



Universiteti i Tiranës
Fakulteti i Ekonomisë
Dega Informatikë ekonomike
Departamenti Statistikë dhe Informatikë e zbatuar



TEMË DIPLOME

RRITJA EKSPONENCIALE E BIG DATA DHE BIZNESET

Studenti:

Joana Braja

Udhëheqësi shkencor:

Prof.Dr. Fatmir Memaj

Tiranë 2018

ABSTRAKT

Ky punim eksploron problemin e rritjes eksponenciale të big data parë nga këndvështrimi i bizneseve. Miliarda të dhëna gjenerohen çdo ditë në rang global. Ky numër pritët të rritet edhe më shumë përgjatë viteve në vazhdim dhe kjo do të ketë ndikim direkt te bizneset, menaxhimi i tyre dhe proceset e biznesit.

Në seksionin e parë, autori rishikon literaturën ekzistuese në lidhje me big data dhe llojet e të dhënave. Identifikohen pesë karakteristikat e big data dhe kategorizohen të dhëna në të strukturuar dhe të pastrukturuar duke u ilustruar me shembuj përkatës. Ky seksion parashtrohet bazën teorike për trajtimin e problemeve në vijim.

Më tej teksti evidenton rritjen eksponenciale të big data dhe shpjegon se çfarë përfitime dhe sfida paraqet kjo rritje për bizneset. Për ta konkretizuar u krye studimi i kompanive Shqiptare dhe marrëdhënies së tyre me big data. U përdor metoda kualitative dhe specifikisht intervistat. U dol në përfundimin se kompanitë shqiptare kanë filluar të ndjejnë sfidat e menaxhimi të të dhënave dhe ndërmarrin masa përkatëse.

Punimi kërkon t'i japë zgjidhje teknologjike dhe menaxheriale bizneseve që përballen me sfidat e big data. Këtë e arrin duke eksploruar gjendjen reale dhe duke sugjeruar zgjidhjet e mundshme. Studimi përqipet ti përgjigjet pyetjeve: Si mund të shfrytëzojmë maksimalisht përfitimet që na sjell big data, duke përballuar sfidat teknologjike dhe duke qëndruar brenda kornizave etike?

Tabela e përmbajtjes

1. Hyrje	4
1.1 Big Data	4
1.1.1 Pesë V-të e Big Data dhe format e të dhënave	4
1.1.2 Llojet e të dhënave	5
2. Trendi i rritjes së Big Data	6
2.1. Mënyrat e përvetësimit të Big Data	7
2.2. Problemet që sjell kjo rritje për bizneset	8
2.3 Sfidat teknologjike dhe sistemet IT	10
2.4. Përfitimet e bizneseve	12
2.4.1 Reduktimi i kostove	12
2.4.2 Vendimarrje më e mirë	13
2.4.3 Optimizimi i proceseve	14
3. Kompanitë shqiptare dhe Big Data	15
3.1. Metoda e studimit	15
3.2 Mbi teknikalitetet dhe intervistat	16
3.3 Mbi kompanitë	17
3.4 Rezultati i studimit	18
3.4.1 Sistemet ERP dhe trendi në të dhëna	18
3.4.2 Këndvështrimi i kompanive lidhur me big data	19
3.4.3 Vendimarrja dhe big data	19
3.5 Përmirësimet e mundshme	21
3.4.1 Përmirësime në marrëdhënie me klientin	21
3.4.2 Përmirësime në vendimarrje	21
4. Sugjerime teknologjike dhe çështje etike të big data	22
4.1 Big Data Hadoop	23
4.2 Data Mining	24
4.3 Çështjet etike të big data	25
4.4 Rast Studimor – Facebook, mundësitë dhe problemet e privatësisë	26
4.5 Kufijtë etik të kompanive	28
5. Konkluzione dhe sugjerime	29
 Referenca	 32

1. Hyrje

Historikisht, të dhënat kanë qenë një aspekt thelbësor në evolucionin njerëzor. Që në prehistori, njerëzit filluan të mbledhnin informacion për botën përreth dhe ta përdornin atë për të mbijetuar.

Sot, mënyra se si i përdorim të dhënat është më komplekse dhe me rritjen e përdorimit të teknologjisë është rritur edhe vështirësia e menaxhimit të të dhënave.

Jetojmë në një kohë ku një sasi shumë e madhe të dhënash gjenerohet çdo sekondë.

Të dhënat gjenerohen nga sensorët, imazhet dixhitale, aktiviteti në rrjete sociale, nga pajisje shtëpiake dixhitale (internet of things). Në masën që janë përhapur këto teknologji, ritmet e rritjes së të dhënave janë marramendëse.

Këtu, hyn në ekuacion Big Data, që është term që përdoret për të kategorizuar të dhënat e sipërpërmendura.

Nëse këto të dhëna analizohen mund të dilet në përfundime të vlefshme për bizneset. Këto të fundit e përdorin informacionin nga analiza e Big Data në avantazhin e tyre: mësojnë më tepër rreth target grupit të tyre dhe përmirësojnë performancën.

1.1 Big Data

Big Data, sikurse u përmend edhe më parë, është term që përdoret për të dhëna në sasi të mëdha. Koncepti i Big Data u krijua në shekullin e 21-të dhe me kalimin e kohës filloi të evoluonte. Për herë të parë filluan të përmendeshin terma për të përshkruar natyrën e të dhënave [Emani, C. K. Nadine, C., Christophe, N. (2015)]. Një studim i vitit 2001 nga Doug Laney foli për tre atributet (tre V-të) që përshkruajnë natyrën e Big Data të cilat ishin vëllimi (volume), shpejtësia (velocity) dhe shumëllojshmëria (variety). Për kohën, këto attribute ndihmuan të kuptohej më mirë se çfarë Big Data nënkupton. Me kalimin e viteve dhe me evoluimin e konceptit, sot njihen 5 attribute të Big Data.

1.1.1 Pesë V-të e Big Data dhe format e të dhënave

- **Vëllimi (Volume)** – Volumi i referohet sasisë shumë të madhe të të dhënave që gjeroen çdo sekondë nga rrjetet sociale, celularët, makinat, kartat e kreditit sensorët M2M, fotografitë videot etj.
- **Shpejtësia (Velocity)** – Ky atribut i referohet shpejtësisë së lartë me të cilën të dhënat krijohen, mbledhen dhe analizohen.
- **Shumëllojshmëria (Variety)** – I referohet llojeve të ndryshme të të dhënave që mund të përdoren. Të dhënat nuk janë vetëm të strukturuar (si emri mbiemri, adresa, numri i telefonit etj). Në fakt, 80% bëjnë pjesë në kategorinë e të dhënave të pastrukturuar si fotot, sekuenca videosh, përditësimi i rrjeteve sociale etj [Demchenko, Y. (2013)].

- **Vlera (Value)** - Kur flasim për vlerën, i referohemi sa e vlefshme është e dhëna që po marrim [Vishal, S., Pravin, D. (2013)]. Të dhënat që mbliidhen janë të pavlefshme nëse nuk përpunohen për të krijuar informacion, Ka disa të dhëna që janë të panevojshme për të dalë në përfundime dhe disa të tjera shumë të rëndësishme pasi janë analizuar.
- **Vërtetësia (Veracity)** - vërtetësia është kualiteti i një të dhëne të besueshme. Sa të sakta janë gjithë këto të dhëna [Sanjay, R. (2015)] .

1.1.2 Llojet e të dhënave

Përveç attributeve që karakterizojnë Big Data kemi dhe kategoritë e të dhënave që janë tre:

1. **Të dhëna të strukturuar**- janë tipe të dhënash të mirë përcaktuara, lehtësisht të aksesueshme dhe mund t'i kërkojmë me thjeshtësi [Zhao, Y., Jie W & Cong Liu, (2014)]. Zakonisht gjenden në databaza relacionare. Shembuj të të dhënave të tilla janë emrat, numrat e telefonit, id-ja e një personi, zip kodi etj. Disa sisteme që përdorin të dhëna të strukturuar janë sistemet e rezervimit të biletave, kontrolli i inventarëve, transaksionet e blerjeve ose të shitjeve, aktiviteti i ATM.
2. **Të dhëna gjysmë të strukturuar**- janë të dhëna që përmbajnë tag-e semantike, por nuk kanë strukturën tipike që lidhen me databazat relacionare. Shembuj të të dhënave gjysmë të strukturuar janë email ka një strukturë të brendshme falë metadatës së tyre (adresa e email, emri i personit që e dërgon, subjekti etj.), por në të njëjtën kohë kanë një fushë jo të strukturuar (në fushë mund të dërgohen audio, video etj.)
3. **Të dhëna të pastrukturuar**- të dhëna që nuk kërkohen me lehtësi në databazë si psh formate si audio, video dhe postimet e rrjeteve sociale. Shembuj të të dhënave të pastrukturuar janë të dhënat nga rrjete sociale si Facebook, Tëitter apo Instagram, vendodhja që gjendet me GPS, fotot dixhitale, paketa office, imazhet me satelitë , të dhënat nga sensorët internet of things.

Është e rëndësishme të njohim llojet e të dhënave, sepse në bazë të tyre kuptojmë se si do t'i manaxhojmë të dhënat dhe si do t'i aplikojmë për të nxjerrë informacione të ndryshme prej tyre. Ky trend në rritje është një mundësi e re për bizneset dhe nëse dinë ta përdorin si duhet trendin, padyshim do të krijojnë avantazh konkurrues.

2. Trendi i rritjes së Big Data

Në vitet e fundit është e dukshme se ka patur një rritje të interesit për Big Data jo vetëm nga ana e akademikëve dhe studjuesve, por edhe nga politikanë dhe drejtues të vendeve të zhvilluara. Administrata Obama në mars të vitit 2012 shpalli se do të investonte 200 milion dollarë në hulumtime për big data [Lakshmi C. V. V., Nagendra, K. (2016)].

Raporte të Korporatës Internacionale të të dhënave parashikojnë se të dhënat globale nga 2005 deri në 2020 do të rriten me faktor 10. Vëllimi i të dhënave globale do të rritet nga 0.13 Zettabyte në 40 zettabyte. IBM vlerëson se 90% e të dhënave të së gjithë botës janë krijuar dy vitet e fundit.

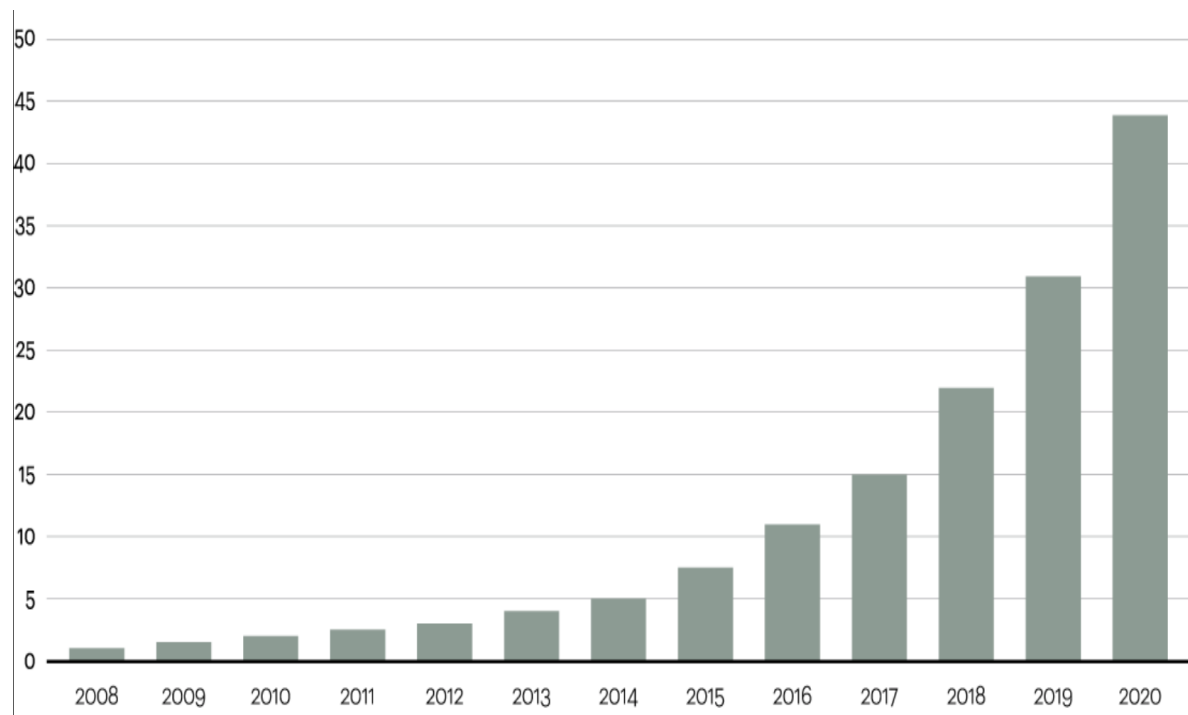


Fig 1.

Burimi: Oracle (2018)

Është parë se rrjetet sociale si Facebook ka 750 milionë përdorues me 350 milion foto që shkarkohen çdo ditë. Twitter ka 320 milion llogari me 500 milion postime në ditë. Por, s'janë vetëm rrjetet sociale përgjegjëse për një rritje të tillë të Big Data.

2.1. Mënyrat e përvetësimit të Big Data

Të dhënat që mund të mblidhen nga faqe si Facebook apo Google janë vetëm një pjesë e big data. Mënyrat e përvetësimit të të dhënave në këto raste janë ose të dhëna që i kërkohen përdoruesit drejtëpërdrejtë, (psh.: të dhënat që kërkon google përpara se të regjistrohesh në Gmail), ose të dhëna që merren nga aktiviteti i përdoruesve në faqe (psh.: shkarkimi i fotove ose lëvizja e mouse-it).

Në fakt, një pjesë e madhe e big data përbëhet nga sensorët ose ndryshe ato që quhen Internet of Things (Iot). Shembuj të këtyre të dhënave janë sensorët e pacientëve në spitale, sensorët e motit, sensorët e vendodhjes (GPS), sensorët dhe kamerat që gjenden në automjete , sensorë në lavatriçe etj.

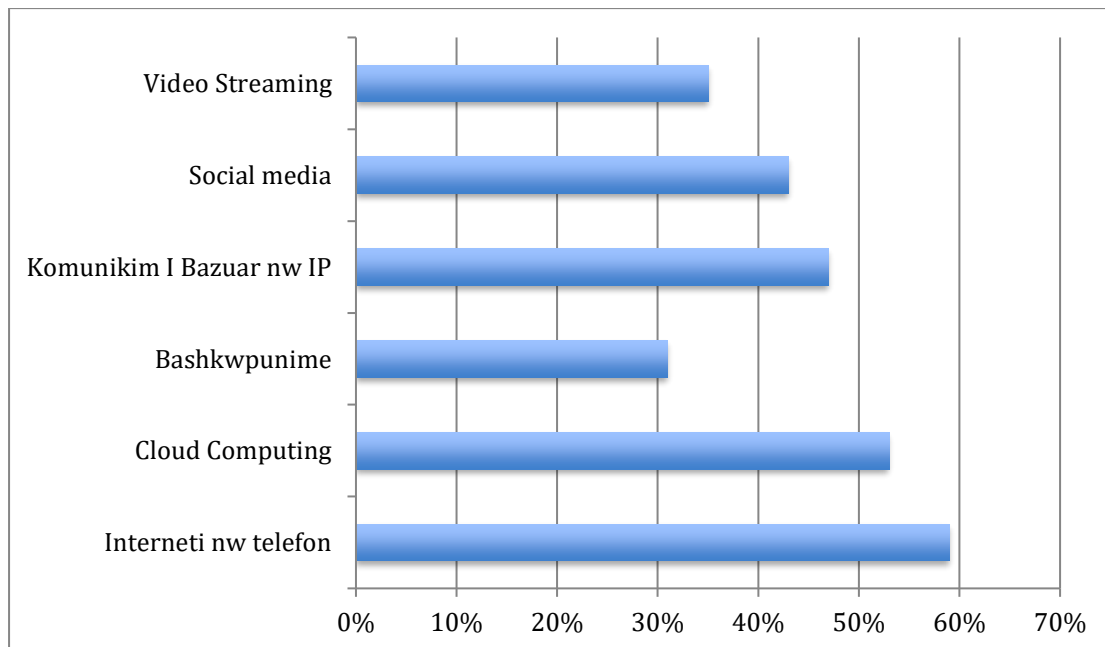


Fig 2.

Burimi: IBM (2017)

Si pasojë e rritjes së pothuajse miliarda pajisjesh të lidhura me njëra tjetrën, internetit në telefon, cloud computing, komunikimit të bazuar në IP, rrjeteve sociale, me ndikim respektivisht me përqindjet e mësipërme, mënyra se si i manaxhohim të dhënat ka për të ndryshuar masivisht vitet që vjen. Kjo rritje me një shpejtësi dhe me një volum të tillë sigurisht që sjell sfida jo vetëm në infrastrukturën kompjuterike në të cilën manaxhohen të dhënat por edhe në mjedisin e bizneseve.

Kompanitë që manaxhojnë të dhëna të përmasave të tilla do të duhet të gjejnë metodologji jo konvencionale që i përshtaten ndryshimeve të tilla. Nga ana tjetër një sasi kaq e madhe të dhënash përbën një mundësi për bizneset.

2.2. Problemet që sjell kjo rritje për bizneset

Përshtatja me ndryshimet është gjithmonë sfiduese për kompanitë. Përshtatja me rritjen në mënyrë eksponenciale të big data nuk bën përjashtim. Është një faktor që pavarësisht se mund të përdoret në të mirë të kompanisë, rrit edhe kompleksitetin e menaxhimit të të dhënave. Më poshtë kam trajtuar disa prej sfidave kryesore që hasin bizneset në përballjen me rritjen massive të të dhënave.

Sasia e madhe e të dhënave të pastrukturuara

Problemi i vetëm nuk është rritja me hapa galopant i gjenerimit të të dhënave. Ajo cka është edhe më e vështirë për tu manaxhuar janë fakti që shumica e të dhënave janë të dhëna të pastrukturuara. Nëse të dhënat tradicionale janë më lehtësisht të trajtueshme, kompleksiteti i të dhënave të pastrukturuara përbëjnë një problem më vete. Kompanitë krahas databazave relacionare duhet të implementojnë databaza më dinamike NoSQL si MongoDB dhe DynamoDB. Përveç ruajtjes së të dhënave duhet edhe të analizohen. Për të analizuar llogaritjet me të dhënat disa zgjidhje janë sisteme të shpërndara kompjuterike si hadoop, hive, Pig etc. [Rizkallah, J. (2017)]

Mungesa e Shkencëtarëve të të dhënave

Numri i ulët i shkencëtarëve të të dhënave është një problematikë globale. Problemi nuk lidhet vetëm me ekzistencën e këtyre të fundit, por dhe me mospërputhjen që kanë në mendime biznesmenët me shkencëtarët e të dhënave. Analistët e të dhënave shpesh nuk kanë vizionin e biznesmenëve kur analizojnë këto të dhëna. Për këtë arsye shpeshherë nuk japin raporte që kanë vlerë të madhe për kompanitë.

Pagat për shkencëtarët dhe analistët e të dhënave janë të larta. Kjo për shkak të kërkesës së lartë për punë dhe ofertës së ulët. Është e kushtueshme për kompanitë të paguajnë analistët më të mirë të të dhënave, por edhe trainimi përkthehet në kosto shtesë. [Half, R. (2016)] Për këtë arsye, shumë organizata preferojnë ti zgjidhin problemet me software automatizimi, gjuhës së makinës, dhe inteligjencës artificiale për të kuptuar të dhënat me kodim minimal. Megjithatë as kjo nuk e zgjidh problemin e mungesës së ekspertëve në fushë.

Analizat në kohë reale

Nevoja për analiza të big data në kohë reale shtojnë edhe më shumë kompleksitetin. Disa biznese si bankat ose kompanitë e siguracionit humbasin informacione të rëndësishme nëse nuk veprojnë në kohë në momentin që gjenerohen të dhënat.

Për ta arritur këtë, kompanitë kanë nevojë për mënyra analizimi që përballojnë të dhëna me vërtetësi dhe shpejtësi (veracity and velocity). Ekzistojnë disa software që e arrijnë këtë si ETL engines, vizualizimi, llogaritja dhe libraritë. Ka dhe dy software që shërbejnë për lehtësimin e ETL si Xplenty, StarfishETL etj.

Siguria e të dhënave më komplekse

Pritet që me volumin e madh të të dhënave vinë dhe vështirësi në sigurinë e informacionit dhe sigurinë dhe integritetin e të dhënave. Burimet nga vinë big data janë të mëdha dhe nyjet dhe kanalet e lidhjes janë të shumta. Kjo sjell risk të lartë dhe mundësi të hacker-ave për të shfrytëzuar dobësitë e sistemit. Duke qenë se të dhënat janë kritike, një incident mund të rezultojë në humbje të mëdha. Kjo i krijon nevojën kompanisë për sisteme të sigurta dhe për praktika po aq të sigurta në kompani.

2.2.5 Sjellja e Organizatës

Jo vetëm analistët dhe shkencëtarët duhet të kujdesen për të dhënat. Që kompania të këtë sukses në manaxhimin e Big Data duhet që kultura dhe struktura e kompanisë të jetë mbështetëse e vendimarrjeve bazuar në të dhëna.

Kjo është nga pjesët më sfiduese, manaxhimi i burimeve njerëzore lidhur me big data. Zëvendësimi i strukturave ekzistuese dhe krijimi i kulturave me në bazë të dhënash kërkon kohë dhe angazhim. Të kuptuarit e rolit të manaxhimit dhe të të dhënave është po aq e vështirë për manaxherët sot. Big Data është një faktor që po ndryshon me ritme marramendëse dhe tepër të shpejta. Si e tillë kërkon përshtatje të kompanisë me këtë situatë sa më herët.

Sipas ekspertëve të PWC, kompanitë duhet të punësojnë drejtues që jo vetëm kuptojnë përsosmërisht të dhënat, por që ndërmarrin iniciativën për ta sfiduar statusin e praktikave ekzistuese dhe sugjerojnë ndryshime përkatëse.

2.3 Sfidat teknologjike dhe sistemet IT

Sfidat teknologjike, lidhen me infrastrukturën e IT-së, sigurinë dhe privatësinë, si dhe sfidat te tjera teknologjike. Së pari, për të përfituar nga big data, për të pasur mundësi për të mbledhur, ruajtur, menaxhuar dhe analizuar të dhënat në një mënyrë më të mirë, teknologjitë e informacionit në organizata janë të rëndësishme. Pa infrastrukturë të mirë IT-je, nuk është e mundur të arrihen avantazhet e big data. Për shkak të rritjes së volumit dhe shumëllojshmërisë së të dhënave, sfida të mëdha kanë lindur në infrastrukturën e organizatës. Sasia e lartë e big data çon në probleme të kapacitetit për ruajtjen e të dhënave, përpunimit dhe shkëmbimit të tyre, që shkon përtej kapacitetit të sistemeve të databazave ekzistuese. Si për shembull grupet e të dhënave që bëhen të disponueshme me anë të big data, përmbajnë sasi informacioni të cilat janë më të mëdha se kapaciteti i një kompjuteri të vetëm. Grupet e të dhënave (Datasets) përbëhen nga të dhëna të mëdha dhe komplekse, të cilat mund të jenë të vështira për t'u mbledhur në një vend të vetëm për ruajtjen e tyre. Kjo gjithashtu paraqet vështirësi në komunikimin e këtyre të dhënave brenda dhe jashtë organizatës dhe për këtë arsye mund të kërkojë kohë për të komunikuar.

Një aspekt i rëndësishëm në trendin e big data është përhapja e madhe e sistemeve të informacionit si Planifikimi i Burimeve të Ndërmarrjeve (ERP), Menaxhimi i Zinxhirit të Furnizimit (SCM) dhe Menaxhimi i Marrëdhënieve me Klientin (CRM). Përdorimi i sistemeve të tilla si ERP në kombinim me rritjen e big data duhet të çojë në një vendimmarrje më të mirë organizative dhe për këtë arsye një rritje në performancën organizative [Abernethy, M. A., Chua, W., Luckett, P. F., & Selto, F. H. (1999)]. Për shembull, duke përdorur këto sisteme, është e mundur që të përdorim inteligjencën e biznesit (Business Intelligence) për të bërë analiza më gjithëpërfshirëse me të dhënat operacionale [Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011)]. Literatura ekzistuese argumenton se përdorimi i sistemeve ERP mund të ofrojë disa përparësi për organizatat. Megjithatë, nuk është aq e lehtë për të arritur këto avantazhe, të cilat shkaktohen nga mënyra se si sistemi përdoret nga organizata. Për të arritur përfitimet e përshkruara të implementimeve ERP dhe për të fituar avantazhe strategjike, riorganizimi i procesit të biznesit ose dizenjimi nga e para e procesit të biznesit kërkohet për përdorim strategjik të sistemit ERP [Davenport, T. H. (2014)]. mënyra se si organizatat marrin qasje më të madhe në informacionin që kërkohet për të marrë vendime më të shpejta dhe më efektive është i rëndësishëm. Kjo do të thotë se është e rëndësishme për të investuar në proceset organizative, jo vetëm në teknologjinë e big data.

Megjithatë, në shumicën e rasteve, kompleksiteti i sistemeve ERP, kufizimet e sistemit dhe mungesa e vullnetit të punonjësve për të ndryshuar mund të shkaktojë pamundësi që të arrihet riorganizimi i procesit të biznesit ose dizenjimi nga e para i procesit të biznesit. Në raste të tilla, është e mundur vetëm përdorimi teknik i sistemit, në të cilin proceset nuk ndryshojnë [Frizzo-Barker, J., Chow-White, P. A.,

Mozafari, M., & Ha, D. (2016)] Kjo sjell pamundesine e sistemeve te organizatës per të ofruar plotësisht avantazhet e big data, sepse sistemi nuk pershtatet me zhvillimet e fundit teknologjike. Për këtë arsye, është e rëndësishme që të ketë një infrastrukturë të IT- së e cila mund të sigurojë plotësisht mundësitë dhe kërkesat për përdorimin e të dhënave të mëdha. Me fjalë të tjera, organizatat kanë nevojë për sisteme të IT-së që bëjnë të mundur pershtatjen me tendencat e fundit siç janë big data, të cilat mund të arrihen me anë të përdorimit strategjik të sistemit.

Përveç infrastrukturës së IT, gjithashtu siguria dhe privatësia e të dhënave është një sfidë teknologjike [Goes, P. B. (2014)]. Databazat e medha, përmbajnë shumë informacione të vlefshme (privacy) dhe mbrojtja e këtij informacioni është një aspekt i rëndësishëm për të mbrojtur të dhënat nga përdorimi i paautorizuar nga palët e treta. Rreziqet në këtë aspekt janë rreziqet rreth sigurisë së pronësisë intelektuale dhe përgjegjësisë dhe rreziqeve të rrjedhjeve të të dhënave [Granlund, M., & Malmi, T. (2002)]. Për të parandaluar këtë, është e domosdoshme që aksesit në keto të dhëna të jetë i disponueshem vetëm për punonjësit që kanë nevojë për këto të dhëna për ekzekutimin e punës së tyre. Për më tepër, është e rëndësishme për të mbrojtur të dhënat kundër hakerëve. Një mënyrë për të rritur besueshmërinë e sistemit është duke përdorur një organizatë të palës së tretë me reputacion të mirë, për shembull një ofrues të sistemeve softuerike si ERP .

Së fundi, disa sfida lindin gjatë grumbullimit të të dhënave, integritetit të të dhënave dhe përpunimit dhe analizimit të tyre. Së pari, sfidat lindin me mbledhjen e të dhënave. Vëllimet e reja të të dhënave çojnë në shumë njohuri të reja dhe të vlefshme. Megjithatë, jo të gjitha të dhënat priten të kenë të njëjtën vlerë të shtuar . Dallimi i informacionit të pavlefshëm nga informacioni i vlefshëm, aktualisht është një aspekt i rëndësishëm në të dhënat e mëdha. Por realizimi i kësaj nuk është i lehtë, pasi format e pastrukturuara të të dhënave që krijohen për shkak të sasive të medha të informacionit, nuk janë lehtësisht të menaxhueshme. Kjo përforcohet më tej nga shfaqja e më shumë sensorëve dhe mënyrave të tjera të matjes së të dhënave, të cilat shpesh bazohen në të dhënat e pastrukturuara [Thornhill, A. (2009)]

Së dyti, shumëllojshmëria e të big data mund të shkaktojë probleme në momentin e bashkimit të të dhënave me njera tjetren [Shaikh, A. A., & Karjaluoto, H. (2015)]. Integrimi i të dhënave të mëdha që vijnë nga burime të ndryshme shkakton shumë pengesa sepse të gjitha këto formate të të dhënave, kanë formatet e tyre të cilat janë të specifikuara dhe të ndryshme. Kjo shumëllojshmëri ka pasoje që bashkimi i këtyre të dhënave të bëhet i vështirë Për më tepër, lindin probleme me ruajtjen e të dhënave për shkak të kompleksitetit. Për shembull, si rezultat i sasive të mëdha të të dhënave të pastrukturuara, numri i dimensioneve të ndryshme të të dhënave është rritur në sasi të mëdha [Sharma, S. (2016)].

Së treti, përpunimi efektiv dhe efikas i të dhënave të mëdha dhe analizimi i tyre pas përpunimit është i rëndësishëm. Analiza është gjenerimi i njohurive dhe inteligjencës,

e cila është e nevojshme për të mbështetur vendimmarrjen dhe objektivat strategjike [McAfee, A. (2002)]. Megjithatë, ky përpunim dhe analizë efikase nuk është gjithmonë i mundur, sepse big data çojnë në sasi të mëdha të dhenash të mbledhura nga sensorë dhe parametra të ndryshëm. Kjo krijon elemente të ndryshme të modelimit, të cilat e komplikojnë interpretimin e këtij informacioni që kërkohet për vendimmarrje [McAfee, A., & Brynjolfsson, E. (2012)].

Për këtë arsye, për të arritur vendim-marrje efektive me anë të big data, është e rëndësishme që big data të kthehen në të dhëna të vlefshme ose smart data, në mënyrë që të sigurohet informacion i saktë me të dhëna të sakta. Vetëm në këtë mënyrë është e mundur të menaxhohen të dhëna në vëllime të mëdha dhe shumëllojshmëri. Për shembull, kjo mund të arrihet duke përpunuar dhe transformuar të dhëna të mëdha të pastruara në të dhëna të strukturuara. Nëse bëhet përpunimi dhe transformimi i tyre, bëhet e mundur të analizohen të dhënat për shembull përmes Business Intelligence që gjeneron diagramë dhe figura, duke çuar në të dhëna të efektshme dhe efikase [Frizzo-Barker, (2016)].

2.4. Përfitimet e bizneseve

Në tregjet konkurruese globale por edhe në tregun shqiptar, të njohurit e klienteve potencialë dhe dëshirat e tyre, është bërë e domosdoshme për të patur avantazh konkurrues. Duke përdorur këto të dhëna për të njohur konsumatorët zvogëlohet barriera biznes-konsumator dhe plotësimi i nevojave të këtyre të fundit bëhet më i thjeshtë. Kompani të fuqishme globale kanë filluar manaxhimin e big data dhe analizën e të dhënave.

Por, analizimi i Big Data nuk është me përfitim vetëm për kompanitë e mëdha. Të gjitha llojet e organizatave mund të përfitojnë prej tyre, si kompanitë që operojnë online ashtu edhe firmat tradicionale. Sipas literaturës së shqyrtuar ka shumë përfitime nga Big Data. Avantazhet dhe përfitimet mund të ndahen në tre kategori kryesore: optimizimi i proceseve dhe produkteve, reduktimi i kostove, vendimarrja më e mirë.

2.4.1 Reduktimi i kostove

Përfitimi i kostove arrihet me anë të ekonomive të shkallës dhe ekonomive të fushëveprimit, të cilat lindin për shkak të rritjes së madhësisë operationale. Kjo është e mundur sepse organizatat i kalojnë të dhënat e tyre në depot e ndërmarrjeve që janë të pajisura posaçërisht për të prodhuar aplikacione analitike [Fanning, K., & Grant, R. (2013)]. Për më tepër, lindin edhe ekonomitë e shkallës dhe ekonomitë e fushës sepse shpejtësia e produkteve dhe shërbimeve në zinxhirin e furnizimit / kërkesës pritet të rritet [Chow-White, P. A., & Green, (2013)]. Një mënyrë tjetër e përfitimeve të kostos është e mundur kur punonjësit

trajnohen të përdorin big data. Në këtë rast, është e mundur të kryhen detyrat më efektivisht, duke rezultuar në reduktime të kostos.

2.4.2 Vendimarrje më e mirë

Big data krijojnë disponibilitet më të madh, shikueshmëri dhe transparencë të informacionit. Teknikat e reja u mundësojnë organizatave të gjejnë modele dhe lidhje reja, gje e cila nuk ishte i mundur pa big data. Këto teknika mund të çojnë drejt menaxhimit më preciz [Demirkan, H., & Delen, D. (2013)]. Për shembull, vendimarrësit mund të marrin më shumë njohuri për sjelljen e klientëve, sepse çdo ndërveprim me konsumatorët prodhon të dhëna. Një mënyrë tjetër për mbledhjen e informacionit të klientëve është e mundur në mënyrë vullnetare, kjo ndodh kur klientët ofrojnë informacionin e tyre personal për organizatën, për shembull për të arritur zbritje shtesë ose veprime të tjera promovuese / përparësi [Abernethy (1999).]. Duke mbledhur dhe analizuar këto të dhëna, krijohen mundësi për të zhvilluar modele që parashikojnë kërkesat e ardhshme të klientëve. Kështu, bëhet e mundur të parashikohen kërkesat e konsumatorëve, duke marrë vendime që bazohen jo në intuitë, por që bazohen në të dhëna [Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A. S., & Buyya, R. (2015)]. Një shembull i kësaj është menaxhimi i inventarit, në të cilin big data mund të çojnë në blerje efikase. Parashikimi i kërkesave të konsumatorit bën që t'i plotësohen më mirë nevojat këtij të fundit. Është gjithashtu e qartë se sa stok nevojitet. Kjo çon në kosto më të ulëta dhe më pak hapësirë të nevojshme për ruajtjen e stokut.

Big data ofrojnë mundësi të tjera. Përveç mundësisë për të parashikuar sjelljen e konsumatorëve të ardhshëm, është gjithashtu e mundur të përcaktohet gjendja aktuale e konsumatorit. Duke analizuar modelet e të dhënave të një klienti të caktuar (për shembull, sjellja e klikimeve në një faqe interneti) mund marrim njohuri për atë që një klient mund të ketë nevojë në një moment të caktuar. Në një mënyrë të tillë organizata mund t'u përgjigjet nevojave specifike të klientëve përmes segmentimit më të ngushtë të konsumatorëve, për shembull duke ofruar produkte të veçanta në kategorinë e kërkimit, duke çuar në aktivitete më efektive të shitjeve të fokusuara në nevojat e klientit. Për shembull, kur një klient shpesh kërkon një kategori të caktuar të produkteve në faqen e internetit të organizatës, organizata e di se klienti ka të ngjarë të jetë i interesuar në këto produkte të veçanta. Në këtë mënyrë, organizata mund të harmonizojë shitjet duke sjellë këto objekte në vëmendjen e klientit, ose ofron zbritje për klientin dhe e mban atë të kënaqur. Duke përdorur këtë mënyrë pune, organizatat përdorin big data për të përmirësuar aktivitetet e shitjes .

Kjo vendim-marrje më e shpejtë dhe më e mirë mund të çojë në një performancë më të mirë për shkak të rritjes së efikasitetit dhe efektivitetit të organizatës dhe strategjisë organizative [Dull, R. B. (2012)]. Kjo është konfirmuar në hulumtimin

e Brynjolfsson (2011), ku thuhet se firmat që kanë vendimmarrje në bazë të të dhënave dhe analizave të biznesit tregojnë performancë më të lartë. Kjo performancë e përmirësuar mundësohet nga disponueshmëria dhe përdorimi i informacionit të saktë për vendimmarrje, i cili krijohet nga big data [Shao, B. B. M. (2016)] .

Së katërti, vendimmarrja më e shpejtë dhe më e mirë është e mundur, sepse të big data rritin mundësitë për punonjësit dhe drejtuesit e organizatës për të bërë pyetje dhe për t'u dhënë përgjigje këtyre pyetjeve, duke çuar në një saktësi më të madhe. Të dhënat e mëdha mund të japin përgjigje më të mira dhe më të vlefshme, të dyja që rezultojnë në uljen e rrezikut dhe rritjen e përgjigjeve të duhura për pyetjet .

Duke përdorur big data, nuk është e mundur vetëm të përmirësohet performanca, por është gjithashtu e mundur të matet më mirë performanca. Kjo është e mundur nga prodhimi i një informacioni më të saktë dhe të detajuar të performancës p.sh. karakteristikat e inventarit, por edhe mungesën e punonjësve. Meqenëse ky informacion tani është i njohur nga organizata, krijohen mundësi për ta shfrytëzuar më mirë këtë informacion. Kështu, është e mundur të marrin vendime më të mira menaxheriale.

2.4.3 Optimizimi i proceseve

Optimizimi i proceseve dhe produkteve të big data u mundësojnë organizatave që të zbulojnë cilat produkte specifike u nevojiten një grupi të caktuar. Në këtë mënyrë është e mundur të harmonizohen nevojat dhe nevojat e klientëve me proceset e organizatës, gjë që bën të mundur lidhjen e klientëve më gjatë me organizatën. Krijohet marrëdhënie afatgjatë me klientin. Kjo pasqyrë që na ofrojnë big data për sjelljen e klientëve gjithashtu mundëson shërbim më të mirë nëpërmjet produkteve dhe shërbimeve të personalizuara. Duke matur modele të blerjes së konsumatorëve, për shembull nëpërmjet përdorimit të mediave sociale, organizatat mund të përcaktojnë cilat metoda janë më efektive për t'i shërbyer klientëve. Për më tepër, me anë të croëdsourcing, big data mund të çojnë në zhvillimin e produktit ose në identifikimin e nevojave të klientëve për biznese të reja. Kështu, organizatat kuptojnë më mirë heterogjenitetin e klientëve ose preferencat personale të tyre.

Përpos kësaj, është e mundur që të optimizohen proceset e biznesit me anë të big data kemi më shumë njohuri mbi zinxhirin e prodhimit. Kjo mund të çojë në më shumë njohuri në lidhje me atë se në çfarë pikë krijohen kosto të panevojshme. Me anë të këtyre diagnozave komplekse, është e mundur të marrim njohuri më të thella në proceset dhe për pasojë marrim vendime më të mira. Kjo çon në më shumë mundësi për zgjidhjen e problemeve dhe optimizimin e proceseve .

3. Kompanitë shqiptare dhe Big Data

Ky kapitull analizon big data në Shqipëri dhe këndvështrimin e kompanive që operojnë në Shqipëri në lidhje me big data. Duke qenë se big data është fenomen relativisht i ri, studimet që janë bërë në big data janë kryesisht studime teorike. Për këtë arsye, vetëm shumë pak studime të gjetura në literaturën për analizimin e big data janë studime praktike. Gjithashtu, është investiguar shumë pak roli i menaxhimit në big data dhe bashkëveprimi i tyre. Për këtë arsye, qëllimi i këtij kapitulli për analizimin e big data do të jetë studimi i kompanive dhe se si ato kanë ndërthurur metodat e menaxhimit me këtë fenomen të ri. Kjo do të na japë një pamje më të qartë se drejt çfarë duhet të tentojnë kompanitë që do të përballen me menaxhimin e të dhënave masive në të ardhmen.

3.1. Metoda e studimit

Në këtë studim është përdorur hulumtimi cilësor (kualitativ). Arsyeja e zgjedhjes së kësaj metode është fakti që literatura mbi big data është fokusuar në teorizimin e fenomenit. Pra, nuk ka studime të mjaftueshme kualitative dhe është parë me vend fokusimi në praktikitetin e big data. Studimi ekzaminon literaturën ekzistuese në një mënyrë që ndihmon vendosjen në praktike të literaturës teorike. Studimi kualitativ fokusohet në mbledhjen e të dhënave specifike nga një grup i vogël. Bazuar në këtë, rezultatet teorike janë marrë dhe janë ilustruar me shembuj konkretë në influencën e big data në menaxhim. Studimi cilësor fokusohet në natyrën e realitetit, por parë nga një këndvështrim social, në mënyrë që të modelojë proceset që lidhen me këtë fenomen, duke nënvijëzuar kufizimet e situatës së hulumtimit. Qëllimi i studimit cilësor nuk është të adoptojë ligjet e përgjithshme, por të investigojë efektet e ndryshme. Është e rëndësishme të zgjedhesh një strategji hulumtimi që u jep përgjigje pyetjeve të studimit dhe çështjeve të hedhura për diskutim. Pra, është e nevojshme zgjedhja e metodës që përmbush objektivat e studimit.

Pra studimi fokusohet në intervista me menaxherë të organizatave në Shqipëri. Marrëdhënia e kompanive që operojnë në Shqipëri ekzaminohet nëpërmjet intervistave të kryera me menaxherë. Një studim kualitativ është i përshtatshëm për studime në fusha të cilat ka hulumtime të limituara. Gjithashtu hulumtimi kualitativ është i përshtatshëm në studime që investigojnë situata komplekse.

Kështu, hulumtimi cilësor (kualitativ) është mëse i përshtatshëm për studimin e marrëdhënieve të kompanive me big data. Kjo për arsye se në Shqipëri këto studime janë të cunguara dhe big data është një fushë komplekse studimi jo vetëm në rastin e Shqipërisë, por edhe në shkallë globale.

Në këtë studim, janë marrë në intervistë dy organizata. Njëra organizatë me përmasa më të mëdha, dhe tjetra kompani më e vogël. Kjo zgjedhje është bërë për të vendosur në kontrast dy tipe kompanish dhe për të krahasuar ndryshimet mes manaxhimit të dy kompanive me madhësi të ndryshme. Gjithashtu, marrja në intervistë e dy kompanive të ndryshme i jep më tepër vlerë praktike studimit. Rezultatet janë të aplikueshme jo vetëm në kompani të mëdha. Kompanitë më të vogla mund të nxjerrin vlerë në manaxhimin e big data.

3.2 Mbi teknikalitet dhe intervistat

Nga rishikimi i literaturës më sipër, janë dhënë avantazhet e big data dhe sfidat e tyre. Ashtu dhe si u përmend, studimet praktike në Shqipëri janë të pakta ose nuk ekzistojnë. Rasti i Shqipërisë është edhe më sfidues sepse ka mungesë të theksuar specialistësh të të dhënave dhe është shumë më e vështirë manaxhimi i të dhënave. Nga kjo na lind pyetja: a mundet për shkak të sfidave të tregut shqiptar të kemi përfitim të mirëfilltë nga big data? Për këtë arsye, qëllimi kryesor i intervistave ishte nxjerrja në pah e sfidave të shtuara të kompanive që operojnë në Shqipëri.

Të dhënat nga intervista janë mbledhur me anë të intervistave F2F (përballë). Intervistat përballë është një nga metodat e hulumtimit cilësor. Qëllimi i intervistave është të marrë mendime dhe perceptime të të intervistuarve dhe ka avantazhin se është e detajuar. Llojet e intervistave që mund të bëheshin ishin të strukturuar, gjysmë të strukturuar dhe të pastrukturuar. U zgjodhën intervistat gjysmë të strukturuar sepse këto lloj intervistash japin rezultate më të detajuara. Në rastin e intervistave gjysmë të strukturuar, intervistuesi ka një list me pyetje të përgatitura paraprakisht, por gjatë intervistës disa pyetje mund të hiqen dhe të tjera mund të shtohen në varësi të nevojave që lindin. Duke planifikuar pyetjet intervistuesi kap të gjitha aspektet e rëndësishme në studimin e tij. Duke patur lirinë për të modifikuar pyetjet dhe për të shtuar ose hequr pyetje sipas nevojës, merr më tepër të dhëna dhe eksploron aspekte që ndoshta nuk ishin menduar më parë.

Ky është një studim deduktiv. Në këtë studim, hulumtimi bazohet në përfundime teorike të nxjerra përpara intervistave dhe mbledhjes së të dhënave. Pyetjet janë bazuar në pyetje të përgjithshme për big data në lidhje me sfidat e tyre dhe përfitimet. Këto pyetje ishin baza e të dy intervistave. Të gjitha pyetjet janë inçizuar me lejen e të intervistuarve në mënyrë që intervistuesi të analizojë të dhënat me më tepër saktësi, dhe të mos mbajë shënime gjatë intervistës, gjë që mund të humbasë fokusin e intervistuesit.

3.3 Mbi kompanitë

Në Shqipëri lidhja e big data me menaxhimin e kompanisë është ekzaminuar shumë pak ose aspak. Është vënë re një mungesë e njohurive praktike në lidhje me big data. Edhe kur ka njohuri në lidhje me të, të pakta janë përpjekjet për të kuptuar influencën e big data në menaxhimin e kompanisë. Për të kuptuar më tepër marrëdhënien e menaxhimit të kompanisë me big data, janë marrë në intervistë dy punonjës të kompanive që janë pjesë e stafit menaxherial dhe kanë marrëdhënie direkte me menaxhimin e të dhënave në kompani. Duke studjuar kompanitë shqiptare dhe big data kuptojmë edhe më mirë problemet që kanë kompanitë kur vjenë puna e menaxhimit të të dhënave dhe kuptojmë më mirë se si t'i zgjidhim praktikisht këto probleme. Ishte mjaft e rëndësishme zgjedhja e dy kompanive të madhësive të ndryshme për të vendosur në kontrast dy rastet dhe për të diversifikuar rezultatet e nxjerra në fund të studimit.

Ndryshimi i kompanive nuk konsiston thjeshtë dhe vetëm në madhësinë e tyre, por edhe në ndryshime të tjera si fusha e operimit, industria dhe avancimi teknologjik. Njëra kompani përdor ERP, kompania tjetër nuk ka sistem të avancuar ERP. Njëra kompani ka më pak se 70 punëtorë dhe tjetra ka më shumë se 600. Si përfundim, mund të themi se dy kompanitë kanë kultura të ndryshme të menaxhimit të të dhënave.

Kompania e parë, me mbi 600 punëtorë dhe me sistem ERP, procesonte sasi të mëdha të dhënash dhe kishte filluar përpjekjen për akomodimin e këtij fenomeni duke avancuar në sistemin ERP. Menaxherët e shikonin potencialin e të dhënave dhe e kanë thelbësor nxjerrjen e vlerës nga të dhënat, por edhe menaxhimin e të dhënave. Kompania e parë kishte implementuar një sistem të ri ERP në gjithë entitetet e organizatës dhe kjo ndihmonte në mbledhjen e të dhënave. Pra kishin filluar përpjekjet dhe eksplorimet e big data në vlerën që mund t'i shtojë kompanisë.

Në kontrast, kompania e dytë, më e vogël dhe më pak e avancuar nga ana teknologjike, jo se nuk kishin njohuri në lidhje me big data dhe me rëndësinë e menaxhimit të të dhënave, por nuk e kishin të mundur të përballonin kostot dhe kompleksitetin e nxjerrjes së vlerës nga big data.

Edhe implementimi i ERP ishte i vështirë për kompaninë në fjalë, pikërisht për kostot dhe kompleksitetin e saj. Kompania e dytë e kuptonte plotësisht rëndësinë e big data. Ishin edhe në dijeni të faktit që për të nxjerrë maksimumin nga të dhënat duhej implementuar ERP, por kostot për ta bërë këtë të fundit ishin shumë të mëdha për kapacitetin e kompanisë.

Megjithatë, pavarësisht ndryshimit të kompanive nga njëra tjetra, të dyja e kuptonin plotësisht rëndësinë e big data, thjesht njëra prej tyre kishte mundësitë

dhe burimet për të bërë dicka në lidhje me fenomenin, kurse tjetra ishte në hapat e para drejt menaxhimit të të dhënave.

3.4 Rezultati i studimit

3.4.1 Sistemet ERP dhe trendi në të dhëna

Intervistat me dy kompanitë e ndryshme sugjerojnë se kompanitë në shqipëri, në përgjithësi punonjësit dhe menaxherët sa po vjen dhe po shtojnë përdorimin e big data. Për më tepër, intervistat sugjerojnë se në të dyja kompanitë po bëhen përpjekje të vazhdueshme për të formuar kulturën e duhur të manaxhimit të big data dhe për zhvillimin në këtë fushë. Për t'ia arritur kësaj, njëra prej kompanive (kompania 1) ka filluar të merret me projektë për t'i dhënë të dhënave rëndësi të vecantë në organizatë.

Kompania 1 ka vënë re se në disa raste sistemi ERP që përdorin dështon në procesimin e të dhënave sepse dizajni i sistemit s'është i duhuri për ta përballuar sasinë e madhe të të dhënave. Për këtë arsye kompania po mendon të blejë një sistem të ri ERP, që të kenë mundësi të ruajnë më tepër të dhëna, në mënyrë që të prodhojnë informacion prej tyre. Në këtë moment kompania 1 po përditësonte (bënte update) sistemin e të dhënave në kompani, dhe më tej në projekt është planifikuar optimizimi i proceseve të organizatës që ta përshtatin me sistemin e ri ERP. Kompania donte të sigurohej se të dhënat do të bëheshin më uniforme dhe të kishin mundësi të ruanin më me lehtësi këto të fundit.

Për më tepër, kompania një njëkohësisht po implementonte CRM në gjithë shtrirjen e organizatës. Ky proces pritej të përfundonte brenda disa vitesh, por na tregon për organizatën se po bën përpjekje të vazhdueshme për t'u angazhuar në manaxhimin e big data dhe për të përballuar trendin eksponencial të rritjes së tyre në vitet në vazhdim. Gjithashtu kompania po punonte me databazën qendrore çka kombinon të dhëna nga sisteme të ndryshme që organizata ka përdorur si në të tashmen ashtu edhe në të shkuarën. Ky kujdes kundrejt databazës shërben si bazë për rritjen e big data në të ardhmen.

Kompania 2, që është më e vogël në madhësi, ka parë se sistemi i vjetër i ERP s'është i mjaftueshëm për sasinë e të dhënave që kanë. Menaxhimin e këtij sistemi e kanë bërë me anë të outsourcing sepse nuk kanë ekspertizën e duhur dhe as mjetet në kompani për t'u kujdesur për të. Megjithatë kompania planifikon që në të ardhmen të konsolidojë sistemin dhe të implementojë një CRM në kompani. Kjo na tregon se si kompania e parë, ashtu edhe e dyta e dinë rëndësinë e

rritjes së të dhënave dhe të avantazheve që u krijojnë këto të fundit dhe bëjnë planifikime afatgjata për implementimin e sistemeve të duhura për të përballuar të dhënat në shtim e sipër.

3.4.2 Këndvështrimi i kompanive lidhur me big data

Për dy organizatat, projektet që po ndërmarrin ose që po planifikojnë të ndërmarrin në të ardhmen, shihen si ndryshime të mëdha dhe drastike për organizatat. Kostot e tyre janë të shumta, si të krijimit apo blerjes së sistemeve, ashtu edhe mirëmbajtja e tyre. Megjithatë, kompanive u ka lindur nevoja për këto ndryshime dhe menaxherët janë të ndërgjegjshëm për rëndësinë e ndryshimit. Kompanitë e kanë vënë re që të dhënat po bëhen sa vjen edhe më të rëndësishme dhe sasia e tyre po shtohet masivisht. Sipas tyre, janë ndryshime progressive që bëhen në treg që nëse kompanitë nuk marrin masa, janë të destinuara të dështojnë. Për shembull, dy kompanitë e intervistuar kanë parë se konkurentët e tyre po i rrisin sa vjen e më shumë përpjekjet për të menaxhuar të dhënat, dhe nëse ata nuk ndjekin ritmin rrezikojnë të dështojnë.

“Edhe pse Shqipëria nuk është në nivelet e një tregu shumë të avancuar krahasimisht me shtetet e tjera të fuqishme, kompanitë kanë filluar të ndjejnë nevojën e menaxhimit efektiv të të dhënave. Nëse një konkurent përdor një sistem të avancuar je i detyruar ta inkorporosh dhe ti nëse do të kapësh ritmin.” - Intervistuesi (Kompani 1)

Shtetet më pak të zhvilluara edhe pse me ritme më të ngadalta ecin drejt një fenomeni global. Kompanitë përpiqen të sjellin risi për të patur avantazh dhe për të shfrytëzuar mundësitë që lindin. Këto kompani kanë ndikim në treg dhe te kompani të tjera që nëse duan të mos falimentojnë janë të detyruara të ecin me risitë që vinë në treg.

“Ka ikur koha kur mbaheshin shënimet me laps edhe letër. Kur them laps edhe letër e kam fjalën dhe për programe ose teknologji të vjetra. Kompania jonë dhe menaxherët janë gjithmonë në përpjekje për të gjetur burime për të automatizuar proceset. E dimë se çfarë avantazhesh sjell e gjithë kjo.” – Intervistuesi (Kompani 2)

3.4.3 Vendimarrja dhe big data

Ndryshimi kryesor që pashë mes kompanive ishte marrëdhënia që kishin të dhënat me vendimarrjen e tyre. Teksa kompania e parë e kishte të rëndësishme vendimarrjen bazuar në të dhënat e mbledhura, e dyta nuk e përdorte gjerësisht. Këtu mendoj kemi hendekun kryesor mes mënyrës tradicionale të përdorimit të të dhënave dhe mënyës moderne përdorimit të

tyre. Përpara disa vitesh, kompanitë prodhonin dhe sistemet shërbenin vetëm për suport. Sot, kjo marrëdhënie ka ndryshuar: sistemet e kompanive janë baza e kompanisë, dhe në bazë të informacionit që marrim prej të dhënave merren dhe vendimet për prodhimin.

Kompania e parë ka burimet e nevojshme që të marrë vendime në bazë të të dhënave dhe në bazë të sistemeve që procesojnë dhe analizojnë të dhënat. Kjo kompani ka ndryshuar proceset brenda saj për t'iu përshtatur kësaj kulture. Për shembull, kompania nuk përdor më lista të printuara nga kompjuteri, por kanë krijuar një mjedis dixhital që është kthyer në bazën e operacioneve të kompanisë. Përpara këtij dixhitalizimi, cdo gjë bëhej manualisht gjë që harxhonte shumë kohë sepse gjithë përpunimi bëhej me dorë. Mbledhja e të dhënave bëhej me dorë dhe po ashtu edhe nxjerrja e informacioni nga këto të dhëna. Përmes zhvillimeve teknologjike kjo situatë u përmirësua për shumë kompani. Me kalimin e kohës dhe menaxherët dhe punonjësit u informuan dhe u ndërgjegjësuat për rëndësinë e të dhënave. Ata tani njohin më mirë mundësitë që të ofrojnë big data.

“Janë bërë ndryshime marramendëse nga e shkuara në të tashmen, tani e dimë më mirë sa të rëndësishme janë të dhënat jo vetëm për të mbështetur prodhimin dhe shërbimin, por kryesisht për të marr vendime në lidhje me çfarë do prodhojmë apo ofrojmë. Firma jonë në shumë raste nuk mjaftohet me të dhënat që gjenerohen gjatë proceseve. Bëjmë outsourcing kompani të tjera që të mbledhin të dhëna për ne. Ndonjëherë analizat i kryejme vetë e në disa raste marrim edhe raporte të gatshme, sidomos në studimet cilësore.” – Intervistuesi (Kompania 1)

“Ndonjëherë paguajmë kompani të tjera që të ba bëjnë studim tregu. Kjo nuk është diçka që bëhet gjithmonë, s'bëhet në mënyrë periodike, vetëm kur është shumë e nevojshme. Përfitjmë nga ekspertiza e kompanive që merren enkas me mbledhje të dhënash dhe studim tregu. Kostot janë të përballueshme dhe më të vogla nga çfarë do të ishin nëse do ta bënim vetë studimin. S'jemi gati për ta bërë vetë studimin e tregut. Por e dimë rëndësinë e të dhënave dhe të studimit të tregut, kuptohet.” – Intervistuesi (Kompani 2)

Edhe pse nuk mund të flitet për big data të madhësisë së kompanive si Facebook apo Google edhe kompanitë Shqiptare janë ndërgjegjësuat për rritjen e sasisë, shpejtësisë dhe volumit të të dhënave. Edhe kompanitë Shqiptare po e ndjejnë fenomenin edhe pse në shkallë më të vogël. Megjithatë kompanitë gjithashtu kuptojnë se nuk është e thjeshtë menaxhimi i big data. Në dy organizatat projektet janë afatgjata dhe duan shumë kohë të arrihen objektivat. Duhet kuptuar gjithashtu, që në shumicën e kompanive në Shqipëri, të dhënat sjajn ende informacion, përveçse në rastet kur flasim për

kompani të studimi të tregut (Market Research). Dhe te këto të fundit besueshmëria e të dhënave është në diskutim. Kompanitë gjithashtu e kuptojnë se të dhënat pa u kthyer në informacion nuk shfrytëzohen në kapacitetin e tyre të plotë. Të dhënat sipas tyre nuk janë qëllimi në vetvete, por mjeti për të arritur qëllimet.

