data analytics correspond to the ways such data is sourced and analyzed.
- BDA build on different technologies and ways of analyzing data from other approaches such as artificial intelligence, as the latter focuses on ways in which machines can think, learn, respond, react, and behave in a way similar to that of humans.

While BDA has seen a rapid increase in interest over the past decade, there is still considerable confusion and a lack of clarity regarding what this notion entails and how it differs from prior areas of research interest, such as business intelligence or data mining. The purpose of this introduction is to clearly outline the key differentiation factors that underpin this set of technologies, as well as the elements of the data that have been critical to the emergence of this new research area. To do so, the report starts by surveying how BDA has been defined within academic research, and specifically within the stream of research that examines the business value and productivity effects of the technology. The notion of big data is first delineated to identify the distinguishing features and requirements it imposes on the technological infrastructure and analytical techniques. In sequence, the term analytics is discussed, and the definitions of BDA that have been presented throughout the last years are compared. Finally, the section ends with a discussion on emergent technologies and their relationship to BDA.

2.1 Defining big data

The notion of big data emerged in academic literature on firm productivity in the early 2010s. While the notion itself has been used in the past within the field of computer science, it gained popularity in the context of the business value of technology after a series of publications sparked interest in using data within the organizational context to improve operations and enhance decision-making (McAfee et al., 2012a, McAfee et al., 2012b, Chen et al., 2012). A series of academic publications that proceeded helped define the concept of what BDA is and how it differs from other types of data that organizations typically handle. In this direction, studies such as that of Wamba et al. (2015) have proposed that what makes big data a distinct and novel concept are the unique properties that such datasets encompass. Specifically, these have been defined in terms of the so-called "Vs", with different publications specifying variations of the underlying features. The majority of empirical studies have defined big data in terms of three primary "Vs", which include volume, velocity, and variety (McAfee et al., 2012b). Volume corresponds to the size that such datasets have due to the large set of observations, multiple variables and timestamps, as well as the different types of formats that pose requirements on storage (Gupta and George, 2016). The expansion in the volume of data over the past decade can be largely attributed to a number of factors, such as the prevalence of smaller and cheaper sensors and devices that can transmit rich data, a phenomenon termed Internet-of-Things (IoT), the rapid uptake of social media platforms and mobile devices, and the increased bandwidth of networks for different types of communication (Sun et al., 2015, Ekbia et al., 2015). Velocity, on the other hand, refers to the pace at which data needs to be collected, analyzed, and utilized (Sagiroglu and Sinanc, 2013). The rapid speed in which data is sourced and leveraged also poses the requirement that it is utilized within a much shorter period than previously required (Al-Jarrah et al., 2015). Since much of the use of big data is based on decision-making, operation management, and other real-time applications, the concept of velocity becomes increasingly important as a quality of the data and its ephemerality (Akter et al., 2016). Finally, variety underscores the plethora of sources from which data emerge, which can include, among others, formats such as text, audio, images, video, networks, and graphics (Constantiou and Kallinikos, 2015). As noted previously, the convergence of several core technological developments have also enabled a plurality in the collection and storage of different data formats. Thus, the concept of variety encapsulates both the large range of different formats that fall within the big data notion, as well as their concurrent utilization to derive meaningful uses (Kar and Dwivedi, 2020).

While the three "Vs" have been a central aspect of defining and differentiating big data from other conventional forms of data that have been used in the past, the literature has also extended the differentiating elements in order to further distinguish this phenomenon from prior forms and structures of data (Dremel et al., 2018, Sivarajah et al., 2017). Several later studies have sought to extend the defining "Vs" to four, five, seven or even more attributes that should characterize such datasets (Mikalef et al., 2019d). For example, the review article by Wamba et al. (2015) finds that, in addition to the three fundamental attributes of big data, several studies also include aspects such as value and veracity, whereas other studies have expanded these to also include variability and visualization (Uddin and Gupta, 2014). While there is considerable discussion regarding whether the extended list of "Vs" constitute the inherent characteristics of the datasets themselves, there is a growing consensus concerning the core defining characteristics. In addition, a differentiating element to separate big data from conventional datasets concerns the set of infrastructure, software, processes, and roles, among others, that surround such datasets. In providing an illustration of how big data has shifted the technological landscape, Abbasi et al. (2016) highlighted the requirements that big data impose on the technological and organizational landscape of contemporary enterprises. Some examples of underlying technologies that are needed for supporting big data include new forms of data storage and analysis of data, such as NoSQL databases, massive parallel processing, as well as cloud computing infrastructures (Corbellini et al., 2017). Taken together, the defining characteristics of big data combined with the technological and organizational requirements they create help define the transition into the era of big data, which separates such datasets from conventional forms of data.

The list of definitions that have been used in the academic literature touches upon elements of both, with some of the definitions explicitly mentioning some of the “Vs” that characterize big data, whereas others remain quite generic, not identifying specific aspects of the data. The table below presents a sample of some of the definitions that have appeared over the years, which indicates not only the large diversity in terms of the way big data have been defined, but also an evolution of the notion to differentiate big data from conventional datasets, separating them from the techniques and methods used to handle and analyze them. While there is considerable diversity in the definitions, an analysis indicates that where big data tends to be consistently different from other forms of data is around its volume, velocity, and variety, and these characteristics impose new technological requirements on organizations for their collection and management. Thus, it is also important to highlight at this point that big data constitutes the raw input to a process of insight extraction and knowledge-generation (Mikalef et al., 2020c). This process, along with the techniques required to actualize this, is described in the following subsection where we focus on the term BDA.

Table 1 Selected definitions of big data

Reference	Definition
Russom (2011)	Big data involves the data storage, management, analysis, and visualization of very large and complex datasets
McAfee et al. (2012b)	Big data, like analytics before it, seeks to glean intelligence from data and translate that into business advantage. However, there are three key differences: Velocity, variety, volume
Gantz and Reinsel (2012)	Big data focuses on three main characteristics: the data itself, the analytics of the data, and presentation of the results of the analytics that allow the creation of business value in terms of new products or services
Bharadwaj et al. (2013)	Big data refers to datasets with sizes beyond the ability of common software tools to capture, curate, manage, and process the data within a specified elapsed time
Sun et al. (2015)	Big data: the data-sets from heterogeneous and autonomous resources, with diversity in dimensions, complex and dynamic relationships, by size that is beyond the capacity of conventional processes or tools to effectively capture, store, manage, analyze, and exploit them
Constantiou and Kallinikos (2015)	Big data often represents miscellaneous records of the whereabouts of large and shifting online crowds. It is frequently agnostic, in the sense of being produced for generic purposes or purposes different from those sought by big data crunching. It is based on varying formats and modes of communication (e.g. text, image, and sound), raising severe problems of semiotic translation and meaning compatibility. Big data is commonly deployed to refer to large data volumes generated and made available on the Internet and the current digital media ecosystems
Akter et al. (2016)	Big data is defined in terms of five ‘Vs:’ volume, velocity, variety, veracity, and value. ‘Volume’ refers to the quantities of big data, which are increasing exponentially. ‘Velocity’ is the speed of data collection, processing and analyzing in the real time. ‘Variety’ refers to the different types of data collected in big data environments. ‘Veracity’ represents the reliability of data sources. Finally, ‘value’ represents the transactional, strategic, and informational benefits of big data

Reference	Definition
Coble et al. (2018)	Big data is defined by several characteristics beyond size, particularly, the volume, velocity, variety, and veracity of the data
Obschonka and Audretsch (2020)	Big data is defined as a large volume of structured, semi-structured, or unstructured data, and a way to collect/produce, process, and analyze these datasets using non-traditional methods

Source: Author

2.2 From raw data to big data analytics

While early academic and user reports on big data often used the term interchangeably with that of BDA, there is a growing consensus that the later represents the tools and techniques used to generate actionable insight from big data (Mikalef et al., 2018c). This distinction was further clarified in the study by Gupta and George (2016), who noted that while big data refers to the large and varied datasets that can be collected and managed, BDA involved the models that can enable us to capture, visualize, and analyze the underlying patterns within this data to provide actionable insight. In effect, BDA encompasses a broader set of aspects, which places a stronger emphasis on the process and tools used in order to extract insights from big data. As a result, BDA includes the data upon which analytics are performed, as well as the tools, infrastructure, and means of visualizing and presenting insight. The definitions that are presented in the table below, particularly those of Lamba and Dubey (2015), highlight this distinction.

Table 2 Selected definitions of big data analytics

Reference	Definition
Mikalef et al. (2018c)	A new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high velocity capture, discovery and/or analysis
Lamba and Dubey (2015)	The application of multiple analytic methods that address the diversity of big data to provide actionable descriptive, predictive, and prescriptive results
Müller et al. (2016)	The statistical modeling of large, diverse, and dynamic datasets of user-generated content and digital traces
Loebbecke and Picot (2015)	A means to analyze and interpret any kind of digital information. Technical and analytical advancements in BDA, which—in large part—determine the functional scope of today’s digital products and services, are crucial for the development of sophisticated artificial intelligence, cognitive computing capabilities, and business intelligence
Vassakis et al. (2018)	Advanced analytic techniques, considering large and various types of datasets to examine and extract knowledge from big data, constituting a sub-process in gaining insights from big data process
Shehab et al. (2021)	The application of advanced analytic techniques including data mining, statistical analysis, predictive analytics

Source: Author

The notion of BDA effectively builds on prior approaches of leveraging data to extract important insights, such as business analytics, data mining, and business intelligence (Hindle et al. 2019). The key distinction that separates BDA from other relevant terms is that it relies on big data, as described earlier, and builds on analytics techniques, methods and technologies that allow for the handling and processing of big data sets (Mikalef et al., 2019a). The underlying technological infrastructure that is needed to support BDA has been described in detail in recent academic research (Talia, 2013). In their study, Abbasi et al. (2016) note that BDA entail changes along the information value chain through the process by which raw data is transformed into insight that can be utilized. Specifically, they note that BDA requires novel technological solutions such as data lakes, in-memory databases, complex event systems, prescriptive analytics tools, as well as NoSQ systems among others. Other articles have centered on the specific models for analyzing big data, such as map reduction, or even techniques such as association rule learning, sentiment analytics, and cluster analytics, among others (Elgendy and Elragal 2014).

A key characteristic of the term BDA is that it is technology and method independent. In other words, there is no specific type of analytics that qualifies as BDA since the term encompasses many different types of analysis and visualization that can be applied to big data (Mikalef et al., 2020c). What the notion of BDA underscores, however, is the process of converting raw data into meaningful representations that can be utilized in a value-adding way. The difference between the notions of big data and BDA has been highlighted in several subsequent studies in order to ground the concepts and enable clarity concerning their potential value to organizations (Mikalef et al., 2018c, Wamba et al., 2015). Extending on this differentiation, in the next section of this report we specifically examine how BDA has been used in organizations, as well as the resulting business value that has been realized. Prior to delving into the relevant empirical findings, we also present a distinction between the notion of BDA with recently emerging buzzwords such as artificial intelligence and generative AI. The purpose of doing so is to clearly separate how these sets of technologies differ or overlap with BDA, and what type of value-generation can be expected from each.

2.3 A distinction between big data analytics and artificial intelligence

An emerging discussion among the research community has been on trying to define the difference between BDA and emerging technologies such as those of artificial intelligence (AI). While there are different and constantly changing definitions of AI, the consensus is that it can be broadly defined as a set of methods and approaches that enable machines to demonstrate intelligence and human-like behavior in terms of performing tasks, solving complex problems, reasoning, and predicting events. An overview of current definitions is presented in the work by Enholm et al. (2021). As with the notion of BDA, the definition of AI also encompasses a wide set of technologies that have been described and categorized in detail in relevant studies such as that of Ertel (2018). A large set of technologies that support AI have been grounded on the cluster of machine learning applications, which follow the principle of problem-solving using data and algorithms to imitate the way humans learn and detect patterns in order to gradually improve accuracy (Collins et al., 2021).

While there is undoubtedly considerable overlap in the types of activities that are connected to BDA and AI, there are also some important differences that separate the two concepts. The key differentiating aspect is that in BDA, the objective is to leverage available big data sets in order to analyze it and derive important insight, whereas AI utilizes this data to learn from it and adjust its output (McAfee and Brynjolfsson, 2017). Thus, the aspect that separates the two concepts is the end goal despite the fact that both heavily rely on big data and use complex methods for analyzing such data. The goal of BDA is to make sense of raw data for predictions, decision-making, and optimizations of problems among others, whereas AI is concerned with the development of applications that exhibit intelligence behavior, and can understand, think, learn, response, react, and behave in a way similar to that of humans (Mikalef and Gupta, 2021). In the next chapter, we examine studies that have looked into the productivity and performance gains of firms that have adopted BDA, and draw our definition of the notion based on the previously mentioned differentiation with AI.

3. Big data analytics and productivity

Section summary

- Studies show that a structured adoption of big data analytics in certain industries (e.g., technology and manufacturing) can result in a 3-7% firm productivity improvement annually.
- The value of BDA is contingent on the industry it is applied in, and the types of operations it is targeted towards.
- Most impacts are indirect and mediated through changes in organizational processes (e.g., process efficiency, product/service innovation, customer engagement/retention).
- Lag effects of 1 to 3 years to realize performance gains characterize BDA projects.
- To realize performance from BDA investments, organizations must develop capabilities that require attention to other complementary resources (e.g., humans and skills, governance practices, cross-department collaboration, opportunities of experimentation).

A key area of focus for research over the past decade has been to understand if BDA can result in productivity gains for firms, and if yes, through what mechanisms and in which forms. Thus, studies have examined several complementary aspects of BDA and value generation and can be broadly categorized as having three objectives. The first objective of research publications has been to examine if there is a causal relationship between the adoption of BDA in firm operations and improvements in productivity or performance. The second objective has been to delve into the mechanisms through which such productivity gains can be realized, as well as to understand more about the shifts BDA introduce concerning the way organizations are structured and compete. Finally, the third objective of research has been to identify some key challenges, obstacles, and common issues organizations face when trying to leverage BDA for productivity gains. Following this broad categorization of literature, this section is also structured accordingly into three sub-sections.

3.1 What effect does BDA have on firm productivity?

One of the primary questions academic research has sought to explore is if there is any causal relationship between the adoption and use of BDA and firm productivity (Akter et al., 2016). Within this stream of literature, researchers have been divided into two approaches in uncovering such effects and establishing causal associations. The two perspectives differ with regard to the unit of analysis they examine, with the former adopting an organizational-centric perspective, whereas the latter focuses on industry-level effects. In other words, studies have been divided into those that examine, on the one hand, how organizations are structured and what effect that has on their operational performance, and those that examine the macro-level effects of adoption rates of BDA and different measures of productivity. The main differentiating aspect in these types of studies concerns the types of data that are used to uncover effects.

The majority of academic articles that examine the impact of BDA on firms' performance build on self-reported primary data, which focus on the firm as the unit of analysis (Popovič et al., 2018). Typically, these studies utilize samples of at least 150 responses from the same industry to examine if the adoption of BDA results in any significant performance gains. As in most of these studies, performance is assessed in a self-reported manner, performance indicators usually include constructs such as a firm's competitive positioning within its industry (Mikalef et al., 2020b), operational efficiency gains (Mikalef et al., 2019c), financial performance (Ghasemaghaei, 2019), customer satisfaction and market share (Hassna and Lowry, 2018), as well as indicators measuring innovation outcomes (Mikalef and Krogstie, 2020, Mikalef et al., 2020a). While some of these studies also utilize secondary data to validate findings and triangulate results, their primary orientation is to explain what organizations should do with BDA and how they should be structured in order to realize productivity gains. Although there is a potentially higher bias in studies that use self-reported data due to overestimation, many of these papers also build on objective data to triangulate findings and control for any overestimation effects.

A much smaller stream of research has sought to explore the issue of BDA investments and productivity on an industry level of analysis, rather than on a firm level. Within this body of work, researchers have studied secondary data in order to understand if investments in BDA within particular industries result in statistically significant productivity gains. Among such studies, there has been a specific focus on economic-related indicators of productivity (Müller et al., 2018). Such econometric-based approaches can accurately highlight the potential return that BDA investments can have on firm economic indicators. Nevertheless, they have the limitation that they do not provide much insight into how BDA investments are leveraged, and for what purposes they are used.

Taken together, these two different approaches of research in exploring the effect that BDA has on firm productivity can allow us to understand if they do indeed result in business value, and if so through what organizational arrangements. This latter part is particularly important since it enables us to identify what key organizational resources need to be invested in for firms to be able to effectively generate value from BDA. It also provides a roadmap for policy recommendations and an organizational implementation of BDA. Before examining the specific findings of the extant research, we draw a differentiation between the different types of indicators that have been used to examine performance effects from adoption of BDA.

Table 3 Sample studies with performance indicators used

Indicator type	Performance indicator	Indicative studies
Objective data	Financial performance	(Müller et al., 2018)
	Firm productivity	(Müller et al., 2018)
	Innovation performance	(Niebel et al., 2019)
	New product development	(Niebel et al., 2019)
	Operational efficiency	(Zhu et al., 2021)

Indicator type	Performance indicator	Indicative studies
	Business growth	(Zhu et al., 2021)
Subjective data	Competitive performance	(Mikalef et al., 2019a, Mikalef et al., 2020b)
	Organizational performance	(Elia et al., 2022, Wamba et al., 2017)
	Business process improvement	(Elia et al., 2022)
	Innovation performance	(Elia et al., 2022, Mikalef et al., 2019b, Mikalef et al., 2018a)
	Financial performance	(Raguseo et al., 2020, Raguseo and Vitari, 2018)
	Market performance	(Saenz et al., 2022)

Source: Author

As shown in the table above, most of the empirical evidence around the productivity effects of BDA comes from studies that use subjective and self-reported indicators to measure performance. Nevertheless, from the limited number of studies that use objective data, there is a clear causal link between the adoption of BDA and important productivity indicators. For example, in the study of Müller et al. (2018), the authors examined a sample of 814 firms during the period of 2008-2014 and found that investments in BDA assets were associated with an average improvement of 3-7 percent in yearly firm productivity. The effects of BDA were more pronounced in industries that are more technologically oriented. Similar effects were also found for highly competitive industries as per the authors, whereas no significant effects were observed in noncompetitive industries. In another study, which focused on German firms, Niebel et al. (2019) found that investments in BDA were associated with an increased likelihood of a firm becoming a product innovator as well as the market success of product innovations. Their results were found to hold true for the manufacturing sector as well as the service sector but were also contingent on the degree to which these firms had invested in IT-specific skills. Their results provide support that BDA are a source of improved firm productivity by enabling the innovation potential.

Apart from these studies, other studies have been built on subjective data to gain deeper insight into the structuring of firms' BDA portfolio and how it affects productivity indicators. Typically, such studies build on primary data collected by key respondents in firms and conduct analyses on samples of between 200-800 firms. In this direction, there have been a plethora of articles that show that a structured adoption of BDA can result in improved competitive performance, which is measured as the perceived advantage a focal firm has over its main competitors. In fact, such studies show that BDA adoption can result in an improved competitive performance of 10-25%. Likewise, other studies have shown that there is a positive and significant effect between the adoption of BDA and innovation outcomes. In this direction, studies have found support that BDA can facilitate product, process, and service innovation outcomes (Mikalef and Krogstie, 2020, Mikalef et al., 2019b, Lehrer et al., 2018, Wamba, 2017, Tan et al., 2015). Finally, another group of studies have examined aspects such as business process performance and market performance in an attempt to measure the impacts that BDA have in making firms more agile and fast to respond. These studies find that investments in BDA can facilitate such operational efficiency providing they are adopted and deployed as part of an organizational strategy (Gupta et al., 2020, Akter and Wamba, 2016, Hassna and Lowry, 2018, Dezi et al., 2018).

Overall, the findings from the two main streams of research provide ample support for the claim that BDA can enable productivity gains for firms. Nevertheless, there are some key contingencies in the findings that need to be highlighted. First, the value of BDA is higher in contexts where there is a high degree of competitiveness and market uncertainty (Müller et al., 2018, Mikalef et al., 2020b). Second, the value of BDA is dependent on a structured adoption within organizational boundaries. In other words, it is necessary for firms to not only invest in the respective technological infrastructure, but also to place weight on the human factor, as well as on governance of such technologies in a way that is in alignment with organizational strategies (Niebel et al., 2019, Mikalef and Krogstie, 2020). This highlights the importance of investing in complementary skills and assets to realize performance gains from BDA. Third, the value of BDA on firm productivity is not always realized in a direct manner. In other words, BDA can prompt significant changes in the operational or competitive strategies that organizations pursue. As a result, it is important to consider what types of uses BDA have within firms, and in what operational domains they are leveraged. Thus, to gain a more nuanced understanding of the performance effects of BDA, it is critical to take a closer look at mechanisms of BDA effects before examining productivity indicators, which may be dependent on other exogenous variables (Mikalef et al., 2017b). While a significant proportion of studies assume a direct effect between adoption of BDA and firm performance gains, a growing body of research now examines indirect effects, which are mediated through new or enhanced business operations. The next sub-section elaborates on these mediating conditions.

3.2 How are effects of BDA on firm productivity measured?

While the body of research that examines the impact of BDA adoption on the productivity of firms has some consistent findings in terms of the magnitude and effect, it does build on a multitude of methods and approaches to explore this relationship. One difference that was described earlier concerns the choice of indicator types (objective vs. subjective) for quantifying investment propensity and productivity effects. Nevertheless, within the body of research, there are also significant methodological differences in how such effects are modeled, and the mechanisms through which they are realized. Thus, while most empirical studies have built on subjective self-reported data adopting a firm-level analysis, there have been other empirical articles using longitudinal and case-study approaches that can provide more contextual information on what applications BDA have within organizational boundaries and how they are leveraged.

As noted previously, as most empirical studies using a quantitative approach have been built on self-reported data, they have managed to capture the mechanisms through which BDA investments result in productivity gains. This has been a limitation that conventional econometric studies have been unable to capture in detail. From these studies, we know that the effects of BDA on firm productivity is not direct, but rather of an indirect nature (Mikalef et al., 2018c). In other words, it is important to understand how firms leverage such investments and for what purposes they are used in order to more accurately capture their effects (Côrte-Real et al., 2020). This outcome also has some important implications concerning the way that performance effects of BDA are measured, since it highlights that apart from economic indicators, studies should also try to capture operational indicators (e.g., market share, operational efficiency, new products or services, business growth). A key finding noted in these studies is that BDA can be used to effectively sense changing conditions, whether these are customer beliefs and requirements, emerging opportunities and threats, or competitive actions and their consequences (Mikalef et al., 2020b). Thus, BDA provides a strategic tool for firms to scan the competitive environment and make data-driven decisions.

A second important mechanism of value-generation through BDA investments is by monitoring and optimizing internal operations (Dezi et al., 2018). Several studies have highlighted the use of BDA for such purposes, which include predictive maintenance (Gupta et al., 2020), business process design and reconfiguration (Braganza et al., 2017), supply chain management optimization (Gunasekaran et al., 2017), and operational monitoring and efficiency (Mikalef et al., 2020b). One of the key findings in these studies is that the embeddedness of many physical devices with sensors that can generate and collect data can allow firms to utilize that for the optimization of processes and closer monitoring of them. These are particularly pronounced in certain industries such as manufacturing, logistics, processing of raw materials, oil and gas, as well as transportation. As a result, the mechanisms of value-generation as well as types of productivity gains that can be expected depend both on the industry as well as the use cases of BDA (Mikalef et al., 2020c).

Another important finding from qualitative articles concerns the process through which BDA are leveraged and effectively used in production. These studies underscore that there are typically long lag effects from when firms invest in BDA to when they can realize productivity gains (Mikalef et al., 2017a). These lag effects are typically within the span from one to three years (Zhu et al., 2021, Ferraris et al., 2019). It is a common practice among firms to launch many BDA initiatives in a sandbox environment, where experimentation is allowed, and there is an assessment of the different dimensions (e.g., access to data, technical skills, technology infrastructure) that need to be matured before applications can be launched (Watson, 2014). Furthermore, some studies have incorporated dimensions of performance that do not typically fall within the monetary productivity indicators, such as waste reduction, environmental performance, and other sustainability goals that BDA can help achieve (Kristoffersen et al., 2021, Kristoffersen et al., 2020).

From the above studies, there are some key findings that can be extracted. First, the productivity indicators related to BDA adoption are contingent upon industry and application uses. It is therefore important to consider that different metrics may be needed to accurately capture the effects in dissimilar industries or contexts. Second, many of the effects of BDA are indirect and dependent on what the focal firm does with the derived insight or new knowledge. Thus, the value of BDA is elusive without considering that it is dependent on how it is acted upon from the relevant managers. Third, there are long lag effects from the time of the adoption of BDA to the actual value generation. As such, these lag effects should be accounted for when trying to identify causal associations. Fourth, there are several other important performance indicators that are conventionally not measured when examining the value of technology. As BDA can have a significant impact on such indicators, including those of environmental performance and sustainability, it is important to account for how it affects these measures.

3.3 Firm challenges in leveraging data analytics?

A large proportion of the research studies on firm productivity from BDA has focused on the issue of the challenges organizations face when trying to realize value from such investments. In this direction, there have been a number of factors that are either technology-, human-, or organizationally-related that have been identified as obstacles for firms attempting to realize value from their BDA investments (Dremel, 2017, Gupta and George, 2016). Since these are challenges related to different aspects of how organizations are set up in order to leverage technology, researchers have coined the term "BDA capability" to denote the organizational capacity to effectively leverage these technologies in a way that drives productivity gains for the firm. Central in this area of research has been the study of Gupta and George (2016), which creates a distinction between tangible (or technology-related), human, and intangible (or organizational) aspects that can drive or hinder value creation. Such an approach to studying the entire organizational capability to effectively use BDA, rather than just the technology in isolation, has been a common practice among recent empirical studies (Bag et al., 2021, Mikalef et al., 2021, Mikalef et al., 2020b, Arunachalam et al., 2018, Wamba et al., 2017).

From a technological point of view, one of the key challenges that organizations face when it comes to leveraging analytics is having access and being able to manage big data sets. This is due to the novel technological infrastructure, large costs associated with storing and processing such data, as well as new processes for filtering, cleansing, and preparing data so that they are ready for analysis (Surbakti et al., 2019). In particular, there are several empirical studies that have examined important barriers for organizations in the BDA paradigm, or obstacles hindering them from being able to create value from their data (Alharthi et al., 2017). In these articles, there are recurrent issues raised on the high costs that are entailed in initiating BDA projects, particularly in relation to uncertain outcomes of their business value (Mikalef et al., 2017a).

Another key hindrance for many organizations is the ability to either hire or train employees with the necessary skill-set to create value from big data (De Mauro et al., 2016). These skills are primarily related the technical implementation of BDA projects, from data engineers, data architects and data scientists, who are responsible for the collection and management of the datasets, to their actual analysis (Costa and Santos, 2017). Furthermore, several studies have highlighted the fact that many firms face challenges in realizing productivity gains due to the fact that there is a lack of management bandwidth that can create a clear roadmap and strategy for how to leverage BDA in operations in a way that generates value for the firm (Mikalef et al., 2018b). For both technical and managerial skills, studies have noted that firms typically experience difficulties in recruiting employees that have up-to-date knowledge (Costa and Santos, 2017).

Finally, another category of challenges in realizing value from BDA has been attributed to a lack of organizational structures and cultures that support data-driven decision making (Gupta and George, 2016). A large proportion of papers report that it is necessary for organizations to perceive BDA as a key strategic initiative that permeates the entire organization, rather than just a technical task from the IT department (Dykes, 2017). In this way, any BDA project that is launched will be strategically aligned with the rest of the activities of the organization, which also facilitates the dissolution of data silos that may exist. As such, approaching BDA as a firm-wide activity ensures that there is transparency and access to key data sources within the organization, and that any launched projects are strategically aligned (Wiener et al., 2020).

4. Policy implications

Section summary

- Incentives and education programs for training employees on the job, plus curricula for students on data science in industry can improve BDA uptake.
- Technological infrastructure, access to data, and financing big data projects are key barriers, particularly for SMEs.
- Uncertainty about data handling and privacy/security concerns impede the use of BDA.
- When assessing the value of big data analytics investments, it is important to study their effects on organizational operations, and industry-specific applications.
- There are no empirical studies conducted in Sweden, so it is important that future studies examine challenges and value-generation paths that are specific to the country.

The synthesis of past empirical research has shown that BDA can result in important productivity gains under a certain set of contingencies. There are undoubtedly some industries or categories of firms that can gain much more from deploying BDA than others (e.g., manufacturing, and high-tech), but there are still a common set of challenges that organizations face when adopting and leveraging BDA. These challenges and obstacles for adoption can serve as important input in providing policy-making recommendations, particularly since they are noted recurrently in academic studies with large sample sizes. For this report, they are presented in three categories that cover complementary aspects of BDA adoption; these include: knowledge and skills, organizational level barriers, and industry-level facilitators.

4.1 Knowledge and skills of individuals

One of the primary barriers for many organizations that plan to adopt BDA is the difficulty in finding employees with the right skill set. Several empirical studies have specifically focused on the issue of knowledge and skills that are needed in the age of BDA, also highlighting some areas where important gaps are currently observed. For instance, the study of Mikalef and Krogstie (2019) looked at the gap that exists compared to what Norwegian computer science graduates receive as education in data science, compared to industry needs. They find that firms require employees with technical, managerial, and soft skills related to data science, and that there is insufficient training in academic institutions on specific tools and technologies that are currently used in industry, as well as on collaborative and cross-domain skills that allow recent graduates to work effectively on big data analytics projects. In addition, the study provided an ordering of the perceived importance firms placed on these skills, which enable them to adopt and use BDA in a competitive way. What was highlighted by this study is that it is important that firms have the sufficient resources and financing required to retrain their personnel with these new skills through flexible learning settings and on-the-job training. Thus, there is a need for further education incentives, as well as a development of study programs at the higher-education level that are in close collaboration with industry requirements. This has been found to be a strong impediment to adopting BDA, particularly for small and medium-sized enterprises (SMEs) that have limited slack and financial resources and a smaller employee base.

4.2 Organizational level enabling conditions

Another important finding that emerged from the analysis of the empirical studies is the unequal adoption of BDA between large firms and SMEs. This outcome has been attributed to the high cost associated with adopting BDA, which many SMEs struggle with (Mikalef et al., 2017a). In particular, purchasing technological solutions for data storage and processing requires considerable investments, which many SMEs are hesitant to make due to unclear benefits from adopting BDA (Iqbal et al., 2018). This has been particularly the case for organizations that operate in non-technology-focused industries that have been lagging significantly in their adoption rates of BDA (Maroufkhani et al., 2023). Thus, incentives that focus on support in providing technological infrastructure to organizations that belong to the SME size-class and in industries that can benefit from BDA have been argued to increase the competitiveness of such firms and their overall productivity (Müller et al., 2018).

One area that was also noted by organizations in the European context, which is particularly challenging to navigate when adopting BDA, was that of data governance and ownership. This issue was highlighted due to the regulatory framework around the General Data Protection Regulation (GDPR)[4] and the Data Act[5], among others. Since many service providers of BDA solutions offered cloud-based solutions, it had been noted that it was not clear for organizations that were handling personal or sensitive data how to implement effectively the appropriate governance structures that align with these regulations. This was particularly the case for organizations that did not have extensive experience in dealing with technology-based initiatives, such as that of BDA, where there was a lack of technical and regulatory knowledge (Kempeneer, 2021). As a result, some recommendations that can be implemented to alleviate this barrier from firms is to establish national guidelines and easy-to-access frameworks with important information on the implementation of data governance and vendor selection.

Finally, a key finding of the analysis of empirical studies highlighted both the indirect nature through which BDA result in value generation for firms, as well as the lag effects that they entail. This outcome highlights two points that need to be considered when evaluating the value of BDA for firms. The first is that performance gains may be quite elusive to capture and require a broader set of indicators to accurately capture BDA effects, which are also contingent upon the industry in which they are used. The second is that BDA investments require considerable time to mature and produce value for firms. This finding underscores the need for providing incentives to firms that give them sufficient time to develop BDA into a core organizational capability. In addition, the

[4] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679
[5] https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113

presence of long lag effects also shows that any assessment of performance gains from such incentives needs to consider the time needed to have such results materialize.

4.3 Industry level facilitators

As noted in the discussion of findings, there are certain industries that are forerunners in the adoption of BDA; at the same time, there are some others that are significantly lagging (Baig et al., 2019). This phenomenon has been identified in several countries and has been linked to the competitiveness and velocity of industries in driving the digitalization of operations. However, there are other factors apart from industry dynamics that have resulted in certain firms either not adopting or being at a very early stage of BDA adoption. These concern the lack of financing schemes and incentives to digitally transform operations, which have also been found to be important determinants of BDA adoption in empirical studies (Maroufkhani et al., 2020). Such financial incentives and stimulus provisions can enable firms to perform the first step in sourcing the necessary technological infrastructure and know-how to leverage their data, and build the technical infrastructure needed to perform analytics to improve productivity.

These measures have been highlighted as important to increase firms use of novel digital technologies and be more competitive in the global market (Manyika et al., 2011). The focus on technological infrastructure incentives for SMEs is crucial, given their significant presence in European business. Developing tailored incentive schemes and financing options is vital to facilitate their investment in technologies like BDA. Many of these SMEs may not yet fully recognize the value of technologies such as BDA, emphasizing the need for policies that promote the adoption and experimentation of digital tools. Additionally, educational and training programs are instrumental in empowering SMEs to enhance their employees' skill sets, making them better equipped to leverage these technologies. Therefore, policies supporting the financing of technological infrastructure should be complemented with incentives aimed at enhancing the skills of employees across various positions for effective use and value generation from BDA.

Finally, an important aspect that emerged from the analysis of studies concerns the limited availability of data on a country level, as well as cross-comparisons between different countries. This issue was especially evident in the case of data from Sweden, where there were no large-scale empirical findings on the productivity gains from BDA, nor the level of adoption and challenges that firms may face. Having such data can enable a more accurate assessment of the strengths and weaknesses of specific sectors, as well as the areas that need support. In addition, including more detailed indicators of productivity gains can enable a more detailed understanding of the types of applications that drive value-creation for firms in different industries, as well as the specific types of gains that can be expected. In this direction, it is important that there are future follow-up studies that utilize a more comprehensive set of variables around BDA investments and other complementary resources, as well as different types of performance indicators.

Referenser

ABBASI, A., SARKER, S. & CHIANG, R. H. 2016. Big data research in information systems: Toward an inclusive research agenda. Journal of the association for information systems, 17, 3.

AKTER, S. & WAMBA, S. F. 2016. Big data analytics in E-commerce: a systematic review and agenda for future research. Electronic Markets, 26, 173-194.

AKTER, S., WAMBA, S. F., GUNASEKARAN, A., DUBEY, R. & CHILDE, S. J. 2016. How to improve firm performance using big data analytics capability and business strategy alignment? International Journal of Production Economics, 182, 113-131.

AL-JARRAH, O. Y., YOO, P. D., MUHAIDAT, S., KARAGIANNIDIS, G. K. & TAHA, K. 2015. Efficient machine learning for big data: A review. Big Data Research, 2, 87-93.

ALHARTHI, A., KROTOV, V. & BOWMAN, M. 2017. Addressing barriers to big data. Business Horizons, 60, 285-292.

ARUNACHALAM, D., KUMAR, N. & KAWALEK, J. P. 2018. Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. Transportation Research Part E: Logistics and Transportation Review, 114, 416-436.

BAG, S., PRETORIUS, J. H. C., GUPTA, S. & DWIVEDI, Y. K. 2021. Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities. Technological Forecasting and Social Change, 163, 120420.

BAIG, M. I., SHUIB, L. & YADEGARIDEHKORDI, E. 2019. Big data adoption: State of the art and research challenges. Information Processing & Management, 56, 102095.

BHARADWAJ, A., EL SAWY, O. A., PAVLOU, P. A. & VENKATRAMAN, N. V. 2013. Digital business strategy: toward a next generation of insights. MIS quarterly, 37, 471-482.

BRAGANZA, A., BROOKS, L., NEPELSKI, D., ALI, M. & MORO, R. 2017. Resource management in big data initiatives: Processes and dynamic capabilities. Journal of Business Research, 70, 328-337.

CHEN, H., CHIANG, R. H. & STOREY, V. C. 2012. Business intelligence and analytics: From big data to big impact. MIS quarterly, 36, 1165-1188.

COBLE, K. H., MISHRA, A. K., FERRELL, S. & GRIFFIN, T. 2018. Big data in agriculture: A challenge for the future. Applied Economic Perspectives and Policy, 40, 79-96.

COLLINS, C., DENNEHY, D., CONBOY, K. & MIKALEF, P. 2021. Artificial intelligence in information systems research: A systematic literature review and research agenda. International Journal of Information Management, 60, 102383.

CONSTANTIOU, I. D. & KALLINIKOS, J. 2015. New games, new rules: big data and the changing context of strategy. Journal of Information Technology, 30, 44-57.

CORBELLINI, A., MATEOS, C., ZUNINO, A., GODOY, D. & SCHIAFFINO, S. 2017. Persisting big-data: The NoSQL landscape. Information Systems, 63, 1-23.

CÔRTE-REAL, N., RUIVO, P. & OLIVEIRA, T. 2020. Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? Information & Management, 57, 103141.

CÔRTE-REAL, N., RUIVO, P., OLIVEIRA, T. & POPOVIČ, A. 2019. Unlocking the drivers of big data analytics value in firms. Journal of Business Research, 97, 160-173.

COSTA, C. & SANTOS, M. Y. 2017. The data scientist profile and its representativeness in the European e-Competence framework and the skills framework for the information age. International Journal of Information Management, 37, 726-734.

DE MAURO, A., GRECO, M., GRIMALDI, M. & NOBILI, G. 2016. Beyond Data Scientists: a Review of Big Data Skills and Job Families. Proceedings of IFKAD, 1844-1857.

DEZI, L., SANTORO, G., GABTENI, H. & PELLICELLI, A. C. 2018. The role of big data in shaping ambidextrous business process management: Case studies from the service industry. Business Process Management Journal, 24, 1163-1175.

DREMEL, C. 2017. Barriers to the adoption of big data analytics in the automotive sector.

DREMEL, C., HERTERICH, M. M., WULF, J. & VOM BROCKE, J. 2018. Actualizing big data analytics affordances: A revelatory case study. Information & Management.

DYKES, B. 2017. Creating A Data-Driven Culture: Why Leading By Example Is Essential [Online]. Forbes. Available: https://www.forbes.com/sites/brentdykes/2017/10/26/creating-a-data-driven-culture-why-leading-by-example-is-essential/#3c3a52806737 [Accessed].

EKBIA, H., MATTIOLI, M., KOUPER, I., ARAVE, G., GHAZINEJAD, A., BOWMAN, T., SURI, V. R., TSOU, A., WEINGART, S. & SUGIMOTO, C. R. 2015. Big data, bigger dilemmas: A critical review. Journal of the Association for Information Science and Technology, 66, 1523-1545.

ELGENDY, N. & ELRAGAL, A. Big data analytics: a literature review paper. Advances in Data Mining. Applications and Theoretical Aspects: 14th Industrial Conference, ICDM 2014, St. Petersburg, Russia, July 16-20, 2014. Proceedings 14, 2014. Springer, 214-227.

ELIA, G., RAGUSEO, E., SOLAZZO, G. & PIGNI, F. 2022. Strategic business value from big data analytics: An empirical analysis of the mediating effects of value creation mechanisms. Information & Management, 59, 103701.

ENHOLM, I. M., PAPAGIANNIDIS, E., MIKALEF, P. & KROGSTIE, J. 2021. Artificial intelligence and business value: a literature review. Information Systems Frontiers, 1-26.

ERTEL, W. 2018. Introduction to artificial intelligence, Springer.

FERRARIS, A., MAZZOLENI, A., DEVALLE, A. & COUTURIER, J. 2019. Big data analytics capabilities and knowledge management: impact on firm performance. Management Decision, 57, 1923-1936.

GANTZ, J. & REINSEL, D. 2012. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. IDC iView: IDC Analyze the future, 2007, 1-16.

GHASEMAGHAEI, M. 2019. Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics. International Journal of Information Management, 102055.

GUNASEKARAN, A., PAPADOPOULOS, T., DUBEY, R., WAMBA, S. F., CHILDE, S. J., HAZEN, B. & AKTER, S. J. J. O. B. R. 2017. Big data and predictive analytics for supply chain and organizational performance. 70, 308-317.

GUPTA, M. & GEORGE, J. F. 2016. Toward the development of a big data analytics capability. Information & Management, 53, 1049-1064.

GUPTA, S., DRAVE, V. A., DWIVEDI, Y. K., BAABDULLAH, A. M. & ISMAGILOVA, E. 2020. Achieving superior organizational performance via big data predictive analytics: A dynamic capability view. Industrial Marketing Management, 90, 581-592.

HASSNA, G. & LOWRY, P. Big Data Capability, Customer Agility, and Organization Performance: A Dynamic Capability Perspective. Proceedings of the 24th Annual Americas Conference on Information Systems (AMCIS 2018), August 16-18, 2018 2018 New Orleans, USA. AIS.

HINDLE, G., KUNC, M., MORTENSEN, M., OZTEKIN, A. & VIDGEN, R. 2019. Business Analytics: Defining the field and identifying a research agenda. Elsevier.

IQBAL, M., KAZMI, S. H. A., MANZOOR, A., SOOMRANI, A. R., BUTT, S. H. & SHAIKH, K. A. A study of big data for business growth in SMEs: Opportunities & challenges. 2018 International conference on computing, mathematics and engineering technologies (iCoMET), 2018. IEEE, 1-7.

KAR, A. K. & DWIVEDI, Y. K. 2020. Theory building with big data-driven research–Moving away from the “What” towards the “Why”. International Journal of Information Management, 54, 102205.

KEMPENEER, S. 2021. A big data state of mind: Epistemological challenges to accountability and transparency in data-driven regulation. Government Information Quarterly, 38, 101578.

KRISTOFFERSEN, E., BLOMSMA, F., MIKALEF, P. & LI, J. 2020. The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies. Journal of business research, 120, 241-261.

KRISTOFFERSEN, E., MIKALEF, P., BLOMSMA, F. & LI, J. 2021. Towards a business analytics capability for the circular economy. Technological Forecasting and Social Change, 171, 120957.

LAMBA, H. S. & DUBEY, S. K. Analysis of requirements for big data adoption to maximize IT business value. 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), 2015. IEEE, 1-6.

LEHRER, C., WIENEKE, A., VOM BROCKE, J., JUNG, R. & SEIDEL, S. 2018. How big data analytics enables service innovation: materiality, affordance, and the individualization of service. Journal of Management Information Systems, 35, 424-460.

LOEBBECKE, C. & PICOT, A. 2015. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. The Journal of Strategic Information Systems, 24, 149-157.

MANYIKA, J., CHUI, M., BROWN, B., BUGHIN, J., DOBBS, R., ROXBURGH, C. & BYERS, A. H. 2011. Big data: The next frontier for innovation, competition, and productivity.

MAROUFKHANI, P., IRANMANESH, M. & GHOBAKHLOO, M. 2023. Determinants of big data analytics adoption in small and medium-sized enterprises (SMEs). Industrial Management & Data Systems, 123, 278-301.

MAROUFKHANI, P., TSENG, M.-L., IRANMANESH, M., ISMAIL, W. K. W. & KHALID, H. 2020. Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. International journal of information management, 54, 102190.

MCAFEE, A. & BRYNJOLFSSON, E. 2017. Machine, platform, crowd: Harnessing our digital future, WW Norton & Company.

MCAFEE, A., BRYNJOLFSSON, E. & DAVENPORT, T. H. 2012a. Big data: the management revolution. Harvard business review, 90, 60-68.

MCAFEE, A., BRYNJOLFSSON, E., DAVENPORT, T. H., PATIL, D. & BARTON, D. 2012b. Big data: the management revolution. Harvard business review, 90, 60-68.

MIKALEF, P., BOURA, M., LEKAKOS, G. & KROGSTIE, J. 2018a. Complementarities Between Information Governance and Big Data Analytics Capabilities on Innovation. European Conference on Information Systems (ECIS). Portsmouth, UK: AIS.

MIKALEF, P., BOURA, M., LEKAKOS, G. & KROGSTIE, J. 2019a. Big data analytics and firm performance: Findings from a mixed-method approach. Journal of Business Research, 98, 261-276.

MIKALEF, P., BOURA, M., LEKAKOS, G. & KROGSTIE, J. 2019b. Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment. British Journal of Management, In press.

MIKALEF, P., BOURA, M., LEKAKOS, G. & KROGSTIE, J. 2020a. The role of information governance in big data analytics driven innovation. Information & Management, 57, 103361.

MIKALEF, P., FRAMNES, V. A., DANIELSEN, F., KROGSTIE, J. & OLSEN, D. H. 2017. Big Data Analytics Capability: Antecedents and Business Value. Pacific Asia Conference on Information Systems, 2018 2017a Langkawi, Malaysia.

MIKALEF, P., GIANNAKOS, M. N., PAPPAS, I. O. & KROGSTIE, J. The Human Side of Big Data: Understanding the skills of the data scientist in education and industry. Global Engineering Education Conference (EDUCON), 2018 IEEE, 2018b Tenerife, Spain. IEEE.

MIKALEF, P. & GUPTA, M. 2021. Artificial Intelligence Capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. Information & Management, Online.

MIKALEF, P. & KROGSTIE, J. Investigating the Data Science Skill Gap: An Empirical Analysis. 2019 IEEE Global Engineering Education Conference (EDUCON), 2019. IEEE, 1275-1284.

MIKALEF, P. & KROGSTIE, J. 2020. Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities. European Journal of Information Systems, 1-28.

MIKALEF, P., KROGSTIE, J., PAPPAS, I. O. & PAVLOU, P. 2019c. Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. Information & Management.

MIKALEF, P., KROGSTIE, J., PAPPAS, I. O. & PAVLOU, P. 2020b. Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. Information & Management, 57, 103169.

MIKALEF, P., PAPPAS, I. O., KROGSTIE, J. & GIANNAKOS, M. 2017b. Big data analytics capabilities: a systematic literature review and research agenda. Information Systems and e-Business Management, 1-32.

MIKALEF, P., PAPPAS, I. O., KROGSTIE, J. & GIANNAKOS, M. 2018c. Big data analytics capabilities: a systematic literature review and research agenda. Information Systems and e-Business Management, 16, 1-32.

MIKALEF, P., PAPPAS, I. O., KROGSTIE, J. & PAVLOU, P. A. 2019d. Big data and business analytics: A research agenda for realizing business value. North-Holland.

MIKALEF, P., PAPPAS, I. O., KROGSTIE, J. & PAVLOU, P. A. 2020c. Big data and business analytics: A research agenda for realizing business value. Information & Management, 57.

MIKALEF, P., VAN DE WETERING, R. & KROGSTIE, J. 2021. Building dynamic capabilities by leveraging big data analytics: The role of organizational inertia. Information & Management, 58, 103412.

MÜLLER, O., FAY, M. & VOM BROCKE, J. J. J. O. M. I. S. 2018. The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. 35, 488-509.

MÜLLER, O., JUNGLAS, I., BROCKE, J. V. & DEBORTOLI, S. 2016. Utilizing big data analytics for information systems research: challenges, promises and guidelines. European Journal of Information Systems, 25, 289-302.

NIEBEL, T., RASEL, F. & VIETE, S. 2019. BIG data–BIG gains? Understanding the link between big data analytics and innovation. Economics of Innovation and New Technology, 28, 296-316.

OBSCHONKA, M. & AUDRETSCH, D. B. 2020. Artificial intelligence and big data in entrepreneurship: a new era has begun. Small Business Economics, 55, 529-539.

POPOVIČ, A., HACKNEY, R., TASSABEHJI, R. & CASTELLI, M. 2018. The impact of big data analytics on firms' high value business performance. Information Systems Frontiers, 20, 209-222.

RAGUSEO, E. 2018. Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. International Journal of Information Management, 38, 187-195.

RAGUSEO, E. & VITARI, C. 2018. Investments in big data analytics and firm performance: an empirical investigation of direct and mediating effects. International Journal of Production Research, 56, 5206-5221.

RAGUSEO, E., VITARI, C. & PIGNI, F. 2020. Profiting from big data analytics: The moderating roles of industry concentration and firm size. International Journal of Production Economics, 229, 107758.

RUSSOM, P. 2011. Big data analytics. TDWI best practices report, fourth quarter, 19, 1-34.

SAENZ, J., DE GUINEA, A. O. & PENALBA-AGUIRREZABALAGA, C. 2022. Value creation through marketing data analytics: The distinct contribution of data analytics assets and capabilities to unit and firm performance. Information & Management, 59, 103724.

SAGIROGLU, S. & SINANC, D. Big data: A review. 2013 international conference on collaboration technologies and systems (CTS), 2013. IEEE, 42-47.

SHEHAB, N., BADAWY, M. & ARAFAT, H. 2021. Big data analytics and preprocessing. Machine learning and big data analytics paradigms: analysis, applications and challenges, 25-43.

SIVARAJAH, U., KAMAL, M. M., IRANI, Z. & WEERAKKODY, V. J. J. O. B. R. 2017. Critical analysis of Big Data challenges and analytical methods. 70, 263-286.

SUN, E. W., CHEN, Y.-T. & YU, M.-T. 2015. Generalized optimal wavelet decomposing algorithm for big financial data. International Journal of Production Economics, 165, 194-214.

SURBAKTI, F. P. S., WANG, W., INDULSKA, M. & SADIQ, S. 2019. Factors influencing effective use of big data: A research framework. Information & Management.

TALIA, D. 2013. Clouds for scalable big data analytics. Computer, 46, 98-101.

TAN, K. H., ZHAN, Y., JI, G., YE, F. & CHANG, C. 2015. Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. International Journal of Production Economics, 165, 223-233.

UDDIN, M. F. & GUPTA, N. Seven V's of Big Data understanding Big Data to extract value. Proceedings of the 2014 zone 1 conference of the American Society for Engineering Education, 2014. IEEE, 1-5.

VASSAKIS, K., PETRAKIS, E. & KOPANAKIS, I. 2018. Big data analytics: Applications, prospects and challenges. Mobile big data: A roadmap from models to technologies, 3-20.

WAMBA, S. F. 2017. Big data analytics and business process innovation. Business Process Management Journal, 23, 470-476.

WAMBA, S. F., AKTER, S., EDWARDS, A., CHOPIN, G. & GNANZOU, D. 2015. How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics, 165, 234-246.

WAMBA, S. F., GUNASEKARAN, A., AKTER, S., REN, S. J.-F., DUBEY, R. & CHILDE, S. J. 2017. Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research, 70, 356-365.

WAMBA, S. F. & MISHRA, D. 2017. Big data integration with business processes: a literature review. Business Process Management Journal, 23, 477-492.

WATSON, H. J. 2014. Tutorial: Big data analytics: Concepts, technologies, and applications. Communications of the Association for Information Systems, 34, 65.

WIENER, M., SAUNDERS, C. & MARABELLI, M. 2020. Big-data business models: A critical literature review and multiperspective research framework. Journal of Information Technology, 35, 66-91.

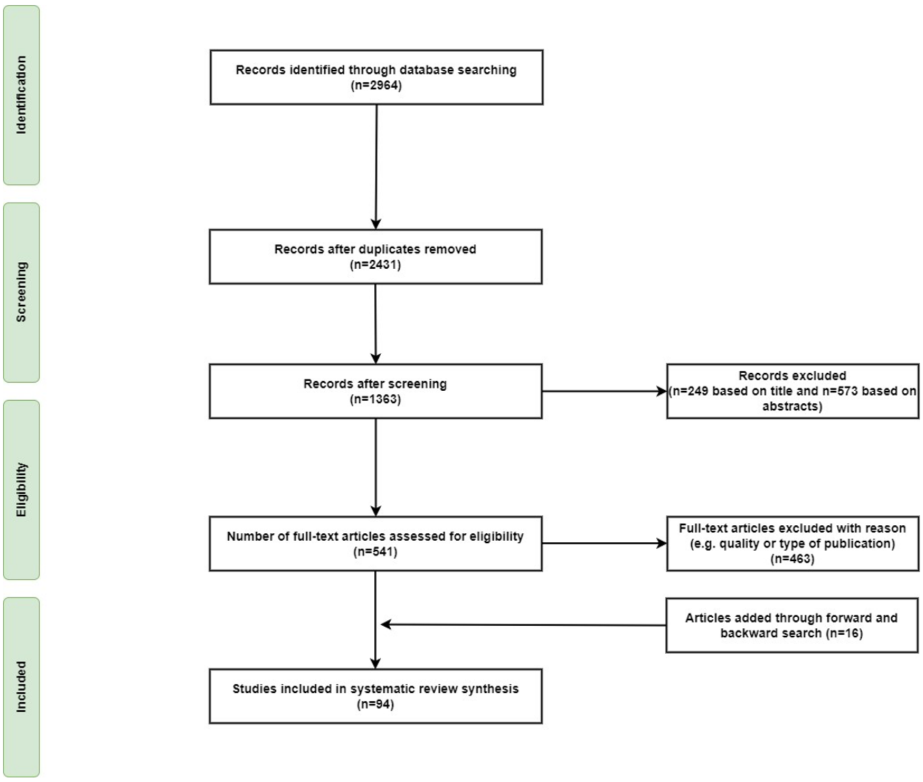
ZHU, S., DONG, T. & LUO, X. R. 2021. A longitudinal study of the actual value of big data and analytics: The role of industry environment. International Journal of Information Management, 60, 102389.

Appendix

Appendix A. Literature review flowchart

The flowchart depicted below represents the four main phases of the literature selection and retention as part of the synthesis of empirical findings for the report. During the first phase (Identification), the use of search strings in the different databases resulted in a total of 2964 studies that were included in the pool of references. The second phase (screening) removed studies that were duplicates, as well as those that were deemed to be irrelevant based on their title or their abstract. After the screening phase, a total of 1363 articles remained. During the third phase (eligibility), the remaining studies were gauged based on their quality, selecting only those that were peer-reviewed and published in reputable journals and conferences, and an in-depth assessment was performed on the content of the main text. At this stage, a total of 463 articles were removed based on quality criteria and the type of publication. During the fourth and final phase (inclusion), a thorough forward and backward search was conducted. This step essentially requires that the final set of articles is analyzed, and important articles that are cited within this pool of papers are evaluated, as well as articles that cite each of the articles in the pool. Through the forward and backward search, 16 additional studies were included in the sample, bringing the total number to 94 studies that were retained for further assessment.

Figure 1 Flowchart of the selection and filtering process of the literature included in the report

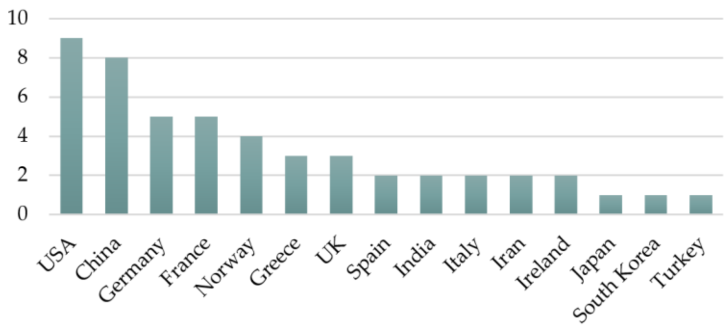


Source: Author

Appendix B. Empirical studies per country

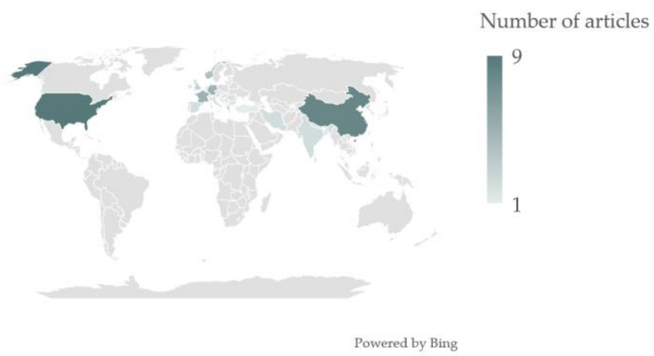
The studies that we retained in our selected pool of papers are mainly empirical results, but they also include some conceptual articles and review papers. From the empirical studies, most papers included information concerning the origin of the sample, and from which country(ies) data were obtained. While most studies had data from a single country, some studies were built on samples from two or more countries. In certain cases, the name of the countries was included in the article, while some other studies only identified large geographical regions as the source of data collection (e.g., Europe, Asia). These studies were not included in the descriptive summary of cases, but only those that specified where the data originated from at a country level. The final subset of empirical studies is visually represented in the graph (Figure 2) and heatmap (Figure 3). Most studies have been conducted in the USA and China, with a substantial proportion done in Europe. From the European countries, the largest samples are from Germany, France, and Norway. Notably, there were no studies that examined a sample from Sweden.

Figure 2 Country representation in empirical studies



Source: Author

Figure 3 Geographical distribution of samples from empirical studies



Source: Author

Appendix C. Tillväxtanalys publikationer inom företags digitalisering, IT-användning och Artificiell Intelligens

År	Titel	Status
2014	Digitaliseringens bidrag till tillväxt och konkurrenskraft i Sverige	Publicerad (Rapport 2014:13)
2014	Hur driver IKT produktivitet och tillväxt? Analyser av kvantitativa data	Publicerad (PM 2014:17)
2017	Digital mognad i svenskt näringsliv	Publicerad (Rapport 2017:02)
2017	IT-användning och företagens produktivitet – Förslag på en indikator för digital intensitet i företag	Publicerad (PM 2017:16)
2018	Hur kan staten främja användandet av digitaliseringens möjligheter i näringslivet	Publicerad (Rapport 2018:01)
2019	Företagens digitala mognad 2018	Publicerad (Rapport 2019:12)
2021	Drivers of AI adoption – A literature review	Publicerad (Rapport 2021:07)
2022	AI-politik för konkurrenskraft	Publicerad (Rapport 2022:02:01)
2022	Varför AI? – möjligheter och utmaningar med AI-implementering i det svenska näringslivet	Publicerad (Rapport 2022:09)
2023	En kvantitativ analys av AI-användning och produktivitet i företag	Publicerad (Rapport 2023:02)
2023	Hur omformar AI näringslivet och hur kan politiken utvecklas	Publicerad (Rapport 2023:04)
2023	Reglering av AI: För lite för sent eller för mycket för tidigt?	Publicerad (Rapport 2023:17)
2024	En kartläggning av BDA-användning och produktivitet bland svenska företag	Kommande (Q1 2024)

På vilket sätt statens insatser bidrar till svensk tillväxt och näringslivsutveckling står i fokus för våra rapporter.

Läs mer om vilka vi är och vad nyttan med det vi gör är på www.tillvaxtanalys.se. Du kan även följa oss på LinkedIn och YouTube.

Anmäl dig gärna till vårt nyhetsbrev för att hålla dig uppdaterad om pågående och planerade analys- och utvärderingsprojekt.

Varmt välkommen att kontakta oss!

Tillväxtanalys

Myndigheten för tillväxtpolitiska utvärderingar och analyser

Tillväxtanalys
Studentplan 3, 831 40 Östersund
Telefon: 010-447 44 00
E-post: info@tillvaxtanalys.se
Webb: www.tillvaxtanalys.se