

## BIG DATA-Г БИЗНЕСИЙН БАЙГУУЛЛАГАД НЭВТРҮҮЛЭХ НЬ

Ш.Эрдэнэбилэг\*, Д.Чимгээ\*\*

**Хураангуй:** Өнөө үед их хэмжээний мэдээлэл асар хурдацтай өсч байгаа бөгөөд үүнд одоогоор ашиглаж байгаа өгөгдлийн сан, мэдээлэл боловсруулах хэрэгслүүд хүчин чадлын хувьд хангалтгүй болсон байна. Их хэмжээний мэдээлэлтэй ажиллах, түүнд шинжилгээ хийхэд шинэ технологи, арга хэрэгслүүд шаардагдахын зэрэгцээ үүнд нийцсэн шинэ технологиуд ч гарч хөгжиж байна. Энэхүү өгүүлэгт Big Data буюу их хэмжээний мэдээлэл гэж юу болохыг тодорхойлж, түүний хөгжлийн чиг хандлагыг мэдээллийн технологийн шинжлэх ухааны бусад хөгжлүүдтэй харьцуулан судалсан. Big Data-ийн бизнес дэх хэрэглээг судлахдаа гадаад болон дотоодын бизнес компаниудаар хэрхэн нэвтрүүлж байгааг судалгааны материалуудаар тодруулж гаргасан ба учирч буй саад бэрхшээл болон түүнийг ашиглан бизнесийн байгууллагууд ашиг, орлогоо нэмэгдүүлэх боломжийн талаар судалсан болно. Big Data-г нэвтрүүлэхийн тулд бизнесийн компаниуд ажиллаж байгаа салбараа судалж, байгаа нөөц боломжуудаа тооцоолж өөрсдийн стратегийг боловсруулахдаа энэ нөхцөл байдлыг бодолцож төлөвлөх хэрэгтэй болсон талаар дүгнэлтүүд гаргасан. Судалгааны ажлын үр дүнд бизнес компаниудад их хэмжээний мэдээллийг ашиглан ашиг орлогоо нэмэгдүүлэхдээ Big Data ашиглах талаар зөвлөмж гаргасан байна.

**Түлхүүр үг:** Их хэмжээний мэдээлэл, Big Data, бизнес, компани

---

\*ХААИС, ЭЗБС-ийн сургууль, Санхүүгийн тэнхимийн багш

\*\*МУИС-ийн, Бизнесийн сургууль, Бизнесийн мэдээллийн системийн хөтөлбөрийн дэд профессор,  
(E-mail) chimgee6847@gmail.com

## Оршил

Дэлхийн улс орнуудад нийгэм, улс төр, эдийн засгийн бүхий л салбарт нь мэдээллийн технологи эрчимтэй нэвтэрч энэ нь хүмүүсийн ажил, амьдрал, дадал, зан төлөвт ихээхэн өөрчлөлтүүдийг авчирсаар л байна. Мэдээллийн энэхүү эрин зуунд асар их хэмжээний мэдээлэл өдөр бүр үүсч байгаа ба цаашдаа ч энэ нь улам ихээр үүсэх нь илт байна. Мэдээллийн технологийн салбарт тэргүүлэгч IDC компаний судалгаанд 2005 оноос 2020 он гэхэд мэдээллийн хэмжээ 130 exabyte<sup>3</sup>-аас 40,000 exabyte хүрнэ. Энэ хугацаанд 2 жил тутам мэдээллийн хэмжээ 2 дахин их болно. Мэдээллийн технологид оруулж буй хөрөнгө оруулалт 40%-аар өсөх ба үүний үр дүнд 1 gigabyte-д оруулж буй хэмжээ \$2.00-оос \$0.20 болж буурна. Мэдээллийн хадгалалтын удирдлага, аюулгүй байдал, их хэмжээний мэдээлэл, үүлэн тооцоололд оруулж буй хөрөнгө оруулалт ихээхэн хурдацтай өсөх болно<sup>4</sup> гэжээ.

Ийнхүү мэдээлэл асар их хэмжээгээр үүсэж байгааг судлаачид дараах үндсэн хүчин зүйлээр тайлбарлаж байна. Юуны өмнө 1990-ээд онд интернет үүсч компьютерууд хоорондоо холбогдсон, том дата төвүүд бий болсон явдал юм. Ухаалаг гар утас мөн нэг чухал хүчин зүйл болж байна. Ухаалаг гар утас хэрэглээнд нэвтэрснээр хүмүүс янз бүрийн аппликэйшн ашиглаж, өөр өөрийн мэдээллийг үүсгэж байна. Байгууллагууд ч бүхий л үйл ажиллагааны талаар мэдээллээ дижитал хэлбэрт оруулдаг болов. Дараагийн нэг хүчин зүйл нь өнөөгийн интернетийн хөгжил буюу Эд юмсын интернет (Internet of Things-IoT) юм. Урьд нь интернетэд зөвхөн хүн л холбогддог байсан бол одоо бүх зүйл холбогддог болж байна. Алсын зайд байршиж байгаа төхөөрөмжүүд үүнд, орон сууцны хаалга, цонх, хөргөгч, галын систем зэрэг, бизнесийн хэрэглээнд машины арилжаа, зогсоолын сул талбайг тоолох, байгууллагын цагийн систем, гэрэл гээд л холбогдож байна. Галын системд асуудал гарлаа гэж мэдрэхэд тэр даруй таны гар утсанд мэдээллийг дамжуулна. Объектыг утасгүй интернетийн сүлжээгээр алсаас удирдана, эд зүйл эргээд мөн мэдээллээ илгээнэ. Эсвэл гудамжинд явж байгаа бүх машинууд машинаас машин холболт (vehicle to vehicle communication)-оор 0.1сек тутамд өөрийнхөө явж байгаа хурд, хурдатгал, байршлаа эргэн тойрны 200 метрт дамжуулдаг болно. Замын хажуугаар системүүд тавьснаар нэгдсэн төв дээр бүх машины мэдээлэл очно гэх мэтчилэн олон зүйл мэдээллээр хэмжигдэж, удирдагдах юм<sup>5</sup>.

<sup>3</sup>Мэдээллийг хэмжих нэгж бөгөөд  $10^{18}$ -тай тэнцүү.

<sup>4</sup>IDC, The digital universe in 2020: Big Data, Bigger Digital Shadow s, and Biggest Growth in the Far East, 2012

<sup>5</sup>[http://www.itzone.mn/en/news/73-Ай Ти Зон компанийн дэд захирал Б.Цэвсүрэн "Их хэмжээний өгөгдөл ба орчин үеийн бизнес аналитикийн хандлага"](http://www.itzone.mn/en/news/73-Ай Ти Зон компанийн дэд захирал Б.Цэвсүрэн ) илтгэл

### **Судалгааны зорилго**

Энэхүү судалгааны зорилго нь Big Data-г бизнесийн байгууллагад нэвтрүүлэхэд арга зүйн зөвлөмж гаргахад нь оршино. Зорилгод хүрэхийн тулд:

- Big Data- гэж юу болохыг тодорхойлж түүний хөгжил, чиг хандлагыг тодруулах
- Дэлхийн улс орнууд болон Монгол улсад Big Data-г нэвтрүүлж буй өнөөгийн байдлыг судлан тодорхойлох
- Big Data-г нэвтрүүлэхэд бизнесийн байгууллагад учирч буй саад бэрхшээлийг тодруулах
- Бизнест Big Data-г үр дүнтэй нэвтрүүлэхэд арга зүйн зөвлөмж гаргах зорилтуудыг тавьсан болно.

### **Судалгааны арга зүй**

Энэхүү судалгааг Big Data-гийн онол арга зүйн үндсийг судалж үндсэн ойлголт, хөгжлийн түүх, чиг хандлага, үзэл баримтлалыг тодруулах үүднээс холбогдох ном, сурах бичиг болон шинжлэх ухааны судлаачдын бүтээлтэй танилцсан. Түүнчлэн Big Data-гийн нэвтрүүлэлтийн талаар гадаадын орнуудын болон өөрийн орны туршлагыг судалж онцлогуудыг тодруулах үүднээс судалгааны материал, ном, хийгдсэн судалгааг харьцуулан судалсан.

Түүнчлэн ярилцлага авах, хэлэлцүүлэг форумд биечлэн оролцох, ажиглалт хийх зэрэг шаардлагатай мэдээллийг цуглуулах шинжлэх ухааны аргуудыг энэхүү судалгаандаа өргөн ашиглалаа.

Судалгааны ажилд Excel, Word, програмыг ашигласан бөгөөд Интернет орчинд мэдээлэл цуглуулж дүн шинжилгээ хийсэн болно.

### **Судалгааны ач холбогдол**

Энэхүү судалгааны ач холбогдол нь дараах зүйлүүдээр тодорхойлогдож байна.

Үүнд:

- Big Data буюу их хэмжээний мэдээлэл гэсэн шинжлэх ухааны хөгжлийн шинэ чиг хандлагыг судалж тодруулсан
- Big Data-гийн онол, арга зүйн үндсийг судалж үндсэн ойлголтуудыг

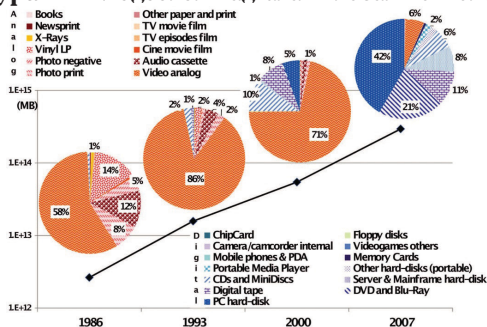
- тодорхойлж гаргасан
- Гадаад орнууд болон дотоодын бизнес байгууллагуудын Big Data нэвтрүүлсэн туршлагаудыг судалж учирч буй саад бэрхшээлийг тодруулсан
- Компаний стратеги, бизнес загварыг өөрчлөхөд Big Data ашиглах боломжтой болсныг тодорхойлж гаргасан
- Big Data-г бизнесийн байгууллагад нэвтрүүлэх арга зүйн зөвлөмж боловсруулсан явдал юм

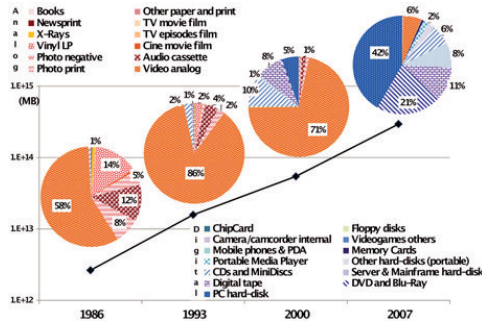
### Big data гэж юу вэ?

Судалгааны материалуудаас харахад сүүлийн хэдэн жилийн турш Big Data гэдэг нь мэдээллийн технологи, маркетинг, хэвлэл мэдээллийнхний зүйр үг хэллэг болжээ. Энэ нь ч аргагүй юм. Тоон технологи нь орчин үеийн хүний амьдралд нэвтрэн орсон. Амьдралын янз бүрийн талаарх мэдээлэл улам нэмэгдэхийн хамт мэдээллийг хадгалах боломжууд ч нэмэгдсэн байна. Мэдээлэл хадгалах глобал технологиудыг харуулсан зургаас орчин үед компьютер, сервер, DVD, Blu-Ray, видео тоглоомоор мэдээллийг түлхүү хадгалах болжээ.

Мэдээллийг хэмжээний хурдацтай өсөлт бодит зүйл болохыг ихэнх шинжээчид хүлээн зөвшөөрч байна. Нийгмийн сүлжээ, хэмжил хийх хөдөлгөөнт тоног төхөөрөмж, бизнес-мэдээлэл гээд зөвхөн эдгээрийг авч үзэхэд л асар их хэмжээний мэдээллийг бий болгох чадвартай эх үүсвэрүүд юм. IDC Digital Universe судалгааны байгууллагын 2012 онд хэвлэсэн судалгаанаас харахад ойрын 8 жил дэлхийд мэдээллийн хэмжээ 40 зеттабайт (zettebytes)-д хүрэх бөгөөд энэ нь дэлхийд амьдарч буй нэг хүнд 5200 гбайт (гигабайт) оногдож байгаатай тэнцүү байх юм.

Зураг 1. Мэдээлэл хадгалах глобал технологи

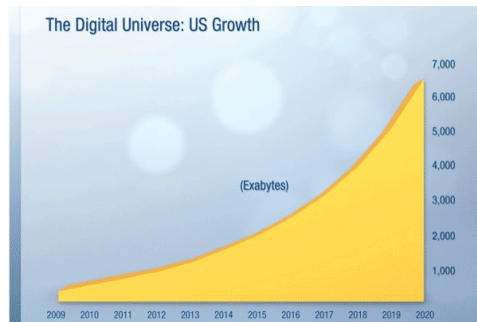




Эх сурвалж: Hilbert and Lopez, 'The world's technological capacity to store, communicate, and compute information, Science, 2011 Global

IDC байгууллагын судалгаагаар АНУ-д 2012-оос 2020 оны хооронд мэдээлэл 898 exabyte-с 6.6 zetabyte хүртэл өсөх бөгөөд энэ нь жилд 25%-с илүү өсөх буюу 3 жил тутам 2 дахин өсөх болно гэж тооцоолж гаргажээ (зураг 2).

**Зураг 2.** АНУ-ийн цуглуулж буй тоон мэдээллийн өсөлт



Эх сурвалж: John Gantz and David Reinsel, IDC Digital Universe in 2020, Big Data..in United States, 2013

Мэдээллийн ихэнх хувийг хүмүүс биш харин бие биетэйгээ болон бусад сүлжээнд харилцаж буй роботууд жишээ нь, мэдрэгч, ухаалаг төхөөрөмжүүд бий болгож байна. Судлаачдын таамаглаж байгаагаар иймэрхүү хурдацтай байвал дэлхий дээр жил бүр мэдээллийн хэмжээ хоёр дахин өснө гэж үзэж байна. Дэлхийд виртуал болон физик серверүүдийн тоо шинэ дата төвүүдийг бий болгох замаар арав дахин өснө гэжээ. Үүнтэй холбоотойгоор эдгээр мэдээллийг үр дүнтэй ашиглах, арилжиж мөнгө болгох хэрэгцээ шаардлага улам

нэмэгдэж байна. Big Data-г бизнест ашиглахын тулд нилээдгүй хэмжээний хөрөнгө оруулалтыг шаардах тул тодорхой нөхцөл байдлыг сайн ойлгох нь чухал юм. Энэ нөхцөл байдал гэдэг нь үнэндээ бол энгийн ойлголт бөгөөд зардлыг багасгах эсвэл борлуулалтыг нэмэгдүүлэх замаар бизнесийн үр ашгийг нэмэгдүүлэх явдал юм.

### Big Data-ийн түүх

Анх их хэмжээний мэдээлэлтэй уялдан “Big Data” гэсэн үг хэллэг, ойлголт гарсан байна. 2001 онд META Групп (META Group) буюу одоогийн мэдээллийн технологийн чиглэлд судалгаа, зөвлөх үйлчилгээгээр дэлхийд тэргүүлэгч Гартнер (Gartner) компаний судлаач Даг Ляней (Doug Laney) мэдээллийн энэ асар их өсөлт, боломжуудыг 3 шинж чанараар 1. **volume**- өгөгдлийн хэмжээ, 2. **velocity**-бодит цаг хугацаанд өгөгдөл үүсч, боловсруулагдаж буй хурд, 3. **variety**- өгөгдлийн олон төрөл гэж тодорхойлсон бөгөөд эдгээр нь улам өссөөр байна гэжээ. Одоо ч Гартнер “3v” загварыг “Big Data”-г тодорхойлохдоо ашиглаж байна. Хөгжлийн явцад “Big Data”-гийн дараагийн “v” буюу шинжүүдийг гаргаж ирж байна. Тухайлбал: **veracity**-өгөгдлийн үнэн зөв байдал, алдаагүй шинжилгээ нь эх өгөгдлийн үнэн зөвөөс хамаарна гэсэн ойлголт, **variability**-өгөгдөл өөрчлөгдөх шинж буюу заримдаа харуулж буй өгөгдөл нийцэхгүй байх нь өгөгдлийг үр ашигтай удирдахад сөрөг нөлөө үзүүлдэг болохыг харуулдаг.

Big Data-д дараах төрлийн өгөгдлүүдийг хамааруулан ойлгож байна. Үүнд: 1) Байгууллагын ердийн мэдээлэл: Хэрэглэгчтэй харилцах удирдлагын (CRM) системүүдийн мэдээлэл, ажил гүйлгээний Байгууллагын нөөц төлөвлөлт (ERP) системийн өгөгдлүүд, вебд хадгалагдсан ажил гүйлгээ, ерөнхий дэвтрийн мэдээлэл гэх мэт; 2) Автомат машин, мэдрэгч төхөөрөмжөөс үүсч буй өгөгдлүүд, жишээ нь, дуудлагын дэлгэрэнгүй бичлэгүүд-Call Detail Records (“CDR”) гэх мэт; 3) Нийгмийн өгөгдлүүд- нийгмийн сүлжээ Twitter, Facebook зэргээр үүсч буй мэдээлэл орж байна.

“Big Data” гэсэн нэрийг анх гаргасан хүн бол Nature сэтгүүлийн редактор Клиффорд Линч бөгөөд 2008 оны 9 сарын 3-д гаргасан “Big Data: How do your data grow?” сэдэвтэй сэтгүүлийн тусгай дугаарт мэдээлэл хэмжээний хувьд асар их өсч байгаа, боловсруулж буй мэдээллийн олон төрөл, технологийн боломжуудыг өгүүлсэн ба материалууд нь бүхэлдээ их хэмжээний мэдээлэл нь ирээдүйн шинжлэх ухаан, техникт хэрхэн нөлөөлөх тухай хамарсан байжээ. Тэрээр их хэмжээний өгөгдөлтэй холбоотой ойлголтыг ажил хэргийн англи ярианы орчинд хэрэглэгддэг Big гэсэн үгтэй адилтган Bid data гэж нэрлэсэн

байв<sup>6</sup>.

Хэдийгээр энэ үг хэллэг судалгааны хүрээнд байсан боловч 2009 оноос ажил хэргийн орчинд өргөн дэлгэрчээ. 2010 онд их хэмжээний өгөгдлүүдтэй ажиллахтай холбоотой анхны бүтээгдэхүүн, шийдвэрүүд гарсан байна. 2011 онд мэдээллийн технологийн томоохон нийлүүлэгч компаниуд болох IBM, Oracle, Microsoft, Hewlett-Packard өөрсдийн стратегийг тодорхойлохдоо их хэмжээний өгөгдлийн ойлголтыг ашиглажээ. 2011 онд Гартнер тэмдэглэхдээ их хэмжээний мэдээлэл нь мэдээллийн технологийн дэд бүтцэд виртуалчлалын дараа 2-рт орох чиг хандлага болно гэсэн байна. Их хэмжээний мэдээллийн технологийг нэвтрүүлснээр үйлдвэрлэл, эрүүл мэнд, худалдаа, улсын удирдлага, хувь хүний нөөцийн шилжилт хөдөлгөөний бүртгэлийг хийж байгаа салбаруудад хамгийн их нөлөө үзүүлнэ гэж таамаглаж байна. 2013 оноос эхлэн их, дээд сургуульд өгөгдлийн болон тооцон бодох шинжлэх ухаан, инженерийн чиглэлээр хөтөлбөрүүд гарч тэдгээрт их хэмжээний өгөгдөл хичээл болж заагдах болжээ<sup>7</sup>.

Хятадын электрон худалдааны аварга Алибагийн AliCloud бизнесийн хэсэг нь хэрэглэгчийн датаг өндөр нууцлалтай хадгалах квантын кодлол (quantum cryptography) гэж нэрлэж байгаа технологийг зарлажээ. Тэд энэ технологийг нэвтрүүлсэн дата төвийг өнгөрсөн долоо хоногт АНУ-д нээсэн байна. Харин Европт ирэх жил ийм дата төвөө нээх гэнэ. Эдгээр төслүүддээ зориулж Алибаба нэг тэрбум ам.долларын хөрөнгө оруулалт хийхээр болжээ. Тэд ирэх арван жилд дата төвийн хэрэгцээ эрчимтэй өснө гэсэн төсөөлөлтэй байгаагаас гадна Алибаба өөрийн платформоор дамждаг гүйлгээний хэмжээг энэ жил \$500 тэрбумд хүргэх зорилтыг хангахад энэ нь хөшүүрэг болно гэж үзэж байгаа юм. Энэ тухай Ханжоу хотод болсон арга хэмжээн дээр мэдээлэхдээ Жэк Ма “Ирээдүйд дата нь хамгийн том үйлдвэрлэлийн түүхий эд болно. Өнөөдрийн ус, цахилгаан, газрын тос шиг өргөн хэрэглээний баялаг болох юм” гэж ярьжээ. Одоогоор клауд бизнесийн хэмжээ нь Алибабад багахан хувийг эзэлж байгаа. Өнгөрсөн хоёрдугаар улирлын санхүүгийн мэдээллээр нийт орлогын 2.6 хувийг л үүлэн тооцооллын орлого эзэлж байжээ.

Их хэмжээний өгөгдөлтэй data warehouse, data mining, Big Data гэсэн гурван ойлголт холбоотой. Data warehouse нь мэдээллийн агуулах гэсэн үг. Энд үнэн, худал болон бодит болон (real time), өнгөрсөн цагийн гэсэн янз бүрийн төлөв байдалд байгаа өгөгдлүүд хадгалагдаж байдаг. Өгөгдлийн агуулахад

<sup>6</sup>[https://ru.wikipedia.org/wiki/Большие\\_данные](https://ru.wikipedia.org/wiki/Большие_данные)

<sup>7</sup><http://www.seas.harvard.edu/computational-science-and-engineering/master-science-cse>.

байгууллагын удирдлагын шийдвэр гаргахад шаардагдах бүхий л төрлийн өгөгдөл, мэдээлэл байна. Data mining нь өнөө үед өгөгдөл олборлолт, өгөгдөл тандалт зэрэг үгээр илэрхийлэгдэж ярьж, бичигддэг ба мэдлэгийг олох, шинэ мэдээллийг хайх, өгөгдлийн судалгаа хийхтэй холбоотой юм. Энэ хоёрын нэг нийтлэг тал нь өнгөрсөн, түүхт мэдээлэл дээр судалгаа шинжилгээ хийдэг тул чиг хандлагыг тодорхойлж шийдвэр гаргахдаа data mining-г ашигладаг. Харин Big Data зөв мэдээллийг өгөхийн тулд өөрөө бодит цагийн, цэгцэрсэн, давхцалгүй өгөгдөл байх ёстой юм. Бодит цагийн өгөгдөл гэдэг нь серверийн дэд бүтцэд маш хоромхон богино буюу миллисекундын зөрүүтэй ирж байгаа өгөгдлийг хэлнэ. Big Data-г миллисекунд тутамд эзэмшигч нь буюу хэрэглэгчид өөрсдөө шинэчилж байдаг<sup>8</sup>.

### **Big Data юунд хэрэгтэй вэ?**

Судалгаанаас Big Data-ийн гурван үндсэн зорилго буюу үүрэг байна гэж үзэв. Үүнд:

- Уламжлалт мэдээллийн сангийн технологиуд үр өгөөжтэй хадгалах боломжгүй болж байгаа өнөөгийн нөхцөлд терабайт, петабайтаар хэмжигдэх асар их хэмжээний мэдээллийг хадгалах
- Тодорхой зохион байгуулалтгүй текст, зураг, видео болон бусад төрлийн мэдээллийг зохион байгуулах
- Зохион байгуулалтгүй өгөгдөл, мэдээлэлтэй ажиллан их хэмжээний мэдээлэлд шинжилгээ хийх, шинжилгээний тайлан бүрдүүлэх, түүнчлэн прогнозчлох загваруудыг бий болгох явдал юм.

### **Big data-г бизнест нэвтрүүлэх нь**

Big Data-ийн төслүүдийн зах зээл нь бизнес аналитик (BA)-ийн зах зээлтэй солбилцож, хэмжээний хувьд улам тэлж байгаа өнөөгийн үед энэ нь сүлжээний технологи, сервер, программ хангамж, техникийн үйлчилгээ зэргийг бүгдийг хамарч байгааг онцлон тэмдэглэхийг хүсэж байна.

Мөн компаниуд үйл ажиллагаагаа автоматжуулан орлого олох баталгаа өгөх системийг бий болгохын тулд Big Data технологи ашиглах нь үр дүнтэй байж болох юм. Орлого олох баталгаа өгөх орчин үеийн системүүдийн онцлог нь нийцэхгүй байгаа зүйлсийг нэн даруй илрүүлэх, гарах алдагдал, эсвэл санхүүгийн үр дүнг бууруулахад хүргэх мэдээллийн гажуудлыг эрт илрүүлэх боломжийг өгч чадах мэдээллийн гүнзгий дүн шинжилгээг өөртөө багтаасан

<sup>8</sup>“Big Data & Analytics Conference 2015”-д тавьсан “Төр засгийн үйл ажиллагаанд их хэмжээний өгөгдлийн шинжилгээг хэрэгжүүлэх нь”, Үндэсний Дата Төв, УТҮГ-ын Захирал З.Энхболдийн илтгэл



байдаг. Ийм нөхцөл байдалд манай бизнес компаниуд дотоодын зах зээл дээр их хэмжээний мэдээллийн технологийн эрэлт хэрэгцээ байгааг хэлж, нотолж байгаа бөгөөд түүний хурдацтай өсөлтийн гол хүчин зүйл нь мэдээллийн өсөлт, удирдлагын шийдвэр гаргалтыг хурдасгах түүний чанарыг нь дээшлүүлэх явдал гэж үзэж байна.

### **Big Data-тай ажиллахад юу саад болж байна вэ?**

Өнөөдөр Big Data ашиглан шийдэж болох салбарын хэмжээнд асуудлууд бодитой байгаа ч хуримтлагдсан тоон өгөгдөл, мэдээллийн зөвхөн цөөн хэсэгт л дүн шинжилгээ хийж байна. Мэдээллийн технологийн зах зээлийн хөгжил нь их хэмжээний мэдээллийн хуримтлал, түүний боловсруулттай холбоотой хүлээлтийг үнэлэх боломжийг өгч байна. Big Data-ийн хэрэгжилтэд саад болдог нэг гол хүчин зүйлс нь төслүүд бөгөөд эдгээр нь өндөр өртөгтэй байдаг, түүнчлэн боловсруулах мэдээллийг сонгох асуудал байгаа юм. Энэ нь өөрөөр хэлбэл ямар өгөгдлийг татаж авах, хадгалах, болон дүн шинжилгээ хийх, ямар өгөгдөлд анхаарал хандуулах хэрэггүй вэ гэдгийг тодорхойлох нь бэрхшээлтэй байдагтай холбоотой.

Манайд бизнесийн төлөөлөгчидтэй уулзаж, ярилцаж байхад тэд Big Data төслүүдийг хэрэгжүүлэхийн тулд энэ талын мэдлэгтэй маркетер, менежер болон шинжээч зэрэг мэргэжилтнүүд хангалтгүй байдаг гэдэг. Үнэндээ гүнзгий нарийн, эсвэл урьдчилан таамаглах шинжилгээ хийж буй эдгээр мэргэжилтнүүдийн ажлын чанараас Bid Data төслүүдэд оруулж буй хөрөнгө оруулалтын үр өгөөжийг олох хурдац шууд хамаардаг. Байгууллагад аль хэдийн бий болсон мэдээллийн тэр их боломжийг мэргэжилтнүүд хуучирсан бизнес процесс, дотоод дүрэм, зохицуулалтаас шалтгаалан ихэвчлэн үр дүнтэй ашиглаж чадахгүй байна. Бизнес компаниуд Big Data төслүүдийг нэвтрүүлэхэд нарийн хэцүү гэж үзэхээс гадна хүрсэн үр дүнг үнэлэх, цуглуулсан мэдээллийн үнэ цэнийг мэдэхэд төвөгтэй гэж харж байна. Маркетер, менежер, шинжээчдийг мэдээлэлтэй ажиллаж тодорхой бизнесийн асуудлыг шийдвэрлэхдээ технологи болон тайланг бүрдүүлэхэд анхаарлаа хандуулахыг шаарддаг.

Их хэмжээний, хурдан урсгалтай мэдээллийг цуглуулахад бодит цаг хугацааны горимд ETL ажиллагааг явуулдаг (энэ нь E-Extract T-Transform L-Load үгүүд бөгөөд энэ нь Гаргах, Хувиргах, Ачаалах гэсэн санааг агуулсан, мэдээллийг хадгалах удирдлагын үндсэн үйл ажиллагаанууд юм). Энэхүү ажиллагаа нь гадаад эх үүсвэрээс мэдээллийг татан авч түүнийг бизнесийн загварын шаардлагад нийцүүлэн хувиргаж, цэвэрлэж улмаар өгөгдлийн агуулах

руу ачаалах замаар явагддаг. ETL-ийг өгөгдлийг нэг хэрэглээнээс нөгөө нэг хэрэглээ рүү шилжүүлэх процесс гэж ойлгохоос гадна шинжилгээнд өгөгдөл, мэдээллийг бэлтгэх хэрэгсэл гэж ойлгох нь зүйтэй юм.

Гадаад эх үүсвэрээс ирж буй мэдээллийн аюулгүй байдлын асуудлыг цуглуулж буй мэдээллийн хэмжээтэй нийцүүлэн шийдэх шаардлага гарна. Big Data-гийн шинжилгээний аргууд зөвхөн мэдээллийн тоо хэмжээний өсөлтийн араас хөгжиж байдаг тул мэдээллийг бэлтгэх, нэгтгэх шинэ аргуудыг ашиглахдаа шинжилгээний платформын шинжүүд дээр анхаарах хэрэгтэй.

Хэдийгээр хүндэлүүд байгаа ч Gartner-ын судалгаагаар 2012 онд дэлхийн томоохон компаниудын 58% нь Big Data төслүүдэд хөрөнгө оруулалт хийж байсан бол 2013 онд энэ нь 68% болж өсчээ. Big Data төслүүдийг нэвтрүүлж буй компаниудад харилцаа холбоо, банк, үйлчилгээний салбарын тэргүүлэх компаниуд орсон байна. Түүнчлэн радио давтамжаар илрүүлэх хэрэгслээс авч буй мэдээллийг ашигладаг жижиглэнгийн худалдааны байгууллагад Big Data төслүүдийг амжилттай хэрэгжүүлж байна.

### **Big Data-гийн төслийг хэрхэн хэрэгжүүлэх вэ?**

Бизнесийн салбарт Big Data-г үр дүнтэй хэрэгжүүлэх програмын болоод техник хэрэгслийн олон боломжтой хослол байна. Бизнесийнхний дунд Big Data-тай ажиллах нь ихээхэн зардлыг мэдээлэл цуглуулахад зарцуулах шаардлага гарна, ялангуяа тэдгээрийг юунд ашиглан хэрхэн боловсруулалт хийхээ өмнө нь мэдэхгүй байхад их эрсдэлтэй гэсэн ойлголт байдаг. Иймээс бизнесийнхэн юуны өмнө их хэмжээний мэдээлэлтэй байхад анхаарах биш харин ямар зорилгоор юунд ашиглах гэж байгаагаар сайн ойлгох нь чухал юм. Энэ тохиолдолд мэдээлэл цуглуулах нь зөвхөн тэр л чухал асуудлыг шийдэхэд зориулсан мэдээллийг авна гэдэгт итгэлтэй байж болно.

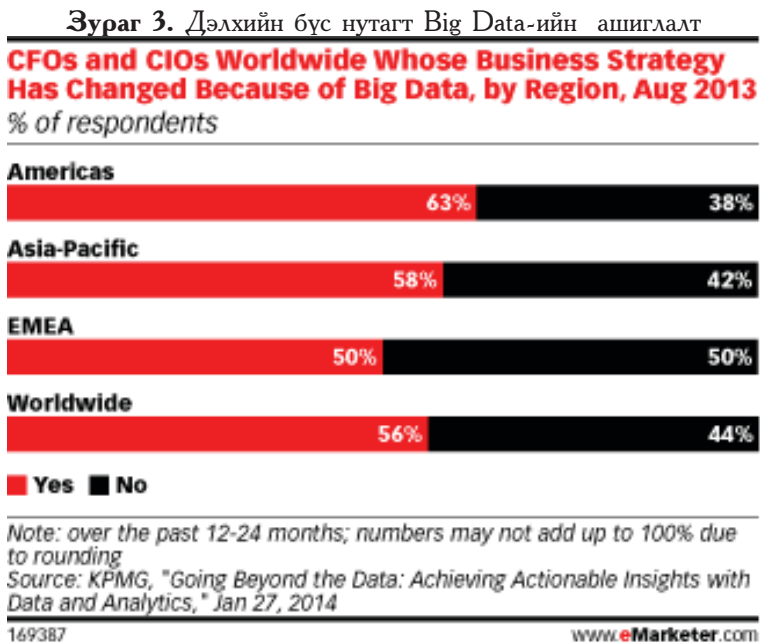
Компани Big Data-тай үр дүнтэй ажиллахын тулд дараах зүйлс анхаарах хэрэгтэй. Үүнд:

- Эхлээд Big Data ашиглан шийдэх тодорхой нэг бизнес асуудлыг тодорхойлох. Яг энэ асуудалтай холбоотой мэдээлэл цуглуулахад анхаарах.
- Байгууллагын шинжилгээ хийх туршлага, дадал нь мэдээлэлтэй ажиллахад их чухал үүрэгтэй. Үүнээс ажлын чанар ч хамаарна. Мэдээллийг ашиглах соёл ч энд чухал.
- Мэдээллийн бизнес утга санааг сайн ойлгодог, шинжилгээний талаар мэргэжсэн мэргэжилтнүүдтэй байх, тэдгээрийг татан оролцуулах нь чухал.

Big Data-г амжилттай хэрэгжүүлж буй туршлагаудыг авч үзлээ. Орос улсад одоогоор холбооны улсын хэмжээгээр мэдээллийн агуулахыг шинжилгээнд зориулан байгуулж эхэлжээ. Үүний үндсэн дээр татварын мэдээллийн статистикийн болон шинжилгээний боловсруулалт хийх нэгдсэн мэдээллийн төв байгуулж байна. Үүнд орон нутгийн түвшний 1200 эх үүсвэрээс шинжилгээний мэдээллийг төвлөрүүлж байна. Бодит цаг хугацаанд их мэдээлэл дээр шинжилгээ хийж буй сонирхолтой өөр нэг жишээ бол 2010 онд байгуулагдсан стартап Synqera компани бөгөөд 2012 онд Simplate платформыг хөгжүүлжээ. Энэ програм нь их хэмжээний мэдээллийг ашиглан худалдан авагчдын мэдээлэл, тэдний худалдан авалт, нас, хүйс, сэтгэл санааны хүртэлх мэдээлэлд шинжилгээ хийдэг ажээ. Гоо сайхны дэлгүүрийн сүлжээний кассууд дээр худалдан авагчийн сэтгэл санааг мэдрэх таних төхөөрөмжийг байрлуулсан байна. Програм нь уг мэдээллийг хүлээн авч хүний сэтгэл санааны байдлыг тодорхойлж шинжлэн дэлгүүрийн хөнгөлөлтийн мэдээлэл, хоног цагийг тодорхойлж уг хүнд тусгай саналуудыг тавьсан мэдэгдэл өгдөг байна. Энэ систем нь худалдан авагчийн тууштай байдал, борлуулалтанд сайнаар нөлөөлж байгааг харуулсан үр дүн өгчээ.

Big Data-г амжилттай хэрэгжүүлж буй өөр нэг компани бол Америкийн глобал кофе үйлчилгээний сүлжээ бүхий Dunkin`Donuts юм. Компани нь бүтээгдэхүүнийхээ борлуулалтанд зориулж бодит цаг хугацаанд мэдээллийг ашиглаж байна. Дэлгүүрүүдийн тоон дэлгэцүүд цаг хугацаа, байгаа бүтээгдэхүүнээсээ хамаараад минут тутамд саналуудыг харуулан өөрчлөгддөг. Кассын чекээр дамжуулан ямар санал хамгийн их худалдан авагчдыг татаж байгаа талаар мэдээллийг өгч байна. Энэ аргыг хэрэглэснээр компани орлогоо нэмэгдүүлж, агуулахын барааны эргэлтээ хурдасгажээ.

Практикаас үзэхэд Big Data нь бикнесийн асуудлыг шийдвэрлэхэд үр дүнтэй байж болохыг харуулж байна. Энд нөлөөлөх хамгийн чухал хүчин зүйл бол стратегия зөв сонгох, инновацийн технологийг ашиглах явдал юм.



Эх үүсвэр: KPMG, emarketer.com-д хэвлэгдсэн

Олон улсын KPMG байгууллагын 2013 онд явуулсан судалгаанаас харахад 2013 оны байдлаар Big Data ашиглахтай холбогдуулан бизнес стратегиа өөрчилсөн компаниудын эзлэх хувь Америкт 63, Азийн бүс нутагт 58, Европ, Ойрхи дорнод болон Африкт 50%, дэлхийн хэмжээнд 56% байна.

Их хэмжээний мэдээлэл олон нийтийн мэдээллийн хэрэгсэл, харилцаа холбоо, тоглоом зугааны салбарт хүчтэй мэдрэгдэж байна. Ийм их хэмжээний мэдээлэл нь улам их, нарийн болж байгаа нь ажиллаж байгаа мэдээллийн сангуудыг удирдах систем, мэдээлэл боловсруулах хэрэгслүүдийн хүчин чадал хангалтгүй болсон байна. Энэ нь ялангуяа “нээлтэй” гэж нэрлэгдэж буй чиг хандлагад: нээлтэй их код, нээлтэй инноваци, нээлтэй мэдээлэлд тодорхой харагдаж байна. Big Data-г үнэ цэнэтэй мэдээлэл болгож хувиргахын тулд тусгай технологи, аналитикийн аргууд шаардагдана. Эдгээр шаардлагыг хангахын тулд бий болсон технологи бол уян хатан, тусгай зориулалтын, хөгжүүлэгч төвтэй Not Only SQL (NoSQL) болон найдвартай, нууцлалыг хадгалсан, удирдагддаг SQL-ийн шийдэл юм. NoSQL-ын судалгаа олон жил хийгдсэн бөгөөд ийм төрлийн шинэ өгөгдлийн санг удирдах системүүд ч одоо гарсаар байна. NoSQL өгөгдлийн санг ихэвчлэн нийгмийн сүлжээний

мэдээллийг цуглуулах, хадгалахад ашиглаж байна. Үүнийг анх Google, Facebook зэрэг их хэмжээний өгөгдөлтэй ажилладаг, бизнес нь ихэвчлэн интернетийн орчинд явагддаг компаниуд өөрсдийн дотоод хэрэгцээнд зориулан ашиглаж эхэлсэн түүхтэй<sup>9</sup>. Одоогийн байдлаар Facebook-ийн мессежийн дэд бүтэц 350 сая-аас илүү хэрэглэгчдийн 15 тэрбум хүнээс-хүнд мессежийг сард хүргэдэг бол чатлах үйлчилгээ нь 300 сая хэрэглэгчийн 120 тэрбум мессежийг сард илгээж байна. Одоо Facebook нь чат, SMS, е-шуудан, мессежийг real-time-д шилжүүлж нэгтгэсэн шинэ Мессежийн хувилбарыг нэвтрүүлэхдээ их хэмжээний өгөгдлүүдтэй ажиллах SQL шийдлийн технологиудыг ашиглаж байна. Түүнчлэн нээлттэй эхийн зарчмаар хөгжиж буй системүүд ч хөгжиж байгаа бөгөөд одоогийн байдлаар 120 орчим ийм систем аль хэдийн бий болжээ. Эдгээрээс онцолбол, их хэмжээний өгөгдлийг хадгалах Apache HBase™ болох Hadoop систем, Apache Cassandra өгөгдлийн санг нэрлэж болно. Өгөгдлийн сангийн системээр тэргүүлэгч Oracle компани шаардлагыг ойлгон хүлээн авч тохирох бүтээгдэхүүнийг хөгжүүлж байна<sup>10</sup>. Үүний нэг жишээ бол Oracle Big Data Appliance систем юм. Энэ систем их хэмжээний мэдээллийг цуглуулан зохион байгуулж Oracle-ийн мэдээллийн санд ачаалахаас гадна байгууллагын үр дүн гүйцэтгэл болон бусад бүх төрлийн мэдээллийг ашиглан real-time аналитик хийх зорилгоор загварчлагдсан бөгөөд ашиглахад хялбар, техник ба програм хангамжийг хамтад нь нэг систем болгон инженерчилсэн систем юм.

### **Монголд Big Data төслийн хэрэгжүүлэлт**

Орчин үеийн бизнесийн байгууллагуудын анхаарал хандуулж буй нэг гол талбар нь цахим бизнес, дүн шинжилгээ болж байна. Сүүлийн жилүүдэд манай дунд болон томоохон бизнесийн байгууллагууд санхүү, хүний нөөц, харилцагчидтай харилцах, хангамж нийлүүлэлтийн зэрэг програм хангамж их авах болсон. Үүний гол зорилго нь үйл ажиллагаагаа хөнгөвчлөх, автоматжуулах байгаа бөгөөд эдгээр системүүдийг 3-5 жил ашигласан компаний хувьд өөрсдийн нилээд хэмжээний цахим мэдээлэл хуримтлуулж байна. Энэ мэдээлэлдээ дараагийн алхамд компани шинжилгээ хийх нь үр дүнгээ сайжруулахад нэн чухал юм. Зөв шийдвэр гаргахын тулд зөв шинжилгээг зөв мэдээлэл дээр тулгуурлан гүйцэтгэх нь үр дүнтэй болно.

2010 оны сүүлээр Монголд хамгийн анх Big Data хэрэгжүүлэх талаар төрийн байгууллага “Үндэсний Дата Төв” албан ёсоор ярьж, ажиллаж эхэлсэн байна. Одоогийн байдлаар төрөөс бусад хэвшлийн байгууллагад мэдээлэл системтэй

<sup>9</sup><http://research.google.com/archive/bigtable.html>, <http://www.facebook.com/notes/facebook-engineering/the-underlying-technology-of-messages/454991608919>

<sup>10</sup><http://www.oracle.com/technetwork/database/nosqldb/learnmore/wp-big-data-with-oracle-521209.pdf>

бөгөөд их хэмжээгээр хараахан цуглагдаагүй байна. Иймд төрийн одоо байгаа мэдээллийг цэгцлэн Big Data-г үүсгэх боломжтой болсон үүний нэг үндэс суурь нь төрийн үйлчилгээний цахим (ТҮЦ) машин болж байв. Энэ нь Data warehouse-ийн буюу хуримтлагдсан бүх мэдээллийг ашиглан иргэндээ түргэн шуурхай үйлчилгээ үзүүлэх зорилготой байсан. Өнөөдрийн байдлаар банкнаас зээл авахад нь шаардлагатай баримт-лавлагааны 500-аад мянган үйлчилгээг иргэдэд ТҮЦ-ээс үзүүлж, төсвөөс 4<sup>11</sup> тэрбум төгрөг хэмнэсэн байна<sup>6</sup>. Энэ бол зөвхөн нэг үйлчилгээнээс хэмнэсэн мөнгө юм. Хэдийгээр энэхүү технологийн шийдэл ТҮЦ нь хүний мөнгө, цаг хугацаа, орон зайг хэмнэсэн давуу талтай байгаа ч зарим тохиолдолд ТҮЦ-ээс өгч буй мэдээлэл алдаатай байх явдал гарсан байна. Энэ нь мэдээлэл үнэн зөв байхын тулд цэгцэрсэн, давхцалгүй байх шаардлагатайг харуулж байна. ТҮЦ-ийг ажиллуулдаг цахилгаан, үнэт цаасыг хэмнэж өмнөхөөсөө улам илүү их хэмнэх боломж гарч байгаа нь Big Data-г үр дүнтэй ашиглах явдал юм. Иргэдийн мэдээлэл нэгдсэн нэг системд давхцалгүй, бодит цагийн, цэгцтэй бүртгэлтэй болсноор бид цаг хугацаагаа хэмнэж, улмаар эдийн засагт ихээхэн үр нөлөөг авчирах боломж харагдаж байна.

Манайд Ай Ти Зон компаний хөгжүүлсэн Compass Mate програм нь байгууллагууд машинуудаа хаана явж байгааг, хэр их шатахуун зарцуулж байгаа зэрэг мэдээллийг цуглуулж авах боломжийг өгч байна. Одоогийн байдлаар арван секунд тутамд мянга орчим машины талаар мэдээлэл маш их цуглагдаж байгаа ажээ. Энэ их хэмжээний мэдээллийг ашиглан байгууллага үйл ажиллагааныхаа үр дүнг сайжруулж болно. Манайд Алкэмист Техноложис компани нь их хэмжээний мэдээллийн хуримтлал дээр ажиллах бизнесийн хэрэгцээг хангахын тулд мэргэжлийн өндөр түвшинд хамгийн дэвшилтэт NoSQL технологи болох MongoDB ашиглан үйлчилгээ үзүүлж байна.

Түүнчлэн Big Data-тай холбоотой 6 сарын төсөл Нийслэл дээр хэрэгжиж байна. Энэ нь ухаалаг хотын суурь платформ, дэд бүтцийг байгуулах төсөл юм. Дэд бүтэц дээрээ суурилад цаашид хэрэглээг хөгжүүлэх зорилготой. Хотын хэмжээнд асар их мэдээлэл цугларах боломжтой. Төсөл хэрэгжсэнээр хаана ямар үйлчилгээний байгууллагууд байна, 24 цагийн эмийн сан, үсчин, дугуй засвар, өрхийн эмнэлгүүд гэх мэт хотын бүхий л мэдээллийг вэбсайт болон аппликэйшнээр өгөх боломжтой болно. Түүнчлэн энэхүү их хэмжээний мэдээлэлд шинжилгээ хийх нь жишээлбэл, бүх автобусны мэдээлэл, тэднийг

<sup>11</sup>“Big Data & Analytics Conference 2015”-д тавьсан “Төр засгийн үйл ажиллагаанд их хэмжээний өгөгдлийн шинжилгээг хэрэгжүүлэх нь”, Үндэсний Дата Төв, УТҮГ-ын Захирал З.Энхболдийн илтгэл

хамгийн зөв маршрутаар яаж явуулах вэ, хэдэн автобус явуулах вэ гэх мэт шийдвэрийг гаргахад хэрэгтэй болно.

Их хэмжээний өгөгдөл, аналитик нь бизнесүүдийн өрсөлдөөний давуу талыг бий болгох стратегийн чухал талбар болж байна. Олон улсын мэдээллийн технологийн судалгааны Gartner Research байгууллагын хийсэн судалгаагаар компаниудын мэдээллийн технологийг ашиглах чиглэлээр их хэмжээний өгөгдөл, аналитик нь 2018 он хүртэл гол, чухал чиг хандлага байна хэмээн үзэж байна.

Бизнесийн байгууллагууд үр дүн, ашгаа дээшлүүлэхийн тулд их хэмжээний мэдээллийг ашиглан аналитик хийж хөгжлийнхээ чиг хандлагыг тодорхойлох боломжтой болжээ. Бизнесийн орчинд хийсэн судалгаанууд бизнесийн компани их хэмжээний мэдээллийг ашиглахын тулд 1) юуны өмнө зардлыг оновчтой болгох, шинээр үнэ цэнийг бий болгохын тулд ямар өгөгдлүүд байгаа, ямар өгөгдлүүд шаардлагатай байгааг тодорхойлох; 2) зах зээлд эдийн засгийн үндэслэлтэй, үр ашигтай гарах, их хэмжээний мэдээллийг удирдах, шинжилгээ хийхэд оновчтой шийдлийг сонгохын тулд бат бэх түншлэлийн харилцааг тогтоох; 3) энэ чиглэлд хэрэгжиж буй хууль эрх зүйг тооцох; 4) хурдан үр өгөөжөө өгөх бодит төслүүдийг байнга хэрэгжүүлэх, удаан хугацааны төлөвлөгөө боловсруулах нь зүйтэй гэсэн дүгнэлтийг өгч байна. Түүнчлэн Big Data-г хөгжүүлэхэд энэ чиглэлээр чадварлаг боловсон хүчин бэлтгэх шаардлагатай. Энэ чиглэлээр нарийн мэргэшсэн мэргэжилтэн хомс байгаа бөгөөд цаашдаа үүний хэрэгцээ шаардлага улам өсөхөөр байгаад дүгнэлт хийж мэргэжилтэнгээ бэлтгэхэд анхаарах хэрэгтэй байна.

### **Дүгнэлт**

Өнөө үед дэлхий даяар их хэмжээний өгөгдөл (Big Data), түүнд дүн шинжилгээ хийж (Analytics) төр, олон нийт, бизнесийн байгууллагын үйл ажиллагаа болон шийдвэр гаргахад ашиглах талаар ярьж, томоохон санаачлага хэрэгжүүлж байна.

Улс үндэстнүүд их хэмжээний өгөгдлийг үндэсний нэг чухал нөөц сурвалж гэж үзэж, түүний аюулгүй байдлыг хамгаалах чиглэлээр бодлого хэрэгжүүлж, ухаалаг ашиглаж байна. Ухаалаг төр, ухаалаг хотын чиглэлээр хэрэгжиж буй томоохон төслүүд их хэмжээний өгөгдөл, түүний дүн шинжилгээнд төвлөрөх болжээ.

Мэдээллийн технологи бол улс орны хөгжлийн хурдасгуур гэж дэлхийн улс орнууд, тухайлбал Монгол улс ч хөгжлийн баримт бичгүүдэд тунхагласан байдаг. Иймд их хэмжээний мэдээлэлтэй ажиллах дэд бүтцийг байгуулах, түүнд тулгуурласан хэрэглээг хөгжүүлэхийн тулд мэдээллийн технологийн болон боловсролын салбарт хөгжлийн урт хугацааны, тууштай бодлого баримтлах нь зүйтэй байна.

Манай бизнесийн байгууллагууд үр дүнгээ сайжруулахын тулд их хэмжээний мэдээллийг ашиглан аналитик хийж хөгжлийнхээ чиг хандлагыг тодорхойлох боломжийг ашиглах нь цаг үеэ олсноос гадна шаардлагатай болсон байна.

### **Ашигласан материал**

- “Big Data & Analytics Conference 2015” -д тавьсан “Төр засгийн үйл ажиллагаанд их хэмжээний өгөгдлийн шинжилгээг хэрэгжүүлэх нь”, Үндэсний Дата Төв, УТҮГ-ын Захирал Э.Энхболдийн илтгэл, 2015
- Hilbert and Lopez, ‘The world’s technological capacity to store, communicate, and compute information, Science, 2011 Global
- IDC, The digital universe in 2020: Big Data, Bigger Digital Shadow s, and Biggest Growth in the Far East, 2012
- John Gantz and David Reinsel, IDC Digital Universe in 2020, Big Data..in United States, 2013
- KPMG, Using Big Data Still a Challenge for Marketers, emarketer.com, 2014
- <http://www.itzone.mn/en/news/73-Ай Ти Зон компанийн дэд захирал Б.Цэесүрэн “Их хэмжээний өгөгдөл ба орчин үеийн бизнес аналитикийн хандлага” илтгэл>
- [https://ru.wikipedia.org/wiki/Большие\\_данные](https://ru.wikipedia.org/wiki/Большие_данные)
- <http://www.seas.harvard.edu/computational-science-and-engineering/master-science-cse>. Bloomberg- <http://www.ikon.mn/n/kxg>