3.5 Përmirësimet e mundshme

Cdo kompani ka specifikat e veta dhe secila ka nevojë për përmirësime specifike, megjithatë janë përmirësime të aplikueshme në pothuajse gjithë kompanitë të cilat do diskutohen në sektorët në vijim. Kompanitë e intervistuar, edhe pse të ndryshme si në fushën e veprimit dhe në madhësi dhe organizim, e dinin shumë mirë rëndësinë dhe nevojën e manaxhimit të të dhënave dhe big data. Këto kompani e dinë gjithashtu se duhen zhvillime në sistemet e tyre dhe përmirësime të tyre.

3.4.1 Përmirësime në marrëdhënie me klientin

Mundësitë që të ofron menaxhimi i të dhënave në raport me klientin janë të pafundme dhe me rëndësi të vecantë për kompanitë. Big data të bën të njohësh më mirë klientin dhe t'i përgjigjesh nevojave të tij më shpejtë. Qëllimi është të kuptosh çfarë është në mendjen e konsumatorit, të parashikosh dëshirat dhe nevojat e tij.

Për shembull, me anë të një sistemi CRM është e mundur të marrësh informacion për konsumatorët strategjik. Me këto sisteme bëhet e qartë ku janë avantazhet më me vlerë për konsumatorët dhe i mban klientët të kënaqur duke i mbajtur të lidhur me kompanitë njëkohësisht. Me rritjen e të dhënave rriten edhe mundësitë për organizatat në treg. Një mënyrë se si mund të arrihet kjo është duke parë në databazë se çfarë kërkon klienti. Konsultohesh me databazën për kërkesat e tij dhe duke u bazuar te profili i tij dhe te të dhënat historike gjen specifikisht çfarë kërkon ky klient.

Kuptojmë nga shembulli i mësipërm se si mund të parashikojmë se çfarë kërkon klienti për ta përdorur në të ardhmen. Kjo ndikon dhe në proceset e organizatës. Duke u bazuar te klientët krijon paraprakisht procese bazuar te ta.

3.4.2 Përmirësime në vendimarrje

Intervistat me dy kompanitë e bënë edhe me të qartë faktin se shikonin mundësi përmirësimi në procesin e vendimarrjes me anë të big data. Uniformiteti në të dhëna ishte shumë i rëndësishëm për ta. Uniformiteti

në të dhëna sjell uniformitet në raporte dhe e bën më të thjeshtë krahasimin e raporteve të mëparshme me raportet aktuale.

Përvec pjesës së uniformitetit të të dhënave, mbledhja e të dhënave me anë të sistemeve më të shpejta dhe të besueshme ndihmon në vendimarrje. Me anë të këtyre sistemeve punonjësit nuk kanë pse të mbledhin të dhëna manualisht dhe ti ruajnë të dhënat me shkrim. Për ta lidhur me vendimarrjen, kjo mënyrë e automatizuar e mbledhjes së të dhënave i bën këto të dhëna më të besueshme e për pasojë dhe vendimet e bazuara në to më të sakta.

Punonjësit falë dixhitalizimit kursejnë kohë dhe mund ta përdorin kohën për proceset e tjera që lidhen me të dhënat, si përshebull analizimi.

Për më tepër, përdorimi i të dhënave sjell përfitime dhe në nivelin operacional që bazohet në punët e përditshme të organizatës. Avantazhi i dixhitalizimit të proceseve në lidhje me të dhënat është që menaxherët marrin vendime më shpejtë. Për shembull, kur ka një problem me funksionimin e kompanisë, të dhënat i japin një ide se çfarë mund të mos shkojë si duhet dhe menaxheri merr vendime më shpejtë.

Të dhënat sjellin edhe përfitime operationale duke përmirësuar komunikimin mes entiteteve të ndryshme. Tradicionalisht, kompanitë kishin sisteme të ndryshme për entitete të ndryshme. Për pasojë sistemet e entiteteve të ndryshme nuk ishin të lidhur me njëri-tjetrin. Entitete të ndryshme nuk e dinin se çfarë ndodhnin të entitetet e tjera. Duke përdorur sistemet e duhura entitetet kanë komunikimin më të thjeshtë mes tyre dhe kjo ndihmon jo vetëm në përsheptimin e proceseve, por edhe në një vendimarrje më të mirë. Duke përdorur të dhënat, kemi edhe standartizim të proceseve. Duke përdorur big data bëhet më e lehtë të analizosh të dhënat dhe përdorimi i big data bën analizat më “inteligjente”. Standartizimi bën që sistemet të kenë njëfarë autonomie dhe të arrijnë të nxjerrin zgjidhje inteligjente me anë të të dhënave që kanë në dispozicion.

4. Sugjerime teknologjike dhe çështje etike të big data

Në sektorët e kaluar është rishikuar literatura në lidhje me big data, është trajtuar problemi i rritjes eksponenciale të big data dhe është studjuar marrëdhënia e kompanive shqiptare me këtë trend. Nga secili prej sektorëve dalim në një përfundim: big data janë sfiduese, kanë përfitimet e tyre dhe më kryesorja mënyra se si menaxhohen big data janë me anë të teknologjisë. Në këtë kapitull, do të trajtohen disa teknologji që janë krijuar enkas për menaxhimin të big data. Këto teknologji janë thelbi i menaxhimit të big data, pa

to nuk do të kishim përfitimet e big data në masën që kanë sot kompanitë që kanë arritur t'i implementojnë me sukses këto teknologji.

4.1 Big Data Hadoop

Hadoop është një mjet (tool) open source nga Apache Software Foundation (ASF). Kur flasim për open source e kemi fjalën për programe të cilat kanë kodin burim të hapur ku të gjithë mund ta modifikojnë. Nëse një funksion i caktuar nuk i përmbush nevojat e tua ti mund ta ndryshosh kodin që t'i përshtatet. Shumica e kodit të Hadoop është shkruar nga Yahoo, IBM, Facebook dhe Cloudera. Është një framework eficient për të kryer procese në nyje të clusters. Clusters dmth grupe sistemesh të lidhura me LAN. Hadoop ofron procesim paralel.

Hadoop përbën më shumë se 90% të tregut të big data. Le të eksplorojmë pak pse është kaq popullor. Hadoop nuk është thjeshtë dhe vetëm sistem ruajtës të të dhënave, por është një platformë për ruajtje dhe procesim të dhënash. Është një sistem i shkallëzuar pra më shumë nyje mund të shtohen në “fly. Është tolerant ndaj gabimeve dhe ashtu si e përmendëm më parë është open source (kodi modifikohet sipas nevojave) [Flair, T. (2016)].

Arkitektura e hadoop

Hadoop punon sipas arkitekturës pardon – skllav (master-slave). Ka një nyje padron dhe një numër të caktuar nyjesh skllav që janë nyje që realisht punojnë. Nyja padron duhet të mbështetet në konfigurimin e hardëare dhe jo thjesht në komoditete të hardëare. Kjo është e rëndësishme sepse nyja padron është qendra e Hadoop cluster [Flair, T. (2016)]. Nyja padron ruan meta datat, që janë të dhënat e të dhënave, kurse nyjet skllav janë nyjet që ruajnë të dhënat. Të dhënat ruhen dhe janë të shpërndara në cluster. Klienti lidhet me nyjen padron për të bërë cdo punë.

Sistemi i shpërndarë i të dhënave Hadoop (HDFS)

Hadoop distributed file system është hapësira kur ruhen të dhënat në Hadoop në mënyrë të shpërndarë (distributed). Në nyjen padron një daemon që quhet namenode ekzekutohet për HDFS-në. Nyjet skllav thërriten si datanode (nyje të dhënash). Namenode-et ruajnë të dhënat metadatat dhe menaxhon datanodet teksa datanode-et ruajnë të dhëna dhe kryejnë detyrën reale.

Hadoop MapReduce

Hadoop MapReduce është shtresa që kryen procesimin e Hadoop. MapReduce është model programimi i dizenuar për të procesuar volumet e mëdha të të dhënave në paralel duke e ndarë punën në sete punësh të pavarura. Duhet të vendosësh logjikën e biznesit në MapReduce dhe ajo është e aftë të procesojë të dhënat në mënyrën e duhur në bashkëpunim me framework. Puna e përfundur pasohet nga përdoruesi të nyjet padron duke u ndarë në detyra të vogla dhe pastaj nga padroni pasohen respektivisht të nyjet skllav.

4.2 Data Mining

Në kapitujt më sipër folëm për llojet e të dhënave dhe se si një problem janë të dhënat të tipeve të ndryshme si të dhëna nga sensorë, të dhëna në formën e fotove, të dhënat tekst. Kombinimi i të gjitha këtyre na jep të dhëna heterogjene të dhëna të grupuara së bashku, por të natyrave të ndryshme. Për procesimin dhe menaxhimin e këtyre të dhënave na ndihmon data mining. Data mining përfshin eksplorimin dhe analizimin e të dhënave të sasive të mëdha. Inteligjenca Artificiale dhe statistika janë fusha të cilat i zhvillojnë këto teknika. Ky seksion eksploron karakteristikat e aplikimeve të procesimit të big data duke e parë nga këndvështrimi i data mining.

Data mining për Big Data

Kryesisht, qëllimi i data mining është ose klasifikimi ose parashikimi. Në klasifikim, data mining i ndan të dhënat në grupe. Për shembull, një marketer mund të ndajë konsumatorët që e pëlqejnë një shërbim të caktuar dhe ata që nuk e pëlqejnë.

Në parashikim, data mining parashikon vlerën e një variabli të caktuar. Për shembull, në këtë rast marketeri do të parashikojë se sa persona do e pëlqejnë një produkt.

Algoritmet që përdoren në data mining vijnë si më poshtë [Ghoting A. and Pednault, E. (2009)]:

- Klasifikimi i pemëve - Një teknikë e data-mining që përdoret për të klasifikuar një variabël të kategorizuar bazuar në matjet e një ose më shumë variablave parashikues. Rezultati është një pemë me nyje dhe lidhje mes nyjesh që mund të lexohen për të formuar rregullat e kushtëzimit if dhe then.
- Regresioni logjik – Një teknikë statistikore që është variant i regresionit standart por zgjeron konceptet që lidhen me klasifikimin.

Prodhon një formulë që parashikon probabilitetin e ndodhjes së një ngjarje.

- Rrjetet neutralë – Një algoritëm software që modelohet pas arkitekturës paralele. Rrjeti konsiston në inputet e nyjeve, shtresave të fshehura dhe nyjeve output. Çdo njësi ka një peshë. Të dhënat cohen te input node dhe algoritmi përshtat peshat derisa ta ndalojë një kriter i caktuar.
- Teknikat clustering si K- Një teknikë që identifikon grupet e rekordeve të ngjashme. Teknika K e clustering përllogarit distancën midis rekordeve dhe drejton rrjedhën të të dhënave historike.

Një shembull i pemëve të klasifikimit do të ishte rasti i një situate të një kompanie telefonike. Marrim rastin kur kjo kompani do të përcaktojë cilët janë klientët që banojnë brenda qytetit dhe duan të ndërpresin lidhjen telefonike.

Kompania telefonike ka informacion për atributet e mëposhtme: për sa kohë e ka përdorur klienti këtë shërbim, sa shpenzon për këtë shërbim, a ka pasur problematika lidhja, ku jeton klienti, çfarë moshe ka, a ka shërbime të tjera përveç shërbimit telefonik dhe nëse vazhdon ta ketë shërbimin.

Në këtë rast, të dhënat ndahen në të dhëna trajnimi dhe të dhëna test. Të dhënat e trainimit konsistojnë në observime dhe në outcome. Algoritmi ekzekutohet te të dhënat e trainimit dhe krijon një pemë që mund të lexohet si një set rregullash. Për shembull, nëse konsumatorët kanë qenë klientë të komoanusë për më shumë se dhjetë vjet dhe janë mbi 60 vjeç, gjasat janë që klienti të vazhdojë me shërbimin e kompanisë. Këto rregulla ekzekutohen te të dhënat test për të përcaktuar se sa i mirë është modeli i krijuar. Krijohet një matrice me të gjitha rastet kur rregullat kanë qenë korrekte ose inkorrekte, pra klasifikohen rastet e suksesit dhe dështimit.

Nëse modeli ka patur më shumë suksese se dështime, ky model përdoret dhe te të dhëna të tjera. Bazuar në model kompania mund të marrë vendime. Për shembull, nëse modeli i nxjerr se disa klientë kanë gjasa më të shumta për t'u larguar, kompania mund t'u bëjë oferta speciale këtyre klientëve.

4.3 Çështjet etike të big data

Duke qenë se kemi një revolucion në big data, një rritje eksponenciale të tyre, një fenomen të ri me të cilat kompanitë po përballen së bashku me përfitimet vinë dhe probleme etike në lidhje me menaxhimin e të dhënave. Pa bërë një analizë të problemeve etike, kompanitë rrezikojnë të bëhen subjekt i problemeve ligjore, por mund të ketë edhe prishje të imazhit të tyre deri në falimentim për shkak të prishjes së imazhit. Kompanitë më të mëdha po hasin probleme në çështjet e menaxhimit të të dhënave. Për të kuptuar më mire se si një kompani mund të bëhet subjekt, do të krijohet një case study për kompaninë Facebook, që është epitoma e kompanisë që menaxhon të dhëna, por që në të njëjtën kohë ka patur dhe probleme ligjore lidhur me menaxhimin e të dhënave. Raste pragmatike si rasti i Facebook na bën të shikojmë realisht se sa thjeshtë mund të bëhemi pre e

cënimet të të drejtave dixhitale të përdoruesve dhe sa probleme ligjore mund të lindin nga kjo.

4.4 Rast Studimor – Facebook, mundësitë dhe problemet e privatësisë

Facebook është një nga rrjetet sociale më të përdorur. Vetë Facebook ka deklaruar se çdo ditë sistemi i tyre proceson 2.5 miliard artikuj dhe 500+ terabyte të dhënash çdo ditë. Çdo ditë ruhen 2.7 miliardë Pëlqime (Like) dhe 300 milionë foto [techcrunch]. Mund të themi me siguri se Facebook është epitoma e simptomit të Big Data , me një sasi kaq të madhe të të dhënave që gjeneron. Duke u nisur edhe nga polemikat e fundit lidhur me Facebook dhe privatësinë e përdoruesve, mendoj se është rasti i përkryer studimor, për të kuptuar jo vetëm se si kjo kompani manaxhon të dhënat, por edhe për të shqyrtuar dilemat etike që vinë me Big Data. Në këtë studim do të shohim se si një nga kompanitë më të suksesshme në botë mbledh të dhënat tona dhe i kthen këto të dhëna në përfitim.

Ashtu si u përmend edhe më lartë, Facebook ka qenë pjesë e polemikave lidhur me Big Data dhe mënyrën se si menaxhon të dhënat e përdoruesve. Skandali nisi me të dhënat që Cambridge Analytica kishte marrë nga përdoruesit e Facebook në formën e një quizi. Këto të dhëna ishin përdorur për të bërë studim psikologjik të përdoruesve. Të dhënat ishin marrë pa leje, dhe skperimentimet psikologjike kishin ndodhur pa dijeninë e përdoruesve. [Jenkins, A. (2018)].

Me sasinë e madhe të të dhënave që Facebook mbledh nga përdoruesit, por edhe me bazën e gjerë të përdoruesve mund të kuptojmë se sa fuqi ka kjo kompani, se sa lehtësisht mund të manipulojë dhe të eksperimentojë me njerëzit dhe me sa lehtësi dhe saktësi mund të prodhojë informacion nga të këto të dhëna. Disa mendojnë, se kjo fuqi të dhënash që Facebook ka është ana tjetër e medaljes së një platforme që ka lehtësuar komunikimin mes njerëzve dhe është një çmim që duhet të paguhet për revolucionarizimin e komunikimit.

Përpos problemeve të privatësisë dhe etikës së përdorimit të të dhënave duhet pranuar se Big Data sikurse mund të keqpërdoret mund edhe të përmirësojë marrëdhënien e bizneseve dhe kosnumatorëve në treg. Motivimi kryesor i facebook nuk është përdorimi i të dhënave për manipulime psikologjike. Motivimi kryesor i përdorimit të të dhënave të Facebook janë reklamat.

Kompanitë që duan të reklamojnë në Facebook përfitojnë nga profilet e detajuara të përdoruesve dhe nga historiku i përdorimit të tyre. Si përfitojnë? Duke mbledhur të dhënat e përdoruesve, kur kompanitë reklamojnë, ato targetojnë vetëm personat që u nevojiten. Psh.: nëse një student ka kërkuar në

facebook për programe studimi master, një universitet privat që dëshiron të reklamojë programet e veta u shfaqet këtyre studentëve. Ose kompani të tjera që duan të reklamojnë produkte kozmetike targetojnë femra të moshës mbi 20 vjeç që janë të interesuara për makeup etj.

Pra të dhënat e marra nga aktiviteti i përdoruesit, shërbejnë për të krijuar profile të detajuara konsumatorësh dhe reklamën e përshtatshme dhe u shfaqen target grupeve të caktuara në bazë të profileve të sipërpërmendura.

Duke përdorur këto teknika, të ardhurat nga reklamën e Facebook u rritën me 129% nga 2011 në 2013. Një raport i “Statista” tregon të ardhurat e Facebook vetëm nga reklamën. Përpos skandaleve të fundit dhe uljes së numrit të përdoruesve, vetëm në vitin 2017 të ardhurat nga reklamën ishin 40 miliard dollarë amerikan. Kjo na tregon se sa tërheqës janë teknikat e përdorimit të Big Data për kompanitë.

Dhe nuk janë vetëm të dhënat e strukturuar që përdoren rëndom. Facebook mbledh të dhëna për mënyrën se si përdorim mouse-in gjatë kohës që gjendemi në facebook, mbledh të dhënat e tipareve tona të fytyrës nga fotot që shkarkojmë etj. dhe gjitha kjo e vendosur në funksion të reklamimit. Facebook mbledh të gjitha të dhënat që mundet, për t’u siguruar se reklamën e duhura shkojnë te grupi i duhur i konsumatorëve.

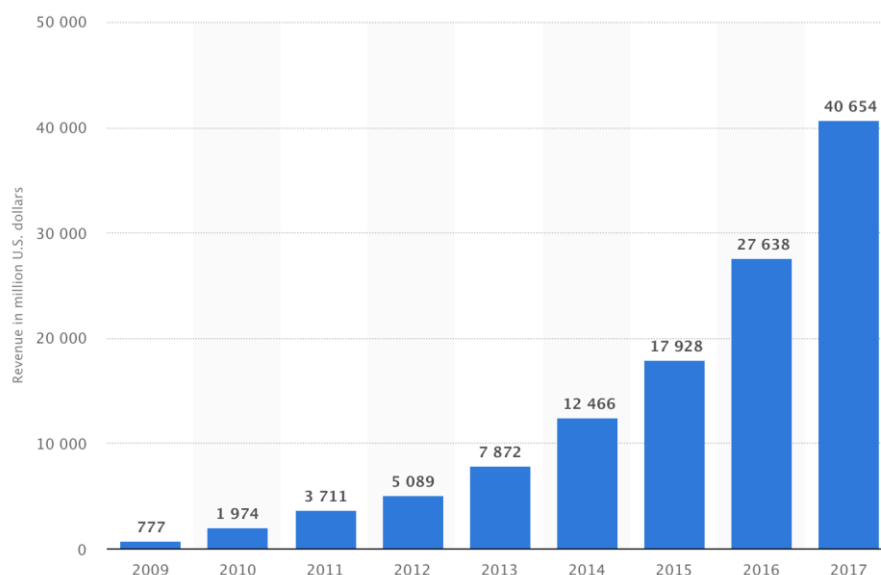


Fig 3.

Burimi: Statista (2018)

Metodat e përdorura janë kritikuar jo pak nga mediat, dhe organizata të mbrojtjes së të drejtave dixhitale si “Ranking Digital Rights” e kanë vlerësuar Facebook me indeksin 53% të sa arrin t’i mbrojtë të dhënat e përdoruesve. Sipas

kësaj organizate, Facebook nuk është i qartë mjaftueshëm te termit e përdorimit, përdoruesit nuk e kuptojnë se si përdoren të dhënat tyre.

Për ta bërë gjithë këtë proces më transparent, Facebook ka krijuar një skuadër që manaxhon të dhënat dhe ka krijuar një faqe që merret specifikisht me studimet e të dhënave nga Facebook [Harman, M. (2018)]. Atje Facebook ndan gjetje të rëndësishme sociale bazuar te të dhënat dhe metodat që përdor për të arritur në këto përfundime.

Në total, mendoj se problemet e Facebook vinë si pasojë e suksesit të madh të kompanisë dhe si pasojë e sfidës që sjell menaxhimi i Big Data. Ka mësim të vlefshme për t'u nxjerrë nga eksperiencia e Facebook. E para që na mëson janë problemet etike dhe ligjore që rrezikojnë të sjellin Big Data. Nëse nuk tregohet transparencë në procesin e përdorimit të tyre, kompanitë rrezikojnë të bëhen objekt i kritikave (që dëmtojnë imazhin e kompanisë), dhe objekt i problematikave ligjore.

4.5 Kufijtë etik të kompanive

Ligjet dhe rregullat i drejtojnë kompanitë kur vjen çështja e big data. Çdo shtet ka ligjet dhe rregullat e veta, por në një mjedis global ku shpesh operojnë kompanitë që gjenerojnë big data janë çështje etike që duhet të konsiderojë çdo kompani që menaxhon dhe analizon të dhënat. Jo çdo gjë që është e mundur të bëhet është edhe e ligjshme. Atje kur ka mundësi të mëdha për kompanitë, po aty ka dhe risqe.

Kur një kompani dëshiron që të aplikojë analiza të big data për të rritur performanën, si duhet ta dinë kompanitë se çfarë është etike?

Ajo çfarë është etike apo jo është subjektive. Ajo që për një individ është etike për dikë tjetër mund të mos jetë. Individët janë ata që përcaktojnë cila sjellje organizative është apo nuk është etike. Kush e përcakton se çfarë është e drejtë apo jo?

Për vetë faktin se çështjet etike janë subjektive dhe shumë persona kanë këndvështrime të ndryshme, kompania duhet të jetë e kujdesshme. Kjo sugjeron se organizatat duhet të konsultohen gjerësisht në mënyrën si përdorin teknologjitë dhe të krijojnë politika që argumentohen dhe shpjegohen shumë mirë për arsyet pse janë krijuar.

Kompanitë duhet të kujdesen për [Chessel, M. (2014)]:

- **Kontekstin** – Me çfarë qëllimi janë mbledhur të dhënat? Me çfarë qëllimi po përdoren të dhënat? Sa larg kontekstit janë analizat dhe mbledhjet e të dhënave? A është gjëja e duhur për t'u bërë?

- **Miratimi dhe Zgjedhja** – Çfarë zgjedhjesh u lihen përdoruesve dhe klientëve? A e dinë këta të fundit se po bëjnë një zgjedhje? A e kuptojnë ata se për çfarë po bien dakord? A kanë ndonjë mundësi për të refuzuar? Çfarë alternativash ofrohen?
- **Arsyeja** - A është thellësia e përdorimit të të dhënave dhe mënyra si përdoren ato të arsyeshme?
- **Pronësia** – Kush i ka në pronësi rezultatet e analizave? Cilat janë përgjegjësitë e personave që i kanë në pronësi?
- **Konsiderimi** – Cilat janë pasojat e të dhënave?
- **Aksesi** – Kush ka mundësinë t'i aksesojë të dhënat?
- **Përgjegjësia** – Si gjenden difektet dhe si mund të riparohen ato? Kush është përgjegjës për këto difekte?

Duke marrë në konsideratë të gjitha këto na krijohen një framework i ndërgjegjshmërisë së etikës. Ky framework është krijuar nga Mbretëria e Bashkuar dhe nga Ireland Technical Consultancy Group (TCG) për të ndihmuar njerëzit të krijojnë politika etike për përdorimin e të dhënave.

Duke qenë të vëmendshëm dhe transparentë në mënyrën si përdorim të dhënat, kompania do t'i limitojë risqet që lidhen me çështjet etike të përdorimit të big data. Me teknologji të mira, me një kulturë të mirë organizate dhe me burimet e nevojshme, duke patur parasysh pyetjet që u ngritën më sipër mund të themi se marrim të mirat nga big data duke limituar riskun për të patur probleme ligjore dhe etike.

5. Konkluzione dhe sugjerime

Me rritjen e avancimeve teknologjike dhe me rritjen e përdoruesve të rrjeteve sociale dhe teknologjisë, po ndodh një revolucion me të dhënat. Ato po shtohen çdo ditë e më shumë, me sasi marramendëse. Ky fenomen i rritjes së të dhënave ne vëllim, shpejtësi dhe sasi të mëdha quhet big data. Big data, parë nga këndvështrimi i kompanive, krijon mundësi të reja në treg, por në të njëjtën kohë edhe sfida të pa ndeshura më parë.

Nga studimet më të fundit, ka dalë se të dhënat po rriten në mënyrë eksponenciale. Nga gjithë të dhënat historike që janë mbajtur, 90% e tyre janë krijuar vetëm dy vitet e fundit. Pritet që numri i të dhënave të shkojë në 40 zetabyte brenda vitit 2020. Kjo rritje me ritme kaq të shpejta është një fenomen i ri që kërkon mënyra jo konvencionale trajtimi. Sigurisht, për të akomoduar këto të dhëna na ndihmojnë teknologjitë e reja që janë krijuar si Data Hadoop apo Data mining.

Por sic e thamë, big data krijojnë sfida, jo vetëm për kompanitë e mëdha të përmasave të Google apo Facebook, por edhe për kompani më të vogla që duan të jenë

konkuruese në treg. Sfidat më të mëdha të kompanive lidhen me sasinë e madhe të të dhënave të pastruara, mungesën e shkencëtarëve të të dhënave dhe kostot shtesë për t'i tërhequr, analizat në kohë reale dhe shpejtësia me të cilën duhet të procesohen të dhënat, çështje të sigurisë së të dhënave dhe kultura e organizatës me aspektin njerëzor, që duhet t'i përshtaten këtyre ndryshimeve.

Përpos sfidave, u fol edhe për përfitime nga big data. Nga të dhënat mund të prodhohet informacion dhe ky informacion i ndihmon kompanitë me vendimmarrjen, reduktimin e kostove dhe përmirësimin e proceseve. Duke e njohur më mirë klientin dhe duke patur të dhëna për proceset brenda organizatës, kompania ka më tepër shance që të ketë sukses. Gjithashtu, me anë të big data mund të identifikohet klientë potencialë dhe të kryejmë vendimmarrje me në qendër konsumatorin.

U fol se big data është fenomen global, po në Shqipëri a mund të flitet për big data? Nëse po deri në çfarë mase? Për t'i dhënë përgjigje këtyre pyetjeve u bë një studim kualitativ i kompanive shqiptare me anë të intervistave direkte me menaxherët e kompanive. U kuptua se kompanitë shqiptare edhe pse në një masë më të vogël kishin filluar të ndjenin nevojën për menaxhimin dhe analizimin e sasive të mëdha të të dhënave. Gjithashtu kuptuam se kompanitë janë të ndërgjegjshme për rëndësinë e të dhënave dhe kanë ndërmarrë projekte afatgjata që të kenë mundësi të zhvillojnë sisteme të përshtatshme për kompanitë (ERP më të mira).

Pas intervistave me kompanitë shqiptare u kuptua më mirë se ku qëndronin problemet e tyre dhe u sugjeruan përmirësime përkatëse dhe teknologji që do e përmirësonin mënyrën e menaxhimit të të dhënave. Këto teknologji përfshijnë Data Hadoop dhe Data mining. Përveç këtyre software dhe algoritmave u sugjeruan dhe databazat e përshtatshme për tipe të ndryshme të dhënash, si ato të strukturuar ashtu edhe të pastrukturuar.

Më tej, u eksplorua mjedisi etik i big data dhe analizimit të të dhënave. Për të kuptuar praktikisht çfarë problemesh mund të lindin u analizua case study i kompanisë Facebook. Atje kuptuam se sa e thjeshtë është të bëhesh pre e problemeve ligjore. As kompani të fuqishme nuk kanë imunitet ndaj çështjeve etike dhe çfarë përcakton shoqëria si të drejtë.

Shpeshherë është e vështirë të përcaktohet se çfarë është etike apo jo për shkak të subjektivitetit të çështjes në fjalë. Për këtë arsye, u trajtuan disa pika që ndihmojnë kompanitë të mos tejkalojnë kufijtë e asaj çfarë është e pranueshme nga ana sociale dhe të mos tejkalojnë kufijtë që përcaktojnë ligjet.

Si përfundim, u krye një analizë e problemit të rritjes eksponenciale të big data, u trajtuan mundësitë e reja dhe sfidat për bizneset. U eksplorua mjedisi i big data në Shqipëri duke u kryer një studim kualitativ. U bënë sugjerimet përkatëse si nga ana teknologjike dhe nga ana menaxheriale dhe u rikujtua se përpos përfitimeve, çdo kompanie i lind përgjegjësia që ti përdori big data me transparencë dhe mendësi. Vetëm në këtë mënyrë, kompanitë mund të gjejnë përgjigje për pyetjet e tyre, dhe në të njëjtën kohë të operojnë me integritet dhe përgjegjshmëri.

Referenca

- Abernethy, M. A., Chua, W., Lockett, P. F., & Selto, F. H. (1999). Research in managerial accounting: Learning from others' experiences. *Accounting & Finance*.
- Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A. S., & Buyya, R. (2015). Big Data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*.
- Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: how does data-driven decision-making affect firm performance? Cambridge: Working Paper, Sloan School of Management, MIT.
- Chessel, M. (2014) Ethics for big data and analytics. Marrë nga https://www.ibmbigdatahub.com/sites/default/files/whitepapers_reports_file/TCG%20Study%20Report%20-%20Ethics%20for%20BD%26A.pdf
- Choi, N. (2017). Top 30 big data tools for data analysis. Marrë nga <http://bigdata-madesimple.com/top-30-big-data-tools-data-analysis/>
- Chow-White, P. A., & Green, S. E. J. (2013). Data Mining Difference in the Age of Big Data: Communication and the Social Shaping of Genome Technologies from 1998 to 2007. *International Journal of Communication*.
- Davenport, T. H. (2014). Big data at work: dispelling the myths, uncovering the opportunities. Boston, Massachusetts: Harvard Business Review Press.
- Demchenko, Y. (2013) The Big Data Architecture Framework (BDAF) Outcome of the Brainstorming Session at the University of Amsterdam.
- Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*.
- Dull, R. B., Gelinas, U. J., Jr., & Wheeler, P. R. (2012). *Accounting Information Systems: Foundations in Enterprise Risk Management*. Mason, OH: South-Western Cengage Learning.
- Emani, C. K. Nadine, C., Christophe, N. (2015) Univ. Bourgogne Franche-Comté, Understandable Big Data: A survey F-21000 Dijon, France
- Fanning, K., & Grant, R. (2013). Big Data: Implications for Financial Managers. *Journal of Corporate Accounting & Finance*.

Flair, T. (2016) Big Picture of Big Data – Top 10 Big Data Trends in 2017. Marrë nga <https://data-flair.training/blogs/big-picture-of-big-data-top-10-big-data-trends-2017/>

Flair, T. (2016) How to Install Hadoop on multi-node cluster. Marrë nga <http://data-flair.training/blogs/install-hadoop-1-x-on-multi-node-cluster/>

Frizzo-Barker, J., Chow-White, P. A., Mozafari, M., & Ha, D. (2016). An empirical study of the rise of big data in business scholarship. *International Journal of Information Management*.

Frizzo-Barker, J., Chow-White, P. A., Mozafari, M., & Ha, D. (2016). An empirical study of the rise of big data in business scholarship. *International Journal of Information Management*.

Ghoting A. and Pednault, E. (2009) An Infrastructure for the Rapid Implementation of Parallel Reusable Analytics, *Proc. Large-Scale Machine Learning: Parallelism and Massive Data Sets Workshop*.

Goes, P. B. (2014). Big Data and IS Research. *MIS*, 3-8.

Granlund, M., & Malmi, T. (2002). Moderate impact of ERPS on management accounting: a lag or permanent outcome? *Management Accounting Research*, 299-321.

Half, R. (2016) Salary Guide. Marrë nga https://www.roberthalf.com/sites/default/files/Media_Root/images/at-pdfs/2017-robert-half-salary-guide-accounting-finance.pdf

Harman, M. (2018) From Start-ups to Scale-ups: Opportunities and Open Problems for Static and Dynamic Program Analysis. Marrë nga <https://research.fb.com/publications/from-start-ups-to-scale-ups-opportunities-and-open-problems-for-static-and-dynamic-program-analysis/>

Jenkins, A. (2018) We're Keeping Track of All of Facebook's Scandals So You Don't Have To. Marrë nga <http://fortune.com/2018/04/06/facebook-scandals-mark-zuckerberg/>

Lakshmi C. V. V., Nagendra, K. (2016). Survey Paper on Big Data. *International Journal of Advanced Research in Computer Science and Software Engineering*.

McAfee, A. (2002). The Impact of Enterprise Information Technology Adoption on Operational Performance: An Empirical Investigation. *Production and Operations Management*.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. Harvard Business Review.

Rizkallah, J. (2017) The Big (Unstructured) Data Problem. Marrë nga <https://www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/#1a54f309493a>

Sanjay, R. (2015) “ Big Data and Hadoop with components like Flume, Pig, Hive and Jaql” International Conference on Cloud, Big Data and Trust 2013, RGPV. ^[1]_{SEP}

Saunders, M., Lewis, P., & Thornhill, A. (2009). Research methods for business students. Harlow, England: Prentice Hall.

Shaikh, A. A., & Karjaluto, H. (2015). Making the most of information technology & systems usage: A literature review, framework and future research agenda. Computers in Human Behavior.

Shao, B. B. M., & Lin, W. T. (2016). Assessing output performance of information technology service industries: Productivity, innovation and catch-up. International Journal of Production Economics.

Sharma, S. (2016). Expanded cloud plumes hiding Big Data ecosystem. Future Generation Computer Systems.

Vishal, S., Pravin, D. (2013) “ HADOOP SKELETON & FAULT TOLERANCE IN HADOOP CLUSTERS ”, International Journal of Application or Innovation in Engineering & Management. ^[1]_{SEP}

Zhao, Y., Jie W & Cong Liu, (2014) “ Dache: A Data Aware Caching for Big-Data Applications Using the MapReduce Framework ”, TSINGHUA SCIENCE AND TECHNOLOGY. ^[1]_{SEP}