

COURSE: INTRODUCTION TO INFORMATION SYSTEM

Lecture 5: Big Data

Lecturer:

SARANCHIMEG Nasanjargal

1-Р ХЭСЭГ: ИХ ӨГӨГДӨЛ

Өгөгдөл бол байгууллагын бусад чухал хөрөнгүүдийн адилаар удирдаж, зохион байгуулж байх шаардлагатай үйл ажиллагааны чухал нөөц юм. Бизнесийн байгууллагууд өөрсдийн гадаад болон дотоод үйл ажиллагааны талаарх чанартай найдвартай мэдээлэлгүйгээр оршин тогтнож, амжилтад хүрч чадахгүй. Тиймээс байгууллага болон байгууллагын удирдлага, менежерүүд өгөгдлийн нөөцийн удирдлагын талаар тодорхой практик мэдлэг туршлагатай байх хэрэгтэй юм. Өдөр бүр байгууллагын төрөл бүрийн системүүдээр маш их хэмжээний өгөгдөл мэдээлэл орж байгаа бөгөөд тэдгээр өгөгдлүүдийг оновчтой зөв зохион байгуулах нь байгууллагын гол асуудал болоод байна. Жишээ нь хэрвээ чи маркетингийн менежер болох бол хэрэглэгчдийн сонирхол болон өнгөрсөн худалдан авалтын талаар асар их мэдээлэл бүхий судалгааны өгөгдөлтэй ажиллах бөгөөд эдгээр мэдээлэл нь чамд бүтээгдэхүүн үйлчилгээгээ зарахад хамгийн чухал учраас тэдгээр их өгөгдөлтэй хэрхэн үр дүнтэй ажиллах нь чиний өгөгдлийн талаарх мэдлэг, чадвар тус болох юм.

БИЗНЕСИЙН КЭЙС ЖИШЭЭ: BESTBUY ХЭРЭГЛЭГЧДИЙН МЭДЭЭЛЛИЙГ АШИГЛАН ҮР ӨГӨӨЖТЭЙ МАРКЕТИНГ ЯВУУЛАХ

“Bestbuy компани 75сая өрхийн талаар 7 жилийн турш цуглуулсан 15Tb бүхий өгөгдлийн сантай. Уг өгөгдлийн санд хэрэглэгчтэй утсаар хийсэн ярилцлагаас гадна сайтад зочилсон мэдээлэл, гүйлгээний талаарх бүхий л мэдээлэл байх бөгөөд тэдгээрийг маш нарийн алгоритм ашиглан боловсруулж нийт өрхүүдийн 3ны 2 буюу 100 сая хувь хүний мэдээллийг хэд хэдэн ангилалд хуваасан байдаг. Ингэхдээ хэрэглэгчидээ Buzz (сүүлийн үеийн шинэ технологи ашиглах сонирхолтой залуус), Party (чинээлэг өндөр боловсролтой залуус), Ray (гэр бүлээ хамгаас эрхэмлэдэг хүмүүс) гэж ангилдаг. Мөн компаний хэрэглэгч байх нийт хугацааны үнэ цэнийг (Customer Lifetime value) тооцож гүйлгээний ашигт байдал болон худалдан авалтын хэмжээний өсөлт эсвэл бууралтаар илрэх хэрэглэгчийн зан төлөвийн өөрчлөлтийг ажилглаж байдаг. Ийнхүү өөрийн хэрэглэгчийн талаарх нарийн мэдээллийг бий болгосноор хувь хүнд зориулсан маркетинг болон хэрэглэгчдийг өдөөх идэвхижүүлэлтийн хөтөлбөрийнхөө үр дүнг нэмэгдүүлж чаджээ.” (Майкл И.Портер, 1985)

Их өгөгдлийн хөгжил

Их өгөгдөл гэх ойлголт харьцангуй шинэ хэдий ч их хэмжээний өгөгдлийн тухай ойлголт 1960-1970 оны үеэс эхтэй юм. Тухайн үед анхны өгөгдлийн төвүүд, өгөгдлийн бааз хөгжүүлэлт эхэлж байсан түүхтэй. Харин “Их өгөгдөл” гэх нэршлийн хувьд 1990-ээд оны эхээр хэрэглэгдэж эхэлсэн байна. Анх 1989-1990 оны хооронд Тим Бернерс-Ли (Tim Berners-Lee), Роберт Кайо (Robert Cailliau) нар World Wide Web платформыг бүтээж, улмаар HTML, URL, HTTP-г хөгжүүлсэн бөгөөд интернэтийн эрин үетэй зэрэгцэн 1996 оноос өгөгдлийг цаасаар бус электрон хэлбэрээр хадгалах нь илүү зардал багатай болж, тэрхүү мэдээлэл рүү хандахад ч хялбар болсон байна.

Улмаар 1998 оноос www.google.com домайн үйл ажиллагаагаа эхлүүлсэн бөгөөд тус онд Карл Стрози “NoSQL” өгөгдлийн сангийн хэлийг хөгжүүлсэн түүхтэй. Улмаар 1999 онд Хэл

Р.Вэриан (Hal R. Varian), Питер Лайман (Peter Lyman) нар “How much information” номынхоо анхны хувилбарыг бичиж, тус номдоо дэлхий дээрх нийт дижитал мэдээллийн хэмжээг тодорхойлохыг оролдож байв.

Түүний дараагаар 2000 оноос хойш интернэт болон вебийн тусламжтайгаар өмнө нь байгаагүй их хэмжээний мэдээллийг цуглуулах, шинжлэх боломж нээгджээ. Тухайлбал, вебийн ачаалал, онлайн дэлгүүрийн тоо нэмэгдсэний дээр Yahoo, Amazon, eBay зэрэг томоохон компаниуд үйл ажиллагаа эхэлж, цаашлаад хэрэглэгчдийнхээ зан төлөвийг даралтын хувь, байршил, хайлтын түүх зэргээс нь шинжиж мэдэхээр ажилласан. Энэ нь цаашлаад өгөгдлийн шинэ ертөнцийг бидэнд нээн өгч байна.

Их өгөгдлийн өсөлт хөгжил нь “Hadoop” зэрэг нээлттэй эхийн технологийн хөгжилтэй салшгүй холбоотой юм. Facebook, YouTube болон бусад онлайн платформ өөрсдийн хэрэглэгчдийн талаар маш их хэмжээний өгөгдөл үйлдвэрлэж байгаа нь олны анхаарлыг татах болов. Улмаар 2005 онд “Hadoop” технологийн хөгжүүлэлт дуусаж, олон нийтэд танилцуулагдаж, “NoSQL” хэл нэрд гарч эхлэв. Эдгээр технологийн тусламжтайгаар өгөгдлийг хямд зардлаар хадгалах боломжтой болсон. Улмаар сүүлийн хэдэн жилд их өгөгдлийн хэмжээ асар хурдтай нэмэгдсэн бөгөөд хэрэглэгчид одоо ч их хэмжээний өгөгдлийг тасралтгүй үйлдвэрлэж байна.

Компьютерын хүчин чадал хурдтай нэмэгдэж буй энэ үед “их өгөгдөл” гэх ойлголтыг тоон утгаар илэрхийлж, хязгаар тогтоох боломжгүй юм. Анх 1965 онд Гордон Мүүр (*Gordon Moore*) нэг чипт багтаах транзиторын тоо буюу компьютерын хүчин чадал 2 жил тутамд 2 дахин нэмэгдэх таамаглал дэвшүүлсэн нь биеллээ олж хожмоо Мүүрийн хууль хэмээн алдаршсан байдаг. Сүүлийн 60 орчим жилийн тооцоогоор компьютерын хүчин чадал 18 сар тутамд 2 дахин нэмэгджээ. Гэсэн хэдий ч Гордон Мүүрийн үзэж буйгаар нэг чипт багтаах транзиторын тоог хязгааргүй ихэсгэх боломжгүй бөгөөд компьютерын хүчин чадлын өсөлтийн хурд цаашид саарч болзошгүй байна.

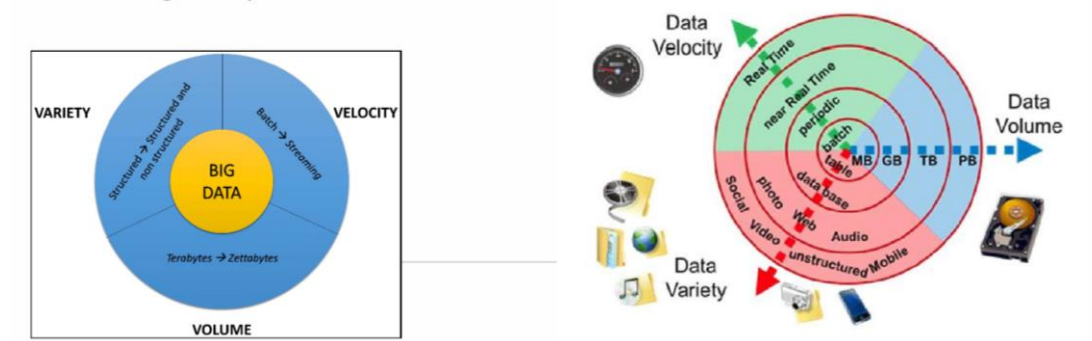
Өдгөө өгөгдөл хөрөнгийн нэг хэлбэр болж хувирч байна. Дэлхийн томоохон технологийн компаниудын хувьд тэдний бүрдүүлсэн өгөгдөл нь байгууллагын үнэ цэнийг тодорхойлох томоохон хүчин зүйл болохын зэрэгцээ шинэ бүтээгдэхүүн боловсруулах гол орц болж байна. Өгөгдлийн хэмжээ нэмэгдэж, хадгалах зардал буурахын хэрээр оновчтой бизнесийн эсвэл бодлогын шийдвэрт ашиглаж байна.

Их өгөгдлийн үнэ цэн нь танд хэр их өгөгдөл байгаагаас бус байгаа өгөгдлийг хэрхэн ашиглаж байгаагаас хамаарна. Өгөгдлийг цэгцлэх, шинжилгээнд тохируулж бэлдэх нь ихээхэн цаг шаардах бөгөөд өгөгдлийг шинжилдэг хүмүүс нийт ажлын цагийнхаа 50-аас 80 хувийг зөвхөн өгөгдлийг цэгцлэж бэлдэхэд зарцуулдаг гэх тооцоо бий. Олон улсын өгөгдлийн корпорацаас (*IDC*) гаргасан тооцоогоор дэлхий нийтээр бизнесийн зорилгоор цуглуулсан бүтэцгүй өгөгдлийн 90 хувьд огт шинжилгээ хийж чадахгүй байна. Шинэ технологи гарч ирснээр өгөгдлийг хадгалах боломж нэмэгдэж байгаа ч өгөгдлийн хэмжээ хоёр жил тутамд хоёр дахин нэмэгдэж байгаа гэдгийг санах нь зүйтэй юм. Тиймээс байгууллагууд тэр их мэдээллийг зохицуулах үр дүнтэй арга хайсан хэвээр л байна. (Д.Болдбаатар, 2023)

Их өгөгдлийн үндсэн ойлголт

(Д.Болдбаатар, 2023) “Их өгөгдөл” нь уламжлалт арга техникээр шинжилж, үнэ цэнтэй мэдээлэл гарган авахад хүндрэлтэй, зарим тохиолдолд бүр боломжгүй гэж хэлж болох асар их

хэмжээний, хурдтай нэмэгдэж байгаа өгөгдлийг хэлнэ. Их өгөгдлийн шинж чанарын талаарх Doug Laneу-ын 2001 онд дэвшүүлсэн “Турван V”-н их өгөгдлийг ойлгох, авч үзэх, уламжлалт өгөгдлөөс ялган таних гол суурь болсон. (Laneу, 2001)-д анх тодорхойлсноор их өгөгдөл нь “volume”, “velocity”, “variety” гэх 3 шинжийг агуулсан байх ёстой байна. Үүнээс гадна сүүлийн жилүүдэд эрдэмтэд, судалгааны байгууллагуудын дэвшүүлээд буй нэмэлт хоёр “V” буюу “veracity”, “value”-г энд нэмж дурдаж болно.



Зураг 1. Их өгөгдлийн шинж чанар

(Д.Болдбаатар, 2023) Их өгөгдлийн дээрх шинж чанарууд нь:

1. **“Volume” (хэмжээ)** – Технологийн дэвшилтэй уялдан өгөгдлийг цуглуулах, хадгалах зардал үлэмж буурсан, хувь хүний ердийн амьдралын хэв маягаар өгөгдөл бий болж дэлхий дээрх нийт өгөгдлийн хэмжээг нэмэгдүүлэх боломж ихэссэн. Өгөгдлийн хэмжээ маш их байх нь их өгөгдөл гэж нэрийдэх эхний шалгуур болох юм.
2. **“Velocity (хурд)** – Электрон төхөөрөмжийн тоо нэмэгдэхтэй зэрэгцэн бизнесийн ертөнц дэх мэдээллийн урсгал асар хурдасч, түүнийг цаг тухайд нь шинжлэх шаардлага нэмэгдэж байна. Тиймээс өгөгдөл үүсэх, дамжуулах хурд нь их өгөгдлийг илэрхийлэх нэг чухал үзүүлэлт гарцаагүй мөн билээ.
3. **“Variety” (төрөл)** – Өгөгдөл ямар ч хэлбэртэй байж болно. Өнгөрсөн хугацаанд цэгцэлж, эрэмбэлсэн бүтэцтэй тоон өгөгдлийг ихээр ашигладаг байсан бол өгөгдлийн шинэ эринд тодорхой бүтэцгүй үг, өгүүлбэр, и-мэйл, видео бичлэг, дуу авиа, санхүүгийн шилжүүлэг гэх зэрэг олон төрлийн мэдээллийг ашиглах болсон.
4. **“Veracity” (үнэн зөв байдал)** – Энэ нь өгөгдлийн эх үүсвэр, төрөл, эхний боловсруулалт зэрэгт итгэж болохуйц байх тухай ойлголт юм. Илүү нарийвчлал өндөртэй, итгэл даахуйц өгөгдлийн хэрэгцээ үргэлж байдаг ч илүү хямд, илүү их өгөгдөл цуглуулах тохиолдолд нарийвчлал, найдвартай байдлыг хайхрахгүй орхих явдал элбэг тохиолддог. Тухайн өгөгдөл үнэн зөв байх нь чухал байдлаараа тэргүүн эгнээнд эрэмбэлэгдэнэ.
5. **“Value” (үнэ цэн)** – Энэ нь цуглуулсан өгөгдлийг ашиглан үнэ цэн бүтээх тухай ойлголт юм. Их өгөгдлийг үнэ цэн нь байгууллага цуглуулсан өгөгдлөөс өөрийн үйл ажиллагааг сайжруулахад шаардлагатай хэр их мэдээлэл гарган авч чадаж байгаагаас хамаардаг

Өгөгдлийн төрөл

Бүтэцлэгдсэн өгөгдөл нь тодорхой зорилгын хүрээнд цэгцтэй ангилан бэлдсэн мэдээллийг бүтэцтэй өгөгдөл гэх бөгөөд түүнийг шинжлэх арга техник сайн хөгжсөн байдаг. Одоогийн бидний ашиглаж буй ихэнх системүүд нь бүтэцчилэгдсэн өгөгдийн сангийн технологийг ашигладаг.

Employee #	Last name	First name	Hire date	Dept. number
005-10-6321	Johns	Francine	10-07-1997	257
549-77-1001	Buckley	Bill	02-17-1979	632
098-40-1370	Fiske	Steven	01-05-1985	598

Зураг 2. Ажилчдын мэдээллийн хадгалж буй хүснэгт

Өгөгдлийн сан. Логик уялдаа холбоотой өгөгдлийн элементүүдийн нэгдсэн цуглуулга юм. Өгөгдлийн сан нь өмнө нь тус тусдаа хадгалагдсан файлуудыг олон янзаар хэрэглэгдэх өгөгдлөөр хангах логик элементүүдийн цуглуулгаруу

нэгтгэдэг. Тиймээс өгөгдлийн сан нь нэгж болон нэгжийн хоорондын харилцаа холбоог дүрсэлсэн өгөгдлийн элементүүдийг агуулдаг.

Бүтэцлэгдээгүй өгөгдөл нь текст, аудио, видео гэх мэт янз бүрийн хэлбэрээр байж болох зохион байгуулахад ярвигтай мэдээллийг хэлнэ. Бүтэцлэгдээгүй өгөгдөл хамгийн их хувийг эзлэх хэдий ч зохион байгуулалт хамгийн их шаардлагатай бүтцийн хувьд түүхий бөгөөд замбраагүй гэж хэлж болох юм. Гэхдээ ихэнхи бүтэцлэгдээгүй өгөгдлийг бүтэцлэж болох ч энэ зардал ихтэй, цаг хугацаа шаарддаг. Мөн тэд бүгд бүтцийн загварт амархан хувирч чаддаггүй.

Хүснэгт 1. Бүтэцлэгдсэн, Хагас бүтэцчилэгдсэн өгөгдлийн ялгаа

	Бүтэцлэгдсэн	Бүтэцлэгдээгүй
Шинж чанар	<ul style="list-style-type: none"> Урьдчилан тодорхойлогдсон өгөгдлийн загвар Ихэвчлэн бичгэн мэдээлэл Хайхад хялбар 	<ul style="list-style-type: none"> Урьдчилан тодорхойлогдоогүй өгөгдлийн загвар Зураг, дуу, бичлэг, текст зэргээр өөр формат байж болно Хайлт хийхэд төвөгтэй
Хадгалалт	<ul style="list-style-type: none"> Холбогдсон өгөгдлийн бааз Өгөгдлийн агуулах 	<ul style="list-style-type: none"> Апליкейшн NoSQL дата бааз Өгөгдлийн агуулах Өгөгдлийн нуур
Өгөгдлийг үүсгэх	Хүн ба машинаар	
Холбогдох аппликейшн	<ul style="list-style-type: none"> Онгоцны захиалгын систем Нөөцийн удирдлага Хүний нөөцийн удирдлагын систем Интерпрайз удирдлагын 	<ul style="list-style-type: none"> Word боловсруулалт Пресантацийн программ Медиа үзэх, засах хэрэгсэл Хиймэл дагуул Мэдрэгчид
Жишээ	<ul style="list-style-type: none"> Он сар Утасны дугаар Регистр Дансны дугаар Үйлчлүүлэгчийн нэр Хаяг Бүтээгдхүүний нэр, код Шилжүүлгийн мэдээ 	<ul style="list-style-type: none"> Текст файл Тайлан Имэйл мессэж Аудио файл Бичлэгэн файл Зураг

Эх сурвалж 1. (Ioanna Constantiou, 2015)

Хагас бүтэцлэгдсэн өгөгдөл нь мэдээллийн бүлэглэлт, шаталбарыг бий болгодог тусдаа өгөгдлийн элементийг дотроо агуулдаг. Баримт бичиг болон мэдээллийн сан нь хоёулаа хагас бүтэцлэгдсэн байж болно. Энэ төрлийн өгөгдөл нь бүтэцлэгдсэн / хагас бүтэцлэгдсэн / бүтэцлэгдсэн мэдээллийн баазын зөвхөн 5-10% -ийг эзэлдэг боловч бизнесийн чухал хэрэглээтэй. (Г.Минжин, 2018) Хагас бүтэцлэгдсэн өгөгдлийн жишээ:

- Тэмдэглэгээний хэл XML хагас бүтэцлэгдсэн баримт бичгийн хэл юм. XML нь хүн ба машины уншигдах форматыг тодорхойлдог баримтжуулалтын кодчилолын дүрмүүд юм.
- Нээлттэй JSON стандарт, (JavaScript Object Notation) JSON нь өөр нэгэн хагас бүтэцлэгдсэн өгөгдөл солилцооны формат юм. Бүтэц нь хэлнүүдийн хооронд солигддог тул JSON нь вэбсайт болон серверүүдийн хооронд өгөгдөл дамжуулахад ашигтай.
- NoSQL Хагас бүтэцлэгдсэн өгөгдөл нь олон NoSQL ("зөвхөн SQL биш") өгөгдлийн сангийн чухал элемент юм. NoSQL өгөгдлийн сан нь өгөгдлөөс схемийг салгаж чаддаггүй учраас холбогдсон өгөгдлийн сангуудаас ялгаатай байдаг. Тиймээс NoSQL-ийг бичлэг болон хүснэгтийн форматад тохирохгүй, янз бүрийн мэдээллийг хялбараар хадгалах боломжийг олгодог. Энэ нь мөн мэдээллийн баазуудын хооронд өгөгдөл солилцох боломжийг амар болгодог.

Их өгөгдлийн боловсруулалт

1. **НЭГТГЭХ:** Өгөгдлийг төрөл бүрийн эх сурвалж, системүүдээс цуглуулах.

Өгөгдлийг цуглуулах эх сурвалжууд:

- Дотоод эх сурвалж:
 - Байгууллагын үйл ажиллагаан үр дүнд цугларсан өгөгдлүүд
 - Байгууллагын гадаад эх үүсвэрээс цуглуулсан эсвэл худалдан авсан мэдээлэл
- Гадаад эх сурвалж
 - Байгууллагын хяналтанд байдаггүй эх сурвалж

Эдгээр эх үүсвэрүүдэд хандах хандалт маш олон төрөл учираас хүлээн авах модулийн эх сурвалжид хандах нарийн төвөгтэй байдлыг хураангуйлах хэрэгтэй. Элемент бүр гарал үүслийн өөр өөр формат, бүтэцтэй

2. **ӨГӨГДЛИЙГ БЭЛДЭХ, УДИРДАХ:** Өгөгдлийг урьдчилан боловсруулах, цэвэрлэн баганад ижил утгыг өөр хэлбэрээр оруулсныг ижил болгох, дутуу утгыг бөглөх, өгөгдлийн бусад алдааг арилгахыг зорьдог.

Өгөгдөл бэлдэх

- Өгөгдлийг мета өгөгдөлд суурилан нэгтгэх
- Өгөгдлийн эх үүсвэрийг баталгаажуулах. Зарим шаардлагад нийцэхгүй байгаа эх сурвалжуудын өгөгдлийн алдааг тодорхойлон цэвэрлэх
- Өгөгдлийг шинжилгээнд бэлдэн шинж чанаруудыг үүсгэх

Өгөгдөл нэгтгэх янз бүрийн өгөгдлүүдээ нэгтгэх, байгаа бүх өгөгдөлд ижил төрлийн хандалтын аргыг бий болгох

Өгөгдлийг хуваах. Загварыг сургахын тулд өгөгдлийн дэд хэсгээр салгаж, сургаж шинэ өгөгдөл оруулахад загвар хэр сайн эсвэл муу ажиллаж байгааг мэдэхэд хэрэглэгдэнэ.

- 3. ШИНЖИЛГЭЭ ХИЙХ:** Бэлтгэсэн өгөгдлүүдэд статистик болон машин сургалтын аргуудыг хэрэглэж, мэдээллийг гарган авах болон ирээдүйн таамаглал хийнэ.
- Decision tree: Өгөгдлийг жижиг хэсгүүдэд задлан хувааж, хэсэг бүрийн ерөнхий энтропийг багасгах
 - Clustering: Өгөгдлийн багц дотороос хэрэгтэй утгыг олох явцын автоматжуулалт
 - Text analysis: Текст аналитикийн алгоритм нь унших үйл явцыг автоматжуулж, магадгүй мянга мянган баримт бичгүүдийг нэгтгэн, товч хураангуй бэлддэг.
 - Neural network: Хүний тархиний мэдрэлийн сүлжээний загвар дээр суурилсан ба NODE-үүд дохиогоор идэвхжиж, бусад NODE-үүдэд идэвхжих хариу дохиог дамжуулдаг.

Их өгөгдөл: Нэр томъёо, технологи

“Data lake” (Өгөгдлийн нуур) нь нуурт цутгах ус адил олон эх үүсвэрээс хүлээн авсан задгай мэдээллийг анхны форматаар хадгалах сав/нуур юм.

“Data Lakehouse” (Өгөгдлийн нуурын байшин) Data lake болон Data warehouse-ын элементүүдийг хооронд нь нэгтгэж зардал багатайгаар хадгалах хэсгийг ийнхүү нэрлэдэг.

“Data warehouse” (Өгөгдлийн агуулах) нь ялгаатай эх үүсвэрээс мэдээллийг цуглуулж, хадгалж түүнээс үнэ цэнтэй мэдээлэл гаргаж өгдөг боловсронгуй электрон агуулах юм.

“SQL” болон “NoSQL” – “SQL (Structured Query Language)” нь бүтэцтэй өгөгдлийг хүснэгт (мөр, баганатай) хэлбэрээр өгөгдлийн санд хадгалах, сангаас өгөгдөл татан авах үйлдлийг энгийн, үр дүнтэй гүйцэтгэх үндсэн арга юм. Харин “NoSQL (Not only SQL)” бүтэцгүй өгөгдөлтэй ажиллахад тохиромжтой юм.

Их Өгөгдлийн систем	Өгөгдлийн агуулахын систем
Яг одооны өгөгдлийг хадгалдаг	Түүхэн өгөгдлийг хадгалдаг
Нарийвчилсан өгөгдлийг хадгалдаг	Нарийвчилсан болон нэгтгэн дүгнэсэн өгөгдлүүдийг хадгалдаг
Өгөгдөл нь хувирамтгай	Өгөгдөл нь тогтмол
Мэдээлэл гаргаж авахын тулд дахин дахин боловсруулах үйл ажиллагаа хийдэг	Цаг хугацааны давтамжтайгаар боловсруулах үйл ажиллагаа хийгддэг
Үйл ажиллагаатай холбоотой мэдээллийг хадгалдаг	Судалгаа шинжилгээнд тулгуурласан мэдээллийг хадгалдаг
Тодорхой хэрэглээний програмтай хамтарч ажилладаг	Судалгаа шинжилгээний чиглэлийн програмуудтай хамтарч ажилладаг
Өдөр тутмын шийдвэрийг дэмждэг	Стратеги шийдлийг дэмждэг
Олон тооны хэрэглэгчтэй	Цөөн тооны менежерүүдэд үйлчилгээ үзүүлдэг

Эх сурвалж 2: (Г.Минжин, 2018)

Зураг 3. Их өгөгдөл, Өгөгдлийн агуулахын ялгаа

Их өгөгдлийн хэрэглээ

(Tiao, 2024)

Бүтээгдэхүүн хөгжүүлэлт: Одоогийн бүтээгдэхүүн үйлчилгээний өмнөх загварууд, тэдгээрийн шинж чанарууд, холбоо хамаарал, борлуулалтын амжилт зэрэгт үндэслэн шинэ бүтээгдэхүүн үйлчилгээний загваруудыг боловсруулах, таамаглах

Засвар сайжруулалтыг урдчилан таамаглах: Асуудлыг тохиолдохоос өмнө илрүүлэн засвар үйлчилгээг үр ашигтай байдлаар төлөвлөх, сэлбэг хэрэгслүүдийг үр дүнтэй ашиглах

Хэрэглэгчийн туршлага: Сошиал мэдиа, вэб хандалт, дуудлагын мэдээлэл зэргээс өгөгдлүүдийг цуглуулан хэрэглэгчийн харилцааг нэмэгдүүлэх, үнэ цэнийг нэмэгдүүлэх. Цаашлаад илүү хувь хүнд чиглэсэн бүтээгдэхүүн үйлчилгээ санал болгох, хэрэглэгчийн сэтгэл ханамжыг нэмэгдүүлэх

Луйвар, хяналт зохицуулалт: Их хэмжээний өгөгдлөөс луйврын шинжтэй өгөгдлийн холбоо хамаарлыг илрүүлэх, хууль эрхзүйн тайлан, мэдээллийг хурдасгах

Үр дүнтэй үйл ажиллагаа: Ирээдүйн эрэлтийг таамаглах, илүүдэл, үлдэгдлийг бууруулахын тулд үйлдвэрлэл, хэрэглэгчийн эргэх холбоо, буцаалт зэргийг шинжлэх, үнэлэх. Одоогийн зах зээлийн эрэлтэд нийцүүлэн шийдвэр гаргалтыг сайжруулахад ашиглагддаг.

Машин сургалт: Машинуудыг програмчлахын оронд их өгөгдлийг ашиглан сургах.

Инноваци: Хүмүүс, байгууллагууд, нэгжүүд, процессуудын харилцаа, хамаарлыг илрүүлэх замаар инноваци бүтээх арга замуудыг тодорхойлох. Санхүүгийн гэх мэт шийдвэр гаргалтуудыг өгөгдлийн тусламжтайгаар сайжруулах гэх мэт хязгааргүй боломжуудыг бий болгож байна.

Их өгөгдлийн асуудлууд

- Өгөгдлийн бодлого, цахим орчны ёс зүй, нийгмийн сэтгэл зүй нь бизнесийн таатай нөхцөл бүрдүүлэхүйц байх
- Нийтээр хүлээн зөвшөөрсөн ёс зүйн хэв шинж бий болгох үүнээс чухал хэрэгтэй
- Мэдээллийг нийтийн ашиг сонирхлын эсвэл арилжааны зорилгоор хэрэглэхдээ түүний олдц, сонголтыг хязгаарлахгүй байх тэнцвэрийг олох,
- Боловсруулалт, шинжилгээний үр өгөөжийг нийтэд ил хэрэглэхдээ хүн, нийгмийн бүлэг, хуулийн этгээдийн ашиг сонирхолд харшлах түвшинг зохистой хэмжээнд хязгаарлах,
- Өдөр тутмын амьдрал, хүн хоорондын харилцаанаас үүсэж байгаа өгөгдлийг ашиглан өртөг бүтээхдээ “хүний хувийн мэдээлэл”, эрхийг хүндэтгэх; инновацийг урамшуулахдаа нийтийн ашиг сонирхлын зорилготой байлгах тэнцвэрийг хангах,
- Хүн, хуулийн этгээдийн мэдээллийг хамгаалахдаа хувь хүн, нийтийн эрх ашгийн төлөө мэдээлэл хуваалцах, хэрэглэх, шинжлэх орон зайг хангахаар зохион байгуулах буюу өгөгдлийг шинэ аргаар боловсруулахыг сэдэлжүүлэхдээ уг мэдээлэл нь нийтийн эрх ашгийн төлөө дамжих, зарагдах боломжтой байхыг чухалчлах;
- Мэдээлэл дамжиж байгаа үйл ажиллагааны ил тод байдлыг хангах ёс зүйтэй байх зэрэг асуудлуудыг тусгасан байна.

БҮЛГИЙН АШИГЛАСАН ЭХ СУРВАЛЖ

- Ioanna Constantiou, J. K. (2015). Шинэ тоглоом, шинэ дүрмүүд: Их өгөгдөл стратегийн тодорхойлолтыг өөрчлөх нь. *Journal of Information Technology*.
- James A. O'Brien, G. M. (2010). *Introduction to information systems*. New York: McGraw-Hill/Irwin.
- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety.
- Ralph M. Stair, G. W. (2018). *Principles of information system*. Boston: Cengage Learning.
- Тiao, S. (2024 оны 03 11). *Big data*. Oracle: <https://www.oracle.com/big-data/what-is-big-data/#best-practices-ээс> Гаргасан
- Г.Минжин. (2018). *Их өгөгдлийн шинжилгээг хөрөнгө оруулалтын шийдвэр гаргалтад ашиглах нь*. Улаанбаатар: СЭЗИС.
- Д.Болдбаатар, Б. Б. (2023). Их өгөгдлийн төв банкин дахь хэрэглээ. *Монгол банкны судалгааны эмхэтгэл*.
- Майкл И.Портер, В. И. (1985). Мэдээлэл өрсөлдөх давуу талыг хэрхэн бий болгодог вэ. *Харвардын бизнесийн сэтгүүл*.