Big Data - အချက်အလက်ပင်လယ်သမုဒ္ဒရာကြီး၏ အနှစ်သာရ၊ စွမ်းအားနှင့် အနာဂတ် မျှော်မှန်းချက်များ

၁။ ၂၁ ရာစု နည်းပညာခေတ်တွင် "Big Data" ဟူသော ဝေါဟာရသည် နေရာအနှံ့အပြားတွင် ပျံ့နှံ့နေပြီး ခေတ်ရေစီးကြောင်းတစ်ခုအဖြစ် အခိုင်အမာ ရပ်တည်လျက်ရှိသည်။ Big Data ဆိုသည်မှာ ရိုးရိုးရှင်းရှင်းအားဖြင့် ကြီးမားလွန်း၊ ရှုပ်ထွေးလွန်းသော အချက်အလက် အစုအဝေးများကို ရည်ညွှန်းခြင်းဖြစ်သည်။ ဤအချက်အလက်များသည် သမားရိုးကျ Data Processing Software များဖြင့် ကိုင်တွယ်နိုင်သည့် ပမာဏထက် များစွာ ကျော်လွန်နေတတ်ပြီး ၎င်းတို့ကို စီမံခန့်ခွဲရန်၊ ခွဲခြမ်းစိတ်ဖြာရန် နည်းပညာအသစ်များနှင့် ချဉ်းကပ်မှုပုံစံအသစ်များကို မွေးဖွားစေခဲ့သည်။

Big Data ၏ အခြေခံလက္ခဏာများ (The 5 V's)

၂။ Big Data ၏ သဘောသဘာဝကို နားလည်ရန်အတွက် ၎င်း၏ အဓိကလက္ခဏာရပ်များဖြစ်သော "5 V's" ကို သိရှိရန် အရေးကြီးပါသည်-

(က) **Volume (ပမာဏ)။** Big Data ၏ အထင်ရှားဆုံး လက္ခဏာရပ်မှာ ၎င်း၏ ကြီးမားလှသော ပမာဏပင်ဖြစ်သည်။ ယခင်က Megabytes၊ Gigabytes များဖြင့် တိုင်းတာခဲ့ကြသော်လည်း ယနေ့ခေတ်တွင် Terabytes၊ Petabytes မှသည် Exabytes၊ Zettabytes၊ Yottabytes အထိ ကြီးမားသော Data ပမာဏများကို ကိုင်တွယ်နေရပြီဖြစ်သည်။ ဥပမာအားဖြင့်၊ ကမ္ဘာတစ်ဝှမ်းရှိ Facebook အသုံးပြုသူများမှ နေ့စဉ် Upload ပြုလုပ်နေကြသော ဓာတ်ပုံများ၊ ဗီဒီယိုများ၊ Status Update များ၊ Comment များ၏ စုစုပေါင်းပမာဏသည် အလွန်ပင် ကြီးမားလှသည်။ Google တွင် တစ်နေ့တာ ရှာဖွေမှုပြုလုပ်သည့် အကြိမ်အရေအတွက်၊ YouTube တွင် ကြည့်ရှုကြသည့် ဗီဒီယို နာရီပေါင်းများစွာ စသည်တို့သည်လည်း Big Data ၏ ကြီးမားလှသော ပမာဏကို သက်သေပြနေကြသည်။

(ခ) **Velocity (အလျင်)။** Big Data သည် အလွန်လျင်မြန်သောနှုန်းဖြင့် စီးဝင်၊ ထုတ်လုပ်၊ သိမ်းဆည်းနေခြင်းသည် ဒုတိယမြောက် လက္ခဏာရပ်ဖြစ်သည်။ ယနေ့ခေတ်တွင် အချက်အလက်များသည် အချိန်နှင့်တစ်ပြေးညီနီးပါး စီးဆင်းနေကြသည်။ ဥပမာ - ကမ္ဘာတစ်ဝှမ်းတွင် စက္ကန့်ပိုင်းအတွင်း ဖြစ်ပေါ်နေသော Credit Card Transaction များ၊ Social Media Platform များပေါ်တွင် တင်လိုက်သော Post များ၊ Like များ၊ Share များ၊ IoT Device များဖြစ်သည့် Sensor များမှ ရရှိသော အချက်အလက်များ၊ စတော့ရှယ်ယာဈေးကွက်၏ အတက်အကျ အချက်အလက်များ စသည်တို့သည် အလွန်လျင်မြန်သော နှုန်းထားဖြင့် စီးဝင်နေကြသည်။

Copyright Mar 20 2025, Myanmar Computer Emergency Response Team, mmCERT/cc

(ဂ) **Variety (အမျိုးအစား)**။ Big Data ၏ တတိယမြောက် အဓိကလက္ခဏာရပ်မှာ Variety ဖြစ်ပြီး ၎င်းသည် ဒေတာများ ပုံစံအမျိုးမျိုးဖြင့် တည်ရှိနေခြင်းကို ရည်ညွှန်းသည်။ ဤတွင်၊ ပုံစံသတ်သတ်မှတ်မှတ်ရှိပြီး Relational Database များတွင် ဇယားကွက် (Table) ပုံစံဖြင့် တွေ့ရလေ့ရှိသော ဝန်ထမ်း သို့မဟုတ် အရောင်း အချက်အလက်များ ကဲ့သို့သော Structured Data များ၊ တိကျသည့် ပုံစံ သတ်သတ်မှတ်မှတ် မရှိသော်လည်း Tag များ၊ MetaData များ ပါဝင်သောကြောင့် အတိုင်းအတာတစ်ခုအထိ ဖွဲ့စည်း တည်ဆောက်ပုံရှိသည့် JSON၊ XML၊ CSV ဖိုင်များကဲ့သို့သော Semi-structured Data များ၊ ထို့ပြင် မည်သည့်ပုံစံမှ မရှိသည့် စာသား၊ ရုပ်ပုံ၊ အသံ၊ ဗီဒီယိုဖိုင်များ၊ Email များ နှင့် Social Media Post များ ကဲ့သို့သော Unstructured Data များ ဟူ၍ အမျိုးအစား ကွဲပြားစွာ တည်ရှိနေကြသည်။ ဤအထဲတွင် Big Data ၏ ၈၀% ခန့်သည် Unstructured Data များ ဖြစ်သည်ကို တွေ့ရသည်။

(ဃ) **Veracity (တိကျမှန်ကန်မှု)**။ Big Data ၏ အရည်အသွေးနှင့် ယုံကြည်စိတ်ချရမှုသည် စတုတ္ထမြောက် အရေးကြီးသော လက္ခဏာရပ်ဖြစ်သည်။ Data များသည် မမှန်ကန်မှု များ၊ ကွဲလွဲမှုများ၊ ဆူညံသံ (Noise) များ ပါဝင်နိုင်သဖြင့် ၎င်းတို့ကို စိစစ်ပြီး အရည် အသွေးမြှင့်တင်ရန်လုပ်ငန်းစဉ်များ ဆောင်ရွက်ရန် လိုအပ်သည်။ အချက်အလက်များ၏ တိကျမှန်ကန်မှုသည် ခွဲခြမ်းစိတ်ဖြာမှု ရလဒ်များအပေါ် တိုက်ရိုက်သက်ရောက်မှုရှိပြီး မှားယွင်းသော ဆုံးဖြတ်ချက်များ ချမှတ်မိစေနိုင်သောကြောင့် အထူးဂရုပြုရမည့် အချက်ဖြစ်သည်။

(င) **Value (တန်ဖိုး)**။ Big Data ၏ အနှစ်သာရမှာ ၎င်းမှ အဖိုးတန်သော အသိပညာများ ထုတ်ယူနိုင်ခြင်းပင်ဖြစ်သည်။ Data များစွာရှိရုံဖြင့် အကျိုးမရှိဘဲ ၎င်းတို့ကို စနစ်တကျ ခွဲခြမ်းစိတ်ဖြာခြင်း ပြုလုပ်ခြင်းဖြင့် အဓိပ္ပါယ်ရှိသော Pattern များ၊ Trend များ၊ ဆက်နွယ်မှုများကို ဖော်ထုတ်ပြီး စီးပွားရေး၊ လူမှုရေး၊ သိပ္ပံပညာရပ်နယ်ပယ် အသီးသီး တွင် အကျိုးရှိစွာ အသုံးချနိုင်မှသာ တန်ဖိုးရှိသည်။ ဥပမာအားဖြင့်၊ ဖောက်သည်များ၏ ဝယ်ယူမှုပုံစံကို လေ့လာခြင်းဖြင့် မည်သည့်ကုန်ပစ္စည်းများကို မည်သည့်အချိန်တွင် အရောင်းမြှင့်တင်သင့်သည်ကို ဆုံးဖြတ်နိုင်ခြင်းသည် Big Data မှ ရရှိသော တန်ဖိုး တစ်ရပ်ဖြစ်သည်။

Big Data နည်းပညာများနှင့် သမိုင်းကြောင်း အကျဉ်း

၃။ Big Data ကို ကိုင်တွယ်ရန် သမားရိုးကျ နည်းပညာများဖြင့် မလုံလောက်တော့ပေ။ ထို့ကြောင့် အောက်ပါ နည်းပညာအသစ်များ ပေါ်ထွက်လာခဲ့ပြီး စဉ်ဆက်မပြတ် တိုးတက်ပြောင်းလဲနေဆဲ ဖြစ်သည်-

(က) **Hadoop (၂၀၀၅)။** Doug Cutting နှင့် Mike Cafarella တို့မှ ဖန်တီးခဲ့သော Apache Hadoop သည် Distributed Computing Framework တစ်ခုဖြစ်ပြီး Big Data ကိုင်တွယ်မှုအတွက် အခြေခံအုတ်မြစ် ဖြစ်လာခဲ့သည်။ Hadoop သည် ကြီးမားသော Data များကို ကွန်ပျူတာ အစုအဝေး (Cluster) များပေါ်တွင် ဖြန့်ဝေပြီး Parallel Processing ပြုလုပ်နိုင်သည်။ ၎င်း၏ အဓိက အစိတ်အပိုင်းများမှာ HDFS (Hadoop Distributed File System) နှင့် MapReduce တို့ ဖြစ်ကြသည်။ HDFS သည် Data များကို ဖြန့်ဝေသိမ်းဆည်းပေးပြီး MapReduce သည် Parallel Processing Model တစ်ခုဖြစ်ကာ ကြီးမားသော Data များကို အပိုင်းငယ်များခွဲ၍ သီးခြားစီ Process လုပ်ပြီး ရလဒ်များကို စုစည်းပေးသည်။

(ခ) **Spark (၂၀၀၉)**။ Matei Zaharia မှ University of California၊ Berkeley တွင် စတင်ဖန်တီးခဲ့သော Apache Spark သည် Hadoop ထက် ပိုမိုမြန်ဆန်သော In-memory Processing Engine တစ်ခုဖြစ်သည်။ Spark သည် Data များကို Memory ထဲတွင် ထားရှိပြီး Process လုပ်သောကြောင့် Hadoop ထက် အဆပေါင်းများစွာ ပိုမိုမြန်ဆန်သည်။ ၎င်းသည် Real-time Data Processing၊ Machine Learning၊ Graph Processing စသည့် လုပ်ငန်းများအတွက် အထူးသင့်လျော်သည်။

(ဂ) **NoSQL Databases**။ သမားရိုးကျ Relational Database များ (ဥပမာ - MySQL၊ Oracle၊ PostgreSQL) သည် Structured Data များအတွက် အဓိကထားသော်လည်း Big Data ခေတ်တွင် ပေါ်ပေါက်လာသော NoSQL Database များ (ဥပမာ - MongoDB၊ Cassandra၊ HBase၊ Redis) သည် Unstructured နှင့် Semi-structured Data များကို ကိုင်တွယ်ရန် ပိုမိုသင့်လျော်သည်။ ၎င်းတို့သည် Scalability နှင့် Performance ပိုမိုကောင်းမွန်ပြီး Horizontal Scaling (Server များ ထပ်တိုးခြင်း) ပြုလုပ်နိုင်သောကြောင့် ကြီးမားသော Data ပမာဏများကို ကိုင်တွယ်ရာတွင် ပိုမိုအဆင်ပြေသည်။

(ဃ) **Cloud Computing Platforms**။ Amazon Web Services (AWS)၊ Google Cloud Platform (GCP) နှင့် Microsoft Azure ကဲ့သို့သော Cloud Computing Platform များသည် Big Data Storage၊ Processing၊ Analysis အတွက် လိုအပ်သော အခြေခံအဆောက်အအုံနှင့် ဝန်ဆောင်မှုများကို ထောက်ပံ့ပေးသည်။ ၎င်းတို့သည် On-Demand Computing Resource များကို ပံ့ပိုးပေးသောကြောင့် လုပ်ငန်းများသည် ၎င်းတို့၏ လိုအပ်ချက်နှင့်အညီ အသုံးပြုနိုင်ပြီး ကုန်ကျစရိတ် သက်သာစေသည်။

Big Data ၏ အသုံးဝင်ပုံ နယ်ပယ်များ

၄။ Big Data သည် နယ်ပယ်ပေါင်းစုံတွင် အသုံးဝင်လျက်ရှိပြီး နေ့စဉ်ဘဝတွင် မသိမသာ ပါဝင်ပတ်သက်နေသည်။ စီးပွားရေးနယ်ပယ်တွင် Big Data သည် များစွာသော အသုံးချမှုများ ရှိပြီး ဈေးကွက်အနေအထားကို ပိုမိုနက်ရှိုင်းစွာ နားလည်ရန်၊ ဖောက်သည်များနှင့် ပိုမိုကောင်းမွန်သော ဆက်ဆံရေး တည်ဆောက်ရန်နှင့် လုပ်ငန်းလည်ပတ်မှုများကို ပိုမိုထိရောက်စေရန် ကူညီပေးသည်။ ဈေးကွက်ခွဲခြမ်းစိတ်ဖြာခြင်း အပိုင်းတွင် Big Data ကို အသုံးပြု၍ ဖောက်သည်များ၏ အပြုအမူ၊ စိတ်ကြိုက်ရွေးချယ်မှုနှင့် ဝယ်ယူမှုပုံစံများကို လေ့လာဆန်းစစ်ကာ ဈေးကွက်၏ လိုအပ်ချက်ကို ခန့်မှန်းနိုင်သည့်အပြင် ပြိုင်ဘက်များ၏ အားသာချက်၊ အားနည်းချက်များကိုလည်း လေ့လာပြီး ထိရောက်သော ဈေးကွက်ဗျူဟာများ ချမှတ်နိုင်သည်။ Customer Relationship Management (CRM) တွင်လည်း ဖောက်သည်တစ်ဦးချင်းစီ၏ အချက်အလက်များကို စုစည်းသိမ်းဆည်းကာ ၎င်းတို့၏ လိုအပ်ချက်များကို ဖြည့်ဆည်းပေးခြင်း၊ ပိုမိုကောင်းမွန်သော ဝန်ဆောင်မှုများပေးခြင်းနှင့် ရေရှည်ဆက်ဆံရေး တည်ဆောက်ခြင်းတို့ ဆောင်ရွက်နိုင်သည်။ ထို့အပြင် လိမ်လည်မှု ထောက်လှမ်းခြင်း အနေဖြင့် Credit Card လိမ်လည်မှု၊ အာမခံလိမ်လည်မှု စသည့် လိမ်လည်မှုများကို Real-time နီးပါး ထောက်လှမ်းနိုင်သည်။ အန္တရာယ် စီမံခန့်ခွဲမှုတွင်လည်း ငွေကြေးဆိုင်ရာ အန္တရာယ်၊ လည်ပတ်မှုဆိုင်ရာ အန္တရာယ်များကို ကြိုတင်ခန့်မှန်းပြီး လျှော့ချနိုင်သည်။ နောက်ဆုံးအနေဖြင့် Supply Chain Optimization တွင် ကုန်ပစ္စည်းများ ထုတ်လုပ်ခြင်း၊ သိုလှောင်ခြင်း၊ ဖြန့်ဖြူးခြင်း စသည့် လုပ်ငန်းစဉ်များ၌ ပိုမိုထိရောက်မှုရှိစေရန် Big Data analytics ကို အသုံးပြုသည်။

၅။ ကျန်းမာရေးနယ်ပယ်တွင် Big Data သည် ရောဂါများကို ပိုမိုနားလည်ရန်၊ ကုသရန်နှင့် ကာကွယ်ရန်အတွက် အရေးပါသော အခန်းကဏ္ဍမှ ပါဝင်လျက်ရှိသည်။ Big Data ကို အသုံးပြု၍ ရောဂါဖြစ်ပွားမှုနှုန်း၊ ပျံ့နှံ့မှုပုံစံနှင့် ရောဂါဖြစ်စေသော အကြောင်းရင်းများကို လေ့လာခြင်းဖြင့် ရောဂါဖြစ်ပွားမှု ပုံစံများကို ထဲထဲဝင်ဝင် သိရှိနားလည်နိုင်ပြီး ထိရောက်သော ကာကွယ်ရေးနှင့် ကုသရေး နည်းလမ်းများ ရှာဖွေဖော်ထုတ်ရာတွင် အထောက်အကူပြုသည်။ ထို့အတူ ကြီးမားသော ဓာတုဒြပ်ပေါင်းများ၏ အချက်အလက်များကို ခွဲခြမ်းစိတ်ဖြာခြင်းဖြင့် ဆေးဝါးသစ်များ ရှာဖွေဖော်ထုတ်ခြင်း လုပ်ငန်းစဉ်ကို ပိုမိုမြန်ဆန်ထိရောက်စေသည်။ လူနာတစ်ဦးချင်းစီ၏ မျိုးရိုးဗီဇ၊ လူနေမှုပုံစံနှင့် ပတ်ဝန်းကျင်ဆိုင်ရာ အချက်အလက်များကို အခြေခံ၍ အသင့်တော်ဆုံး ဆေးကုသမှု ပေးနိုင်မည့် လူတစ်ဦးချင်းစီအတွက် သီးသန့်ဆေးကုသမှု ဖြစ်ပေါ်လာစေရန်လည်း Big Data က အဓိကကျသော အခန်းကဏ္ဍမှ ပါဝင်သည်။ ဆေးရုံစီမံခန့်ခွဲမှုတွင်လည်း လူနာများ၏ အချက်အလက်များကို စနစ်တကျ စီမံခန့်ခွဲခြင်း၊ ဆေးရုံအရင်းအမြစ်များကို ထိရောက်စွာ အသုံးချခြင်းနှင့် လူနာစောင့်ရှောက်မှု အရည်အသွေး မြှင့်တင်ခြင်း စသည်တို့အတွက် များစွာ အထောက်အကူပြုသည်။

၆။ သိပ္ပံနယ်ပယ်တွင် Big Data သည် သဘာဝလောက၏ နက်နဲရှုပ်ထွေးသော ဖြစ်စဉ်များကို ပိုမိုနားလည်သဘောပေါက်ရန် အဓိကကျသော ကိရိယာတစ်ခု ဖြစ်လာသည်။ ရာသီဥတုပြောင်းလဲမှု လေ့လာခြင်းတွင် ကမ္ဘာတစ်ဝှမ်းမှ ရာသီဥတုဆိုင်ရာ အချက်အလက် အမြောက်အမြားကို စုဆောင်း၊ ခွဲခြမ်းစိတ်ဖြာပြီး ရာသီဥတုပြောင်းလဲမှု ဖြစ်စဉ်များကို နားလည်ရန်၊ အနာဂတ် ရာသီဥတု အခြေအနေများကို ခန့်မှန်းရန်နှင့် ယင်း၏ သက်ရောက်မှုများကို လျှော့ချရန် နည်းလမ်းများ ရှာဖွေရာတွင် အသုံးပြုသည်။ မျိုးရိုးဗီဇဆိုင်ရာ သုတေသနတွင်လည်း လူသားမျိုးရိုးဗီဇအပေါ်အဝင် သက်ရှိများ၏ မျိုးရိုးဗီဇ အချက်အလက် ကြီးမားစွာကို လေ့လာခြင်း၊ ရောဂါများနှင့် မျိုးရိုးဗီဇ ဆက်နွယ်မှုကို ဖော်ထုတ်ခြင်းနှင့် မျိုးရိုးဗီဇဆိုင်ရာ ရောဂါများကို ကုသနိုင်မည့် နည်းလမ်းများ ရှာဖွေရာတွင် အရေးပါသော အခန်းကဏ္ဍမှ ပါဝင်သည်။ ထို့အပြင် နက္ခတ္တဗေဒ ပညာရပ်တွင်လည်း အာကာသမှ ရရှိသော ကြီးမားလှသော အချက်အလက်များကို ခွဲခြမ်းစိတ်ဖြာခြင်း၊ ဂြိုဟ်များ၊ ကြယ်များ၊ ဂလက်ဆီများ၏ ဖြစ်ပေါ်မှု၊ ပြောင်းလဲမှုများကို လေ့လာခြင်း စသည်တို့အတွက် Big Data ကို မရှိမဖြစ် အသုံးပြုလျက်ရှိသည်။

၇။ အစိုးရ၏ လုပ်ငန်းဆောင်တာများတွင် Big Data သည် ပိုမိုကောင်းမွန်သော မူဝါဒများ ချမှတ်နိုင်ရန်၊ ထိရောက်သော ဝန်ဆောင်မှုများ ပေးနိုင်ရန်နှင့် ပြည်သူ့လုံခြုံရေးကို မြှင့်တင်နိုင်ရန် အတွက် အရေးပါသော အခန်းကဏ္ဍမှ ပါဝင်လာသည်။ မြို့ပြစီမံကိန်းတွင် မြို့ပြများ၏ လူဦးရေ ထူထပ်မှု၊ ယာဉ်ကြောပိတ်ဆို့မှု၊ အခြေခံအဆောက်အအုံ လိုအပ်ချက် စသည်တို့ကို လေ့လာဆန်းစစ်ပြီး ပိုမိုနေချင်စဖွယ်ကောင်းသော၊ သွားလာရလွယ်ကူသော၊ သဘာဝပတ်ဝန်းကျင်နှင့် သဟဇာတ ဖြစ်သော မြို့ပြများ ဖန်တီးရန်အတွက် အထောက်အကူပြုသည်။ ရာဇဝတ်မှု တိုက်ဖျက်ရေးတွင် ရာဇဝတ်မှု ဖြစ်ပွားမှုနှုန်း၊ ဖြစ်ပွားမှုပုံစံ၊ ဖြစ်ပွားမှုနေရာ စသည်တို့ကို လေ့လာဆန်းစစ်ပြီး ရာဇဝတ်မှုများကို ကြိုတင်ကာကွယ်နိုင်ရန်နှင့် ရာဇဝတ်ကောင်များကို ဖော်ထုတ်ဖမ်းဆီးရာတွင် အသုံးပြုသည်။ အလားတူပင် အကြမ်းဖက်မှု တန်ပြန်တိုက်ဖျက်ရေးတွင်လည်း အကြမ်းဖက်သမားများ၏ ဆက်သွယ်မှုများ၊ လှုပ်ရှားမှုများကို စောင့်ကြည့်လေ့လာပြီး အကြမ်းဖက်တိုက်ခိုက်မှု အန္တရာယ်မှ ကာကွယ်ပေးသည်။ ပြည်သူ့ကျန်းမာရေး ကဏ္ဍတွင်လည်း ရောဂါဖြစ်ပွားမှုများကို စောင့်ကြည့်လေ့လာခြင်း၊ ကူးစက်ရောဂါများ ပျံ့နှံ့မှုကို ထိန်းချုပ်ခြင်းနှင့် ထိရောက်သော ကျန်းမာရေးဆိုင်ရာ မူဝါဒများ ချမှတ်ခြင်းတို့တွင် Big Data ကို အသုံးပြုလျက်ရှိသည်။

၈။ လူမှုကွန်ရက်များတွင် Big Data သည် အသုံးပြုသူများ၏ အပြုအမူနှင့် သတင်းအချက်အလက် ပျံ့နှံ့မှုကို နားလည်ရန် အဓိကကျသော အခန်းကဏ္ဍမှ ပါဝင်သည်။ လူမှုကွန်ရက် အသုံးပြုသူများ၏ Post များ၊ Like များ၊ Share များ၊ Comment များ နှင့် Search Query များ အစရှိသည့် ကြီးမားလှသော အချက်အလက်များကို ခွဲခြမ်းစိတ်ဖြာခြင်းဖြင့် ၎င်းတို့၏ စိတ်ဝင်စားမှုများ၊ အပြုအမူများကို နားလည်နိုင်ပြီး ပိုမိုသက်ဆိုင်ရာ အကြောင်းအရာများနှင့် ပစ်မှတ်ထား ကြော်ငြာများကို ပြသနိုင်သည်။ ဤသို့ဖြင့် အသုံးပြုသူများအတွက် ပိုမိုအကျိုးရှိသော အတွေ့အကြုံကို ဖန်တီးပေးနိုင်သလို စီးပွားရေး

လုပ်ငန်းများအတွက်လည်း ကြော်ငြာထိရောက်မှုကို မြှင့်တင်ပေးနိုင်သည်။ တစ်ချိန်တည်းမှာပင် လူမှုကွန်ရက်များပေါ်တွင် ပျံ့နှံ့နေသော သတင်းမှားများ၊ သတင်းတုများကို ဖော်ထုတ်နိုင်ရန်နှင့် ၎င်းတို့၏ ပျံ့နှံ့မှုကို ထိန်းချုပ်နိုင်ရန်အတွက်လည်း Big Data Analytics ကို အသုံးပြုလျက်ရှိပြီး သတင်းအချက်အလက် အမှန်များ ရရှိရေးကို အထောက်အကူပြုသည်။

၉။ စက်မှုလုပ်ငန်းများတွင် Big Data သည် ထုတ်လုပ်မှု လုပ်ငန်းစဉ်များကို ပိုမိုကောင်းမွန်အောင် လုပ်ဆောင်ခြင်း၊ ကုန်ကျစရိတ်များ လျှော့ချခြင်းနှင့် စွမ်းဆောင်ရည် မြှင့်တင်ခြင်းတို့အတွက် အရေးပါသော အခန်းကဏ္ဍမှ ပါဝင်သည်။ စက်ပစ္စည်းများ၏ စွမ်းဆောင်ရည်ကို စောင့်ကြည့် လေ့လာခြင်းနှင့် ကြိုတင်ပြုပြင်ထိန်းသိမ်းမှုများ ပြုလုပ်ခြင်းဖြင့် စက်ရုံများရှိ စက်ပစ္စည်းများမှ Sensor Data များကို အချိန်နှင့်တပြေးညီ စုဆောင်းကာ ခွဲခြမ်းစိတ်ဖြာခြင်းကြောင့် စက်ပစ္စည်းများ ပျက်စီး နိုင်ခြေကို ကြိုတင်ခန့်မှန်းနိုင်ပြီး ပြုပြင်ထိန်းသိမ်းမှုများကို ကြိုတင်စီစဉ်ဆောင်ရွက်နိုင်ကာ စက်ရပ် နားချိန်ကို လျှော့ချပြီး ထုတ်လုပ်မှုစွမ်းအားကို မြှင့်တင်နိုင်သည်။ ထုတ်လုပ်မှုဖြစ်စဉ် တစ်လျှောက် လုံးတွင် Data များကို စုဆောင်းပြီး ခွဲခြမ်းစိတ်ဖြာခြင်းဖြင့် ထုတ်ကုန်များ၏ အရည်အသွေးကို စောင့်ကြည့်နိုင်ပြီး ချို့ယွင်းချက်များကို အချိန်မီ သိရှိနိုင်သည့် အရည်အသွေးထိန်းချုပ်မှု စနစ်ကိုလည်း Big Data အသုံးပြု၍ လုပ်ဆောင်နိုင်သည်။ ၎င်းအပြင် စက်ရုံများ၊ အဆောက်အအုံများတွင် စွမ်းအင်သုံးစွဲမှု ပုံစံများကို လေ့လာပြီး စွမ်းအင်သုံးစွဲမှု အလေအလွင့်များကို လျှော့ချနိုင်ရန်နှင့် စွမ်းအင် ချွေတာနိုင်မည့် နည်းလမ်းများ ရှာဖွေဖော်ထုတ်သည့် စွမ်းအင်စီမံခန့်ခွဲမှုတွင်လည်း အသုံးချနိုင်သည်။

Big Data ၏ စိန်ခေါ်မှုများ

၁၀။ Big Data သည် အကျိုးကျေးဇူးများစွာ ရှိသော်လည်း၊ လက်တွေ့အကောင်အထည်ဖော်ရာတွင် စိန်ခေါ်မှုများစွာလည်း ရှိနေပါသည်။

(က) **Data Privacy (အချက်အလက်လုံခြုံရေး)။** Big Data တွင် လူပုဂ္ဂိုလ်တစ်ဦးချင်းစီ၏ ကိုယ်ရေးကိုယ်တာ အချက်အလက်များ အမြောက်အများ ပါဝင်နိုင်သဖြင့် ၎င်းတို့ကို လုံခြုံစွာ ထိန်းသိမ်းရန်နှင့် ပုဂ္ဂိုလ်ရေးဆိုင်ရာလွတ်လပ်ခွင့်ကို ကာကွယ်ရန် အလွန် အရေးကြီးသည်။ ဥပဒေကြောင်းအရ ကျင့်ဝတ်သိက္ခာအရ ပုဂ္ဂိုလ်ရေးဆိုင်ရာလွတ် လပ်ခွင့်ကို လေးစားလိုက်နာရန် လိုအပ်ပြီး Data များကို ခွင့်ပြုချက်မရှိဘဲ အသုံးပြုခြင်း၊ ပေါက်ကြားခြင်းတို့မှ ကာကွယ်ရမည်။ GDPR (General Data Protection Regulation)၊ CCPA (California Consumer Privacy Act) စသည့် Data Privacy Regulation များကို လိုက်နာရန် လိုအပ်သည်။

(ခ) **အချက်အလက်လုံခြုံရေး။** Big Data များကို ခွင့်ပြုချက်မရှိဘဲ ဝင်ရောက်ကြည့်ရှုခြင်း၊ ပြင်ဆင်ခြင်း၊ ဖျက်ဆီးခြင်းတို့မှ ကာကွယ်ရန်အတွက် လုံခြုံရေး အစီအမံများ ချမှတ်

Copyright Mar 20 2025, Myanmar Computer Emergency Response Team, mmCERT/cc

ထားရန် လိုအပ်သည်။ Data Encryption၊ Access Control၊ Intrusion Detection System စသည့် နည်းပညာများကို အသုံးပြုပြီး Data security ကို မြှင့်တင်နိုင်သည်။

(ဂ) **အချက်အလက်စီမံအုပ်ချုပ်မှု။** Big Data ၏ အရည်အသွေး၊ တိကျမှန်ကန်မှု၊ ညီညွတ်မှု၊ ရရှိနိုင်မှုတို့ကို ထိန်းသိမ်းရန်အတွက် စည်းမျဉ်းစည်းကမ်းများ၊ မူဝါဒများ၊ လုပ်ငန်းစဉ်များချမှတ်ပြီး စနစ်တကျ စီမံအုပ်ချုပ်ရန်လိုအပ်သည်။ Data Governance Framework တစ်ခု တည်ဆောက်ပြီး Data Lifecycle တစ်လျှောက်လုံးတွင် Data Quality ကို စောင့်ကြည့်ထိန်းသိမ်းရမည်။

(ဃ) **အချက်အလက် ပေါင်းစည်းခြင်း။** မတူညီသော အရင်းအမြစ်များ၊ မတူညီသော ပုံစံများဖြင့် တည်ရှိနေသော Data များကို ပေါင်းစပ်ပြီး တစ်စုတစ်စည်းတည်းဖြစ်သော မြင်ကွင်းတစ်ခု ရရှိအောင် ပြုလုပ်ရခြင်းသည် နည်းပညာပိုင်းဆိုင်ရာ စိန်ခေါ်မှုတစ်ရပ်ဖြစ်သည်။ Data Integration Tools and Techniques များ (ဥပမာ - ETL (Extract၊ Transform၊ Load)) ကို အသုံးပြုပြီး Data များကို ပေါင်းစပ်ရသည်။

(င) **ကျွမ်းကျင်မှု ကွာဟချက်။** Big Data ကို ထိရောက်စွာ ကိုင်တွယ်နိုင်ရန်အတွက် ကျွမ်းကျင်သော Data Scientist၊ Data Engineer၊ Data Analyst၊ Data Architect စသည့် ပညာရှင်များ လိုအပ်သော်လည်း လက်ရှိတွင် ဤနယ်ပယ်တွင် ကျွမ်းကျင်သူ ရှားပါးမှု ပြဿနာ ရှိနေသည်။ Big Data နည်းပညာများကို သင်ကြားပေးနိုင်မည့် သင်တန်းများ၊ ပညာရေးအစီအစဉ်များ တိုးချဲ့ဆောင်ရွက်ရန် လိုအပ်သည်။

(စ) **အခြေခံအဆောက်အအုံကုန်ကျစရိတ်။** Big Data ကို သိမ်းဆည်းရန်၊ စီမံရန်အတွက် လိုအပ်သော Hardware နှင့် Software များသည် ဈေးကြီးနိုင်ပါသည်။

Big Data ၏ အနာဂတ်

၁၁။ Big Data သည် ဆက်လက်တိုးတက်ပြောင်းလဲနေဆဲ နည်းပညာတစ်ရပ်ဖြစ်ပြီး အနာဂတ်တွင်လည်း နယ်ပယ်အသီးသီးတွင် ပိုမိုအရေးပါသော အခန်းကဏ္ဍမှ ပါဝင်လာမည်မှာ သေချာပါသည်-

(က) **Artificial Intelligence (AI) နှင့် Machine Learning (ML) တို့၏ ပေါင်းစပ်မှု။** Big Data နှင့် AI၊ ML တို့သည် အပြန်အလှန် အမှီသဟဲပြုနေကြသည်။ Big Data သည် AI၊ ML Model များကို လေ့ကျင့်ရန်အတွက် လိုအပ်သော Data များ ပံ့ပိုးပေးပြီး AI၊ ML တို့သည် Big Data မှ အဖိုးတန်သော Insight များ ထုတ်ယူရာတွင် အသုံးပြုသည်။ နောင်တွင် AI၊ ML တို့သည် Big Data Analysis လုပ်ငန်းစဉ်များကို ပိုမို Automated ပြုလုပ်ပေးနိုင်မည်ဟု မျှော်လင့်ရသည်။

Copyright Mar 20 2025, Myanmar Computer Emergency Response Team, mmCERT/cc

(ခ) **Edge Computing ၏ အရေးပါမှု။** Edge Computing ဆိုသည်မှာ Data များကို Data Source (ဥပမာ - Sensor) နှင့် နီးကပ်သော နေရာ (ဥပမာ - စက်ရုံ၊ ကား) တွင် Process လုပ်ခြင်းဖြစ်သည်။ ဤသို့ဖြင့် Latency ကို လျှော့ချနိုင်ပြီး Real-time Processing ပြုလုပ်နိုင်သည်။ IoT Device များ တိုးပွားလာသည်နှင့်အမျှ Edge Computing သည် Big Data Processing တွင် ပိုမိုအရေးပါလာမည်ဖြစ်သည်။

(ဂ) **Quantum Computing ၏ အလားအလာ။** Quantum Computing သည် လက်ရှိ Classical Computer များထက် များစွာ ပိုမိုမြန်ဆန်သော Processing Power ကို ပေးစွမ်းနိုင်သော နည်းပညာသစ် ဖြစ်သည်။ Quantum computing သည် Big Data Analysis၊ Machine Learning စသည့် နယ်ပယ်များတွင် တိုးတက်မှုများ ဖြစ်ပေါ် စေနိုင်သည်ဟု မျှော်လင့်ရသည်။

နိဂုံး

၁၂။ Big Data သည် ၂၁ ရာစု၏ အဖိုးတန်ဆုံး သယံဇာတများထဲမှ တစ်ခုဖြစ်ပြီး နည်းပညာ လောကတွင် အရေးပါသော အခန်းကဏ္ဍမှ ပါဝင်နေသည်။ ၎င်းသည် စီးပွားရေး၊ လူမှုရေး၊ သိပ္ပံပညာရပ် နယ်ပယ်အသီးသီးတွင် တိုးတက်ပြောင်းလဲမှုများကို ဖြစ်ပေါ်စေနိုင်သော အလားအလာကောင်းများစွာ ရှိသော်လည်း ၎င်း၏ စိန်ခေါ်မှုများကိုလည်း ကျော်လွှားနိုင်ရန် လိုအပ်သည်။ Big Data ကို တာဝန်ယူမှုရှိရှိ၊ ကျင့်ဝတ်နှင့်အညီ လုံခြုံစိတ်ချရစွာ အသုံးချနိုင်ခြင်းဖြင့် လူသားတို့၏ ဘဝများကို ပိုမိုကောင်းမွန်စေမည့် အနာဂတ်ကို ဖန်တီးနိုင်မည် ဖြစ်ပါသည်။ Big Data နည်းပညာများ စဉ်ဆက် မပြတ် တိုးတက်ပြောင်းလဲနေမှုနှင့်အတူ ၎င်း၏ အသုံးချနယ်ပယ်များလည်း ပိုမိုကျယ်ပြန့်လာမည်ဟု ယုံကြည်ပါသည်။

Copyright Mar 20 2025, Myanmar Computer Emergency Response Team, mmCERT/cc