

## UOT 004.02

**Qasımova R.T.**

AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan  
[rena.gasimova@science.az](mailto:rena.gasimova@science.az)

### **BÖYÜK HƏCMLİ TİBBİ VERİLƏNLƏRİN ANALİTİKASI: MÖVCUD PROBLEMLƏR VƏ İMKANLAR**

*Səhiyyə sahəsində böyük verilənlər (BV) inkişaf edən perspektiv sahə olub, böyük həcmli verilənlər yığımının analizi və xərclərin azaldılması ilə nəticələrin yaxşılaşdırılmasını təmin etmək üçün istifadə edilir. Verilənlərin həcmnin artması və real zaman rejimində onların analizinə olan tələbat BV-nin analitikasının yaranmasına gətirib çıxarır. Məqalədə səhiyyə sahəsində yaranan BV-nin analitikasının mövcud problemləri araşdırılır, BV-nin əsas xüsusiyyətləri tədqiq olunur. BV-nin müxtəlif üstünlükləri, imkanları və xarakteristikaları müəyyənləşdirilir və bir sıra tövsiyələr verilir.*

**Açar sözlər:** verilənlər xəzinəsi, böyük verilənlər, böyük verilənlərin analitikası, biometrik verilənlər, genetikə, isbat edici tibb, genom analitikası, MapReduce, Hadoop.

#### **Giriş**

Bu gün İnternet texnologiyalarının növbəti inkişaf dövrü müşahidə olunur. Elə əşyalar yaranır ki, onlar ötürücülərin köməyi ilə bizim haqqımızda və bizi əhatə edən dünya haqqında informasiya toplayır. Gartner tədqiqat şirkətinin proqnozuna görə, 2020-ci ilə kimi İnternetə 30 milyard qurğu qoşulacaq, bunların da əksəriyyəti smartfon və kompüterlər kimi kommunikasiya vasitələri deyil, idman qolbaqları, İnternetə çıxışı olan soyuducular və digər əşyalar olacaqdır. Bu qurğular verilənləri toplayacaq və insanların qayğısına qalmaq üçün onlarla mübadilə edəcəklər [1].

Bu səbəbdəndir ki, son zamanlar BV (*ing. Big Data*) ideyası aktuallaşmış və bu sahəyə olan maraq daha da artmışdır. Hazırda Big Data termini böyük həcmli verilənlərlə işləyən bütün sferalarda: tibbdə, iqtisadiyyatda, bank fəaliyyətində, istehsalatda, marketinqdə, telekommunikasiyada, veb-analitikada və s. istifadə edilir. Yaxın illərdə tibbdə, kriminalistikada, şəhər infrastrukturunda və digər sferalarda Big Data inqilabı gözlənilir. Belə ki, tibb sferasında cari vəziyyətin təhlili göstərir ki, daşınan qurğular sahəsində fəal tədqiqat işləri gedir, bunların gələcəkdə kütləvi hal alacağı gözlənilir. Müxtəlif daşınan qurğular arasında təkcə saatları, eynəkləri, linzaları və bədəndə olan müxtəlif qurğu-plastırları deyil, həm də dəri altına yeridilən elektron mini-qurğuları görəcəyik. Mütəxəssislər bu istiqamətdə eksperimentlərin könüllülərin çatışmazlığı və həmçinin eksperimentlərin bahalığı üzündən çox yavaş templərlə getdiyini qeyd edirlər. Bu səbəbdən xroniki xəstəliyi olan və müxtəlif xəstəliklərə məruz qalan insanların sayı, təəssüf ki, getdikcə artır [2]. Məsələn, belə daxili ötürücülər diabetlə bağlı olan problemi həll edə bilirlər. Çünki insan real vaxt rejimində öz vəziyyətini izləyərək mövcud proqram vasitəsilə qanında şəkərin səviyyəsinə nəzarət edə bilər.

Belə ötürücülərin tətbiqi və inkişafı böyük həcmli verilənlər axını yaradır ki, bunları da nəinki sadəcə operativ emal etmək, həm də sonrakı analiz üçün saxlamaq lazımdır. Bundan başqa, bir sıra problemlər meydana gəlir ki, bunları da BV ilə iş zamanı həll etmək lazım gəlir. Birinci problem ondan ibarətdir ki, insanların fərdi məlumatları (verilənləri) və ümumiyyətlə, müştərilərdən daxil olan bütün informasiya emal olunmaq üçün şirkətlərə verilməlidir. Məsələ ondadır ki, bu zaman informasiya pulsuz təqdim olunur, analiz edilir, kompüterlə emal edilir, yeni məhsul yaradılır və sonra yenidən insanlara satılır. İnsanların fərdi məlumatlarından bu şəkildə istifadə etməyin hüquqi əsasları olmalıdır.

Tibbdə bu istiqamətdə işlər daha mürəkkəbdir. Məsələn, heç də bütün insanlar müayinədə öz genomlarını verməyə hazır deyillər. Ona görə ki, bunun nə olduğunu və nəticələrin necə ola biləcəyini yaxşı bilmirlər. Bununla yanaşı həkimləri pasiyentlərin xəstəliklərini sirr kimi saxlamağa borclu edən tibb etikası və həkim sirri mövcuddur. Beləliklə, şirkətlər konkret müştərilərin maraqlarına xidmət edən statistik verilənlərini, yaxşı olar ki, toplayıb saxlasınlar. Bütün

pasiyentlərin verilənlərinin analizinin informasiya metodlarının inkişafı gözlənilməyən nəticələr verə bilər. Qeyd etmək lazımdır ki, daşınan qurğuların spesifikasiyası çoxlu sayda, lakin strukturlaşmış xarakterli informasiya axınına nəzərdə tutur, bu da onun emalı prosesini asanlaşdırır. Bu informasiyaları istifadə etməklə şirkətlər pasiyentin vəziyyətinin yaxşılaşdırılmasına yönələn müəyyən işləri həyata keçirə bilərlər. Tibb sahəsində toplanan böyük həcmli verilənlərdən çox kiçik zamanda böyük sürətlə analiz aparmaqla bilik əldə etmək və analiz nəticələrini təqdim etmək aktual problemlərdən biridir. Bu mühüm faktı nəzərə alaraq, tibb sahəsində BV-nin analitikasının mövcud problemlərinin araşdırılması və tədqiq olunması olduqca zəruridir.

### Səhiyyədə “Big data” mənbələri

Səhiyyə sahəsi tarixən BV həcmi yaratmışdır. Bu gün verilənlərin böyük hissəsinin kağız şəklində saxlandığı bir vaxtda böyük həcmli verilənlərin rəqəmsallaşdırılması meylləri müşahidə olunur. Xərclərin azaldılması ilə yanaşı tibbi yardımın keyfiyyətinin yüksəldilməsi üçün BV geniş spektrli tibbi funksiyalara, o cümlədən qərar qəbulunun klinik dəstəklənməsinə, xəstəliklərin monitorinqinə, əhəlinin sağlamlığının idarə edilməsinə şərait yaradır [3, 4].

Səhiyyə sahəsində BV öz həcminə görə deyil, verilənlərin müxtəlif tip olmasına və onların emal sürətinə görə səciyyəvi xarakterə malikdir. BV-yə həkimin yazılı qeyd və təyinatları, tibbi vizuallaşdırma, laboratoriya, aptek, sığorta və digər inzibati verilənləri, pasiyentin elektron xəstəlik tarixçəsindəki verilənləri, vericilər tərəfindən generasiya olunan maşın verilənləri, sosial mediada, o cümlədən tvitlər, bloqlar, Facebook və digər platformalardan olan məlumatları, həmçinin veb-səhifələri və konkret pasiyentlər haqqında təxirəsalınmaz tibbi yardım haqqında olan informasiya, tibbi jurnallardakı yeniliklər və məqalələr və s. daxil edilir. Pasiyentin (xəstənin) sağlamlığı və rifahı ilə bağlı verilənlər birlikdə səhiyyə sahəsində BV-ni təşkil edir [5, 6].

Tibb elə sahələrdən biridir ki, burada verilənlər böyük həcmdə, çox böyük sürətlə və müxtəlif formatlarda məlumatlar, məsələn, rəqəm (analizlərin verilənləri, tibb kartlarında yazılar), video (məsələn, ultrasəs müayinəsi), foto (tomoqrafiya, rentgenoqrafiya), elektrosiqnalları qeyd edən aparatdan texniki siqnallar (elektroensefaloqramma, elektrokardiogramma) və s. toplanır. Bunları, sadəcə olaraq, bir araya yığmaq, birgə emal etmək və analiz üçün uyğun şəkllə salmaq çox çətin olur. Bu elektron tibb verilənlərini ənənəvi proqram təminatı ilə emal etmək çətin, hətta mümkün deyildir. Əsas problem ondan ibarətdir ki, məlumatların əksəriyyəti müxtəlif mənbələrdən müxtəlif formatlarda daxil olur və müxtəlif indeksləşmə sxemlərindən istifadə olunur. Onlar verilənlərin idarə edilməsinin ənənəvi alətləri və metodları ilə asan idarə olunmurlar [7].

Qeyd edildiyi kimi, səhiyyədə təyinatına görə BV çox böyük sürətlə müxtəlif mənbələrdən və müxtəlif formatlarda toplanan verilənlərə aiddir. Daxili mənbələrə, məsələn, elektron tibb yazılarını, klinik qərarların qəbulunun dəstəyi sistemini və s., xarici mənbələrə isə coğrafi cəhətdən paylanmış hökumət mənbələrini, laboratoriyaları, aptekləri, sığorta şirkətlərini və s. aid etmək olar. Səhiyyə sahəsində BV-nin mənbə və tiplərini aşağıdakı kimi təsnifatlandırırlar [8, 9]:

- *maşınlar vasitəsilə yaradılan verilənlər* (ing. *machine-to-machine, M2M*), buraya distant ötürücülərin, sayğacların və digər həyati mühüm qurğuların göstəriciləri daxildir;
- *biometrik verilənlər*, bunlara barmaq izi, genetika, gözün torlu qişasının skanlaşması, rentgenlər, digər tibbi təsvirlər, qan təzyiqi, nəbz göstəriciləri və digər bu kimi oxşar verilənlər aiddir;
- *insanın generasiya etdiyi verilənlər*, buraya həkimin qeydləri, elektron poçt və kağız sənədlər kimi strukturlaşdırılmamış və qismən strukturlaşdırılmış verilənlər aid edilir;
- *BV ilə razılaşmalar*, buraya əsasən səhiyyə sahəsində tələblər, qismən strukturlaşdırılmış və strukturlaşdırılmamış formatlarda əlçatan olan digər hesab yazıları aiddir;
- *veb və sosial media verilənləri*, buraya Facebook, Twitter, LinkedIn, bloqlar və onlara oxşar digər layihələrdə olan qarşılıqlı əlaqə verilənləri aiddir (həmçinin veb-saytların, smartfon tətbiqlərin tibbi xidmət planları da buraya daxil ola bilər).

Qeyd etmək lazımdır ki, artıq xüsusi bazalara, xəzinələrə toplanan verilənlər gəlir gətirə bilən kapitalla çevrilir. Formalaşan Big Data bazarını bir neçə sektora bölmək olar:

- Big Data-dan şəxsi işin analitikası və effektivliyinin yüksəldilməsi üçün istifadə edilməsi;
- Big Data-nın kənar agentlərə satılması;
- Big Data-nın müxtəlif agentlərə yenidən satılmaq üçün vasitəçilər tərəfindən toplanması;
- Öz Big Data-larına çıxış verən müştərilərə analitika üzrə xidmətlərin təqdimatı və s.

Səhiyyənin informasiyalaşdırılması pasiyentlərin elektron yazılarının və tibbi kartlarının, klinikanın informasiya sistemi ilə idarə edilməsinin, İnternet-servislərin tətbiqi ilə bağlıdır. Araşdırmalar göstərir ki, informasiya texnologiyalarının (İT-nin) inkişafı və tətbiqi üzrə lider olan ABŞ-da 15%-dən çox həkim klinikalarda tamamilə elektron yazılara keçib və kağız sənədləri doldurmur. Qalanları kağız blank və kartlardan qismən istifadə edirlər. Belə elektron yazıların, tibbi kartların informasiya sistemi ilə klinikanın idarə edilməsi və s. böyük həcmli verilənlər axını yaradır. BV axınını analiz etmək imkanına malik olmaq üçün bütün mümkün parametrlərin qeydiyyatını və bütün yazıları tamamilə elektron formaya keçirmək, həmçinin informasiyanın müxtəlif klinikalardan vahid verilənlər bazasına toplanmasını təmin etmək lazımdır. Bütün verilənlərin elektron daşıyıcılara köçürülməsi, rəqəmsallaşdırılması prosesi *datafikasiya* (ing. *datafication*) adlanır [10].

### **Səhiyyə sahəsində BV-nin analizi problemləri**

Səhiyyə sahəsində BV-nin analizi layihəsi üçün konseptual əsas ənənəvi tibbi informatikada olduğu kimidir. Əsas fərq emalın necə yerinə yetirilməsindədir. Adi səhiyyə analitikası layihəsində analiz fərdi kompüter və ya noutbuk kimi avtonom sistemlərdə qurulmuş biznes-analiz alətlərinin köməyi ilə yerinə yetirilə bilər. Burada BV təyinatına görə böyük həcmli olduğundan emal bölünür və bir neçə qovşaqlar üzrə yerinə yetirilir. Paylanmış emal konsepsiyası bir çox onilliklər ərzində mövcud olmuşdur. Yeni olan onun çox böyük həcmli verilənlər yığımının analizində istifadəsidir. Belə ki, bu gün tibb işçiləri sağlamlıqla bağlı əsaslandırılmış qərar qəbul etmək üçün BV xəzinəsindən istifadə edirlər. Verilənlər xəzinəsində (VX) verilənlərin emalı mərhələləri verilənlərin *toplanmasından, təmizləmədən, yükləmədən, analizdən* və nəhayət, analizin nəticələrinin *təqdimatından* ibarətdir. Bu mərhələlərin hər birində verilənlər üzərində xüsusi əməliyyatlar yerinə yetirilir [11]. Qeyd etmək lazımdır ki, əgər BV-nin analizi üçün VX texnologiyalarının tətbiqinə cəhd olsa, onda yalnız alqoritmlərin analizinə deyil, həm də verilənlərlə işin bütün mərhələlərinə diqqət yetirmək lazımdır.

Səhiyyədə yüzlərlə terabayt və ekzabayt həcmində BV-nin mövcud metodologiyalarla və ya alətlərlə toplanması, idarə edilməsi, saxlanması və onlardan faydalı informasiyanın əldə edilməsi ciddi problemdir. Həm strukturlaşdırılmış, həm də strukturlaşdırılmamış informasiya ilə işləmək, daha dərin intellektual analiz aparmaq və analizin nəticələrini vizuallaşdırılmaq BV-nin analitikasının əsas məsələlərindəndir. Verilənlərin həcmnin artması və real zamanda onların analizinə olan tələbat BV-nin ən əsas problemlərindən sayılan BV-nin analitikası (ing. *Big Data Analytics*) istiqamətinin yaranmasına gətirib çıxarmışdır. Strukturlaşdırılmamış verilənləri cədvəl və ya qrafikə gətirmək üçün Big Data analitikası üçün xüsusi İT-alətlər tətbiq etmək lazımdır (məsələn, R proqramlaşdırma dili və ya Hadoop proqram tətbiqi və s.) [12].

Big Data problemləri əsasən çox sürətlə böyüyən böyük həcmli informasiyanın real vaxt rejimində emalı, axtarışı, təsnifatlandırılması, analizi ilə bağlıdır. Tətbiq olunduğu sahələrdən asılı olmayaraq, BV-ni təsvir etmək üçün ümumi xarakteristikalar mövcuddur. Bu xarakteristikalar BV-nin əsas problemlərini özündə əks etdirməklə, üç əsas qrupa bölünür: *həcm* (ing. *volume*), *sürət* (ing. *velocity*) və *müxtəliflik* (ing. *variety*). İngilisdilli mənbələrdə bunu «3V» də adlandırırlar. Bu parametrlərin konvergensiyası BV-ni təyin etməyə və digər verilənlərdən fərqləndirməyə kömək edir [13, 14].

Səhiyyədə də BV-nin analitikası üç əsas xarakteristika ilə təsvir olunur. Sağlamlıqla bağlı verilənlər uzun müddət yaranır və fasiləsiz olaraq toplanaraq BV əmələ gətirir. Artıq səhiyyənin mövcud BV-nin həcmi şəxsi tibbi yazılardan, təsvirin radiologiyasından, klinik sınaqların

ötürücülərindən, insanın genetik və əhalinin genom ardıcılığı verilənlərindən və s. ibarətdir. BV-nin 3D vizuallaşma, ötürücülərin genom və biometrik göstəriciləri kimi yeni formaları da eksponensial artıma kömək edir. Bu gün verilənlərin idarə edilməsi sahəsindəki nailiyyətlər, xüsusən virtuelləşmə və bulud hesablamaları BV-nin daha effektiv toplanması, saxlanması və emalı üçün platformanın inkişafına kömək edir. Verilənlər real vaxtda yüksək sürətlə toplanır. BV-nin analizi üçün bu verilənlər birləşməlidirlər. Yeni verilənlərin daimi axını yeni problemlər yaradır. Toplanan və saxlanılan verilənlərin müxtəlifliyi dəyişsə də, onlar analiz, müqayisə və qərar qəbulu üçün vacibdir. Son zamanlar böyük həcmli tibbi verilənlərin müntəzəm monitorinqi keçirilir (məsələn, qlükozanın gündəlik diabetik ölçülməsi, arterial təzyiqin ölçülməsi və elektrokardiogramma diaqnostikası aparılır və s.). Nəticədə bir çox tibbi hallarda real vaxt rejimindəki daimi verilənlər (anesteziya üçün əməliyyat zalında monitorlar, ürək xəstəsinin yatağında monitorlar və s.) həyatı xilas edir. Real vaxt rejimində BV-nin analitikasını tibbin bütün ixtisasları üzrə yerinə yetirmək imkanı səhiyyə üçün ən aktual problemlərdəndir [15].

Sağlamlıq vəziyyəti haqqında verilənlərin xarakteri inkişaf etdiyinə görə həcm, sürətin və müxtəlifliyin də analitik metodları işlənilib. Yalnız elektron tibb kartlarında və digər strukturlaşdırılmış formatlarda toplanmış verilənlərin analizi aparılmır. Big Data mənbələrinin əsas hissəsini də strukturlaşdırılmamış və qismən strukturlaşdırılmış verilənlər təşkil edir. Həm strukturlaşdırılmış və qismən strukturlaşdırılmış, həm də strukturlaşdırılmamış informasiya ilə işləmək səhiyyə verilənlərini maraqlı və mürəkkəb məsələyə çevirir. Strukturlaşdırılmış verilənlər elə verilənlərdir ki, onları asan saxlamaq, analiz etmək və maşınla manipulyasiya etmək olar. Səhiyyə elə sahələrdəndir ki, burada strukturlaşdırılmış və qismən strukturlaşdırılmış verilənlər cihazların göstəricisindən və kağız sənədlərin elektron tibbi sənədlərə çevrilməsi nəticəsində alınan verilənlərdən ibarətdir. Tibbə, həmçinin strukturlaşdırılmamış verilənlər də generasiya olunur, bunlara tibb bacısı və həkimin əli ilə yazılmış tibbi sənədləri, kağız reseptləri, tomoqrafiya, rentgenoqrafiya, ultrasəs müayinəsi və digər təsvirləri aid etmək olar.

Artıq hazırda səhiyyə sahəsində strukturlaşdırılmış və strukturlaşdırılmamış verilənlərin yeni axınları fitnes-qurğulardan, genetikə və genomikadan, sosial tədqiqatlardan, kütləvi informasiya vasitələrindən və digər mənbələrdən daxil olur. Lakin bu verilənlərin az bir hissəsi hazırda toplanıb saxlanıla bilər. Verilənlər elə təşkil olunmalıdır ki, onları kompüterlərin köməyi ilə manipulyasiya etmək və faydalı informasiyanı məzmununa görə analiz etmək mümkün olsun. Tibbi təbiiqlər, strukturlaşdırılmamış verilənləri strukturlaşdırılmış verilənlərə çevirmənin avtomatlaşdırılması da daxil olmaqla, müxtəlif verilənlərin kombine edilməsi və çevrilməsinin daha effektiv üsullarına ehtiyac duyurlar. Səhiyyə sahəsində BV-nin potensialı ənənəvi verilənlərin yeni formalı verilənlərlə həm fərdi, həm də əhali səhiyyəsinə uzlaşmasındadır. Məsələn, əczaçılıq məhsulları istehsalçıları əhalinin klinik verilənlər yığımını genomiki verilənlərlə inteqrasiya edə bilsəydilər, onda istehsalçılar keçmişdəkinə nisbətən dərman müalicəsinin ən yaxşı keyfiyyətinə daha tez çıxış əldə edə bilərdilər.

Tədqiqatçılar səhiyyədə dördüncü xarakteristikanı da qeyd edirlər. Belə ki, tibb sahəsində toplanan məlumatların həqiqiliyi (*ing. veracity*) də önəmlidir. Yəni BV, analitika və nəticələr səhv deyil və etibara, inama layiq olan nəticələndir. Çünki əsasında qərar qəbul ediləcək verilənlərin nə qədər dəqiq və ya şübhəli olması sual doğurur. Əlbəttə, həqiqilik məqsəddir, amma hələlik reallıq deyildir. Məsələn, ötürücülər vasitəsilə alınan verilənlər daha etibarlıdır, nəinki sosial media verilənləri. Həqiqilik BV-nin tələblərinə uyğun olmaq üçün, eyni zamanda, platforma, alqoritm, metodika və alətlərin miqyaslılığını və məhsuldarlığını nəzərdə tutur. Strukturlaşdırılmış və strukturlaşdırılmamış BV-nin arxitekturu, analitikası və alətləri ənənəvi biznes-analitika (*ing. Business intelligence, BI*) alətlərindən çox fərqlənir. Məsələn, səhiyyə sahəsində BV-nin analizi bir neçə serverlərdə (“qovşaqlar”da) verilənlərin paylanmış emalı şəklində paralel hesablamalar paradigmasından istifadə etməklə yerinə yetirilir. Eyni zamanda, verilənlərin intellektual analizinin model və metodları, həmçinin statistik metodlar, vizuallaşdırma metodları, alqoritmləri BV-nin xarakteristikaları nəzərə alınmaqla aparılır. Səhiyyədə verilənlərin həqiqiliyi maliyyə verilənlərində olduğu kimi məsələlərlə (xüsusilə, ödəyici tərəfdən) qarşılaşır: pasiyentlərin sayı,

ödəyicisi, təminat kodu, pul vahidinin həqiqiliyi və s. digər həqiqilik parametrləri də səhiyyə üçün vacibdir. Bunlara diaqnozlar, müalicələr, reseptlər, prosedurlar, nəticələrin doğruluğu və s. aiddir.

Verilənlərin keyfiyyət məsələləri səhiyyə sahəsində iki səbəbə görə ciddi narahatlığa səbəb olur. Birincisi, insanların sağlamlığı dəqiq informasiyanın olmasından asılıdır, ikincisi isə tibbi verilənlərin, xüsusilə strukturlaşdırılmamış verilənlərin keyfiyyəti daim dəyişir və çox vaxt düzgün olmur (qeyri-dəqiq tərcümələr, reseptlərdə pis xətt ən tipik nümunələrdəndir) [16]. Səhvlərdən qaçmaqla və xərcləri azaltmaqla tibbi yardımın əlaqələndirilməsinin yaxşılaşdırılması yüksək keyfiyyətli verilənlərdən, dərman vasitələrinin təhlükəsizliyi sahəsindəki nailiyyətlərdən, diaqnostikanın effektivliyindən və dəqiqliyindən asılıdır. Lakin müxtəlifliyin və sürətin artması verilənlərin analizi və qərar qəbulundan əvvəl özünü təmizləmə qabiliyyətinə maneə törədir. Aparat və proqram təminatının dəyərinin azaldığı bir vaxtda bu məsələlər həll edilməlidir ki, səhiyyə sahəsində BV-nin analitikasının potensialından maksimum istifadə edilsin.

Buna görə də müxtəlif tip strukturlaşdırılmamış informasiyanın toplanması və saxlanması, axtarışı, təhlükəsizliyi, analizi kimi məsələlərin həlli üçün yeni alətin yaradılmasına hər zaman ehtiyac yaranır. Lakin arxitektur və platformaların sayı, eyni zamanda alətlərin mövcudluğunda açıq mənbəli kodların üstünlük təşkil etməsi və s. kimi digər məsələlər nəzərə alınmalıdır. Həmçinin aparat və proqram təminatının dəyərinin azaldığı bir vaxtda metodologiyaların işlənməsi və rahat interfeyslərin zəruriliyi məsələləri həll edilməlidir ki, səhiyyə sahəsində BV-nin analitikasının potensialından maksimum istifadə edilsin. Böyük həcmli tibbi verilənlər üzərində analiz aparmaqla verilənlərdə gizli qanunauyğunluqların və assosiasiyaların aşkarlanması mühüm qərarların qəbul edilməsində, insanların rifahının yaxşılaşdırılması, həyatlarının xilas edilməsində və digər məsələlərin həllində çox önəmlidir. Beləliklə, səhiyyə sahəsində BV çoxluğundan istifadə edərək daha əsaslandırılmış qərarlar qəbul edilir. Mütəxəssislər səhiyyə sahəsində BV-nin analitikasının potensialına aşağıdakıları daxil edirlər [17, 18]:

- klinik və iqtisadi cəhətdən effektiv müalicə metodlarını müəyyənləşdirmək üçün pasiyentin xarakteristikalarının analizi (müalicənin dəyəri və nəticələri əsasında analiz və alətlər təklif edərək istehlakçının davranışına təsir edirlər);
- həyat tərzi dəyişən pasiyentlərin aşkarlanması üçün qabaqcıl analitik vasitələrin (məsələn, seqmentləşdirmə və proqnoz modelləşdirilməsi) tətbiqi;
- profilaktikanın dəstəklənməsi üzrə proqnostik tədbirlərin aşkarlanması üçün xəstəliyin geniş miqyaslı profilaktikası;
- tibbi prosedurlar haqqında verilənlərin toplanması və nəşri (bununla da baxış və müalicə sxemi müəyyən edilən zaman pasiyentlərə kömək göstərilir);
- qabaqcıl analitik sistemlərin tətbiqi yolu ilə möhtəkirliyin aşkarlanması, proqnozlaşdırılması və minimallaşdırılması;
- real vaxt rejimində iddiaların həlli və s.

Səhiyyə sahəsində verilənlərin həcmnin yaxın illərdə kəskin artması gözlənilir. Bundan başqa, səhiyyənin xərclərinin ödənilməsi modeli dəyişilir. Konstruktiv istifadə və işin nəticələrinə görə mükafatlandırma müasir səhiyyədə mühüm amillər kimi meydana çıxır. Səhiyyə təşkilatı üçün mövcud alətləri, infrastrukturunu, BV-nin effektiv istifadə metodlarını əldə etməklə, həyatı mühüm amil olan gəlirləri artırmaq, xərcləri optimal idarə etmək, maliyyə göstəricilərini yaxşılaşdırmaq və artırmaq mümkündür.

### **Səhiyyədə BV-nin analitikasının imkanları**

Hazırda səhiyyə təşkilatları tərəfindən BV-nin rəqəmsallaşdırılması, birləşməsi və effektiv istifadəsi bir çox üstünlüklər əldə etməyə imkan verir. Potensial üstünlüklər xəstəlikləri daha asan və effektiv şəkildə müalicə etmək üçün onların vaxtında üzə çıxarılması daxil olmaqla, konkret fərdin, əhalinin sağlamlığını idarə etmək və səhiyyədə möhtəkirlik faktlarını daha tez, effektiv aşkarlamaqdan ibarətdir. Bu gün bir çox mürəkkəb məsələlər var ki, onları BV-nin analizi ilə həll etmək olar. Mütəxəssislər bunlara aşağıdakıları daxil edirlər [19, 20]:

– böyük həcmli arxiv verilənlərini qiymətləndirərək müəyyən hadisə və ya nəticələrin əvvəlcədən proqnozlaşdırılması (məsələn, xəstənin tibbi müəssisədə qalma müddəti);

– cərrahiyyəni seçəcək pasiyentlərin təyin edilməsi;

– cərrahiyyə müdaxiləsinin xeyri olmayan pasiyentlərin müəyyənlişməsi;

– ağırlaşmalar və tibbi ağırlaşma riskli və ya digər qeyri-müəyyən xəstəlik riskli pasiyentlərin üzə çıxarılması;

– xəstəliyin artması və xəstəlik halında inkişaf riskli pasiyentlərin, xəstəliyin inkişafına səbəb olan amillərin, mümkün patoloji halların aşkarlanması və s.

Həmçinin qeyd edilir ki, BV aşağıdakı üç sahədə xərcləri və səmərəliliyi artırmağa kömək edə bilər:

– **Klinik əməliyyatlarda**, pasiyentlərin müalicəsi və diaqnostikasının iqtisadi cəhətdən daha effektiv üsullarını müəyyən etmək üçün səmərəliliyin müqayisəli tədqiqi zamanı;

– **Tədqiqat və layihələndirmədə**, buraya elmi məqsədli qurğuların istehsal xərcini azaltmaq üçün proqnoz modelləşdirilməsi aiddir. Klinik sınaqların keçirilməsini yaxşılaşdırmaq üçün statistik alətlər və alqoritmlərin tətbiqi, eyni zamanda yaxşı müalicə üçün xüsusi pasiyentlərin toplanması, yeni müalicə metodlarının bazara sürətlə çıxarılmasına şərait yaradır. Nəticədə pasiyentlərə aid klinik sınaqların və yazıların analizi bu göstəricilərin müəyyən edilməsi və məhsul bazara çatmazdan əvvəl mənfi nəticələrin aşkarlanması üçün lazım olur;

– **İctimai səhiyyədə**, buraya ictimai səhiyyə sahəsində nəzarətin təkmilləşdirilməsi məqsədilə xəstəliklərin modellərinin analizi və xəstəliklərin yayılmasının izlənməsi, dəqiq məqsədli vaksinlərin daha sürətli inkişafı, böyük həcmli verilənlərin faydalı informasiyaya çevrilməsi və s. aiddir. Bu informasiyadan tələbatların aşkarlanması, xidmətlərin göstərilməsi, həmçinin böhran vəziyyətinin proqnozlaşdırılması və qarşısının alınması, xüsusən əhəlinin maraqları üçün istifadə oluna bilər. Bundan başqa, mütəxəssislər səhiyyə sahəsində BV-nin analitikasının aşağıdakı sahələrə kömək edəcəyini göstərir [21]:

– İsbatedici tibb, müxtəlif strukturlaşdırılmış və strukturlaşdırılmamış, maliyyə, operativ, klinik verilənləri, həmçinin genom verilənlərini analiz etmək, müalicəni nəticələrlə uyğunlaşdırmaq, xəstəlik riski olan pasiyentləri əvvəlcədən proqnozlaşdırmaq və daha effektiv xidməti təmin etmək üçündür. İsbatedici tibb fundamental elmi verilənlərə əsaslanan tibbi biliklərin təşkili konsepsiyasıdır. Bu halda şəxsi təcrübə, həmkarların nüfuzu ikinci dərəcəli əhəmiyyətə malikdir. Burada tibbi və idarəetmə qərarlarının qəbulunda başlıca prinsip yalnız obyektiv faktın əsas götürülməsidir. Ədəbiyyatda tibb elminin bu bölməsinin digər adları mövcuddur: faktlara əsaslanan tibb (ing. *fact-based medicine, FBM*), isbatlara əsaslanan təcrübə (ing. *evidence-based decision making, EBDM*), faktlara əsaslanan qərar qəbulu (ing. *fact-based decision making, FBDM*), isbatedici səhiyyə (ing. *evidence-based health care, EBHC*). Tibbi təcrübəyə isbatedici tibbin tətbiqinin əsas ideyası həkimin fəaliyyətinə insan amilinin təsirini minimallaşdırmaqdır [22];

– **Genom analitikası**, genin sekvenləşməsini effektiv və qənaətlə yerinə yetirmək, genom analizini tibbi yardımın müntəzəm prosesinin bir hissəsi etməkdir;

– Möhtəkirliyin analizi, möhtəkirlik, israf və sui-istifadə etmə hallarının azaldılması üzrə çoxsaylı iddia sorgularının sürətlə analiz edilməsidir;

– Məsafədən monitorinq, *real vaxt* rejimində sürətlə yerdəyişən BV-nin təhlükəsizliyinin monitorinqi və eyni zamanda, hadisələrin mənfi proqnozu üçün toplanması və analizidir;

– Pasiyent profilinin analitikası, pasiyentlərin profili üçün qabaqcıl analitik vasitələrin tətbiqi nəzərdə tutulur (məsələn, seqmentləşdirmə və proqnoz modelləşdirilməsi). Yəni həyat tərzi dəyişən pasiyentlərin effektiv şəkildə aşkarlanması, həmçinin konkret xəstəliyin inkişaf riski olan pasiyentlərə (məsələn, şəkərli diabet) profilaktiki kömək üçündür.

Eyni zamanda, səhiyyə sahəsi üçün BV-nin analitikasının aşağıdakı üstünlükləri qeyd edilir [23]:

– mənfi nəticəli riski olan pasiyentlərin dəqiq təyini;

- pasiyentlərə əsaslandırılmış qərarların qəbulu və sağlamlılıqlarının daha effektiv idarə edilməsi, həmçinin daha sağlam həyat tərzini sürmək üçün lazımı informasiyanın verilməsi;
- yaxşı fayda verməyən və ya çox baha olan prosedurların, proqramların və proseslərin müəyyən edilməsi;
- təkrar hospitallaşmanın və uyğun olaraq nizamlanan müalicə planlarının riski artıran və ya əlavə effektlər yaradan ətraf mühit və həyat tərzini amillərinin aşkarlanması yolu ilə azaldılması;
- nəticələrin evdəki monitorlarla həyati mühüm orqanların öyrənilməsi yolu ilə yaxşılaşdırılması;
- əhəlinin sağlamlığının pasiyentlərin xəstəliklərinin yayılması və ya təbii fəlakətlər zamanı zəifliklərinin aşkarlanması yolu ilə idarə edilməsi (nəticədə klinik, maliyyə və əməliyyat verilənləri bir yerdə olur ki, bu da resurslardan məhsuldar və real vaxt rejimində istifadəyə imkan verir) və s.

Məlumdur ki, tibb sahəsində qərarlar pasiyentlərin şəxsi verilənlərinin toplanması və saxlanması tələb edir. Belə BV bazası artıq inkişaf etmiş dövlətlərdə yaradılır (məsələn, ABŞ-da ştatlar səviyyəsində). Orada olan informasiyanı yalnız peşəkar həkimlər, orta tibbi personal, əczaçılar, sığorta şirkətlərinin əməkdaşları, məhkəmə orqanları, analitiklər əldə edilə bilirlər. Qeyd etmək lazımdır ki, BV xəzinədə düzgün şəkildə saxlanmalıdır ki, onları analitika və statistika üçün istifadə etmək mümkün olsun. Eyni zamanda onların üçüncü ələ ötürülməsi halında verilənlərin fərdi məxsusluğu kodlaşdırılmalıdır. Mütəxəssislər tibb sahəsində baş verən dəyişiklikləri müqayisə edərkən haqlı olaraq tibbi statistik göstəricilərə əsaslanan tibb (ing.Data Driven Medicine) adlandırırlar. Eyni zamanda bu tibb Big Datanın “ekosistemi” termininin meydana gəlməsinə gətirib çıxardı. Bu zaman nəzərə almaq lazımdır ki, Big Data ekosistemi solumla qarşılıqlı əlaqədədir və ona təsir edir. Big Data ekosistemi genişlənir, ona müxtəlif həyat sahələrindən vətəndaş haqqında verilənlər daxil olur. Bu məlumatlara kredit kartın bütün əməliyyatlarını, fərdi kompüterdə toplanan log-fayl verilənlərini, mobil telefonun yerini təyin edən verilənləri, təqaüd fondu və ya sığorta şirkətinin verilənlərini və çoxlu digər verilənlər bazalarını aid etmək olar.

Tədqiqatlardan məlum olur ki, artıq bu gün BV axınlarını birləşdirmək cəhdləri edilir, bu da iqtisadiyyatın ayrı-ayrı sektorları üçün real fayda və qənaət gətirir. İndi elmi tədqiqatlar nəzəri düşüncələrdən daha çox elmi eksperimentlərə köklənmişdir. Hər bir vətəndaş gələcəkdə müxtəlif BV bazalarında öz verilənlərinin kodlaşdırılması və dekodlaşdırılması prosesləri üçün identifikasiya nömrəsi alacaqdır. Genişlənən BV ekosistemi peşəkar birliyin konsolidasiyasını tələb edir, belə ki, verilənlərin saxlanması formatları interoperabel mühitini təmin etməlidir. Unifikasiyalı metod üzrə kodlaşdırılmış müxtəlif bazalardan müxtəlif verilənlər vahid analiz üçün uyğun olmalıdırlar.

Mütəxəssislərin fikrincə, Big Data ideologiyasının tətbiqi üzrə yaxın zamanlarda görülən işlər arasında aşağıdakıları göstərmək olar [24, 25]:

- kağızdan tam imtina etməklə bütün yazıların elektron daşıyıcılara köçürülməsinə cəhd;
- tibbi informasiya sistemlərində toplanan BV-nin analizi vasitəsilə müalicə prosesinin optimallaşdırılması;
- poliklinikalara təhkim olunmuş əhali kontingenti üzrə informasiyanın yerli bazarların lazımı dərmanlarla zəmanətli təmin edilməsi məqsədilə dərman istehsalçılarına təqdim edilməsi və s.

Pasiyentlərin hüquqları yalnız həyatın və müalicənin keyfiyyətinin yaxşılaşdırılması üçün Big Data-nın istifadə imkanları ilə deyil, həm də konfidensiallığı və bazarda Big Data dövrünün hüquqi tənzimlənməsi ilə əlaqədardır. Həmçinin istənilən vətəndaş ödənişsiz olaraq öz şəxsi verilənlərinə giriş əldə etmək, onları nüsxələmək hüququna malik olmalıdır. Hazırda dövlətlərdə informasiyanın yaradıcısı, sahibi və istehlakçısı kimi iştirak edən bütün kateqoriya şəxs və təşkilatların qarşılıqlı münasibətlərini tənzimləyən normativ-hüquqi sənədlər işlənir. Big Datanın əsas əhəmiyyəti həm retrospektiv, həm də real vaxt rejimində verilənlərə proqnozlar tərtib etmək imkanındır. Bunun üçün ölkənin müvafiq mərkəzi icra hakimiyyəti orqanı, tibb mütəxəssislərinin rəhbərləri, tibbi informasiya sistemlərini (TİS) işləyib hazırlayanlar və analitiklər arasında əlaqə yaratmağa imkan verən şərait olmalıdır. Mütəxəssislər onlara aşağıdakıları aid edirlər [26]:



- həkimin ilkin yazılarının tərkibini genişləndirməyə və saxlanma strukturunu modifikasiya etməyə imkan verən TİS-in istifadəsi;
- pasiyentin cari vəziyyətinin göstəricilərini müəyyənləşdirən və onları xəstənin mütəxəssis tərəfindən qəbul vaxtından kənar tibb təşkilatlarının verilənlər bazasına daxil edən metodların (texnologiyaların) inkişafı;
- verilənlərin emalının regional mərkəzlərinin (VEM) hesablama güclərinin TİS-in ilkin yazılarının emalı və ümumiləşdirilməsi alətlərinin effektivliyinin yüksəldilməsilə yanaşı formalaşdırılması və daim artırılması;
- səhiyyə sahəsində Big Data ilə iş metodlarının inkişafı məsələlərinin prioritet kimi müəyyən edilməsi və s.

### **Səhiyyədə BV-nin emalı və analizi üçün Hadoop texnologiyası**

Mütəxəssislər bilavasitə İT-dən Big Data üçün texnologiyalar kateqoriyasına verilənlərin kütləvi paralel emalı (ing. *Massively Parallel Processing, MPP*) platformalarında analitik sistemləri, verilənlərin bulud servisini, MapReduce texnologiyalarını və NoSQL tipli paylanmış verilənlər bazasının idarəetmə sistemlərini (VBİS) və s. daxil edirlər. Apache Software Foundation-un layihəsi olan Hadoop daha geniş yayılmış texnologiyadır, paylanmış hesablama mühitində BV-nin (petabayt miqyasında) emalı və analizi üçün əsas platformadır, MapReduce modelinin açıq kodlu (ing. *open access*) sistemidir. Hadoop iki əsas komponentdən təşkil olunmuşdur: Hadoop MapReduce və Hadoop Distributed File System (HDFS). Burada MapReduce paralel hesablamalara, HDFS paylanmış fayl sistemi isə verilənlərin idarə edilməsinə əsaslanır [27–29].

Hazırda Big Data emalı üçün açıq kodlu Hadoop hər bir istifadəçiyə əlçatandır və bu sistem pulsuz yayılır. Hadoop, MapReduce kimi platformaların açıq kodu səhiyyə sahəsində BV-nin analitikasının tətbiqi üçün də şərait yaratmışdır. Səhiyyənin ənənəvi analitika alətləri çox rahat və şəffafdırlar. Analitikanın ənənəvi alətlərinin alqoritmi və modelləri oxşar olduqları halda BV-nin istifadəçi interfeysləri tamamilə müxtəlifdir. Lakin digər tərəfdən BV-nin analizi alətləri, proqramlaşdırma çox mürəkkəbdir və müxtəlif vərdişlərin tətbiqini tələb edir. Hadoop verilənlərin və analitika alətlərinin ikili təşkilatçılıq rolunu oynaya bilər. O, böyük potensial təklif edir və müəssisələrə indiyə qədər idarə və analiz edilməsi çətin olan verilənləri istifadə etməyə imkan verir. Hadoop müxtəlif strukturlaşdırılmış və ümumiyyətlə, strukturlaşdırılmamış BV-ni emal etməyə imkan verir. Lakin Hadoop sistemini qurmaq, sazlamaq və inzibatçılıq etmək mürəkkəb ola bilər, həm də bu sistemlə işləməyi bacaran istifadəçiləri tapmaq asan deyildir [30].

Eyni zamanda bu platformalar açıq kodlu sistemlər olduğu üçün onlardan istifadə rahat deyil. Yəqin ki, bu səbəblərə görə təşkilatlar Hadoop-un tam tətbiqinə hazır deyillər. Əlavə platforma və alətlərin ekosistemi Hadoop paylanmış platformasını dəstəkləyir. Mənfi cəhəti texniki dəstəyin və minimal təhlükəsizliyin olmamasıdır. Səhiyyə sahəsində bu, əlbəttə, mühüm çatışmazlıqdır. Bundan başqa, qeyd edildiyi kimi, bu platformalar, alətlər son istifadəçilərdən səhiyyə sahəsində böyük proqramlaşdırma bacarığına malik tipik olmayan vərdişlər tələb edir. Eyni zamanda, səhiyyə sahəsində BV-nin yeni olduğunu nəzərə alaraq, idarəetmə, mülkiyyət hüquqları, konfidensiallıq, təhlükəsizlik və standartlar daxil olmaqla, bir çox məsələlər həll edilməlidir. Qabaqcıl metod və mövcud analitik texnologiyalar pasiyentin sağlamlığı ilə bağlı çoxlu sayda mövcud tibbi verilənlərə tətbiq edilməklə daha yaxşı nəticələr alınabilir. Bu nəticələr də gələcəkdə əhalinin sağlamlığının qorunması məqsədi ilə tətbiq oluna bilər. Fərdi və qrup verilənlər hər həkimə konkret pasiyent üçün daha münasib müalicə variantı müəyyən etməyə və qərar qəbulu prosesində istifadə etməyə kömək edə bilər.

### **Bioinformatikadan fərdiləşdirilmiş tibb**

Molekulyar genetikanın ayrıca bölməsi olan genomika fenomenal sürətlə inkişaf edir. Bu gün bir çox firmalar genomun intellektual analizi metodları üzərində təhlillər aparırlar. Onların fikrincə, bu yaxın beş il ərzində fərdiləşdirilmiş tibb adı hala çevriləcək və geniş yayılacaqdır.



Bütün bu nailiyyətlər BV-nin emalı metodlarının inkişafı sayəsində mümkün oldu. Müxtəlif tibbi verilənlərin avtomatlaşdırılmış emalı metodlarının tədqiqində məhz Big Data texnologiyaları yeni nəsil texnologiya, arxitektura olmaqla, BV-dən çox kiçik zamanda böyük sürətlə analiz aparmaqla bilik əldə etmək və analizin nəticələrini təqdim etməyə imkan verir. Big Data texnologiyalarında istifadə olunan xüsusişdirilmiş alqoritmlər hesablama yükünü on dəfələrlə azaldır, yeni veb-platformalar isə genetik verilənlər petabaytına birgə çıxışı mümkün edir.

Fərdiləşdirilmiş tibbin inkişafına yanaşmalar indi çoxdur, lakin son nəticədə onlar konkret pasiyentin genetik verilənlərinə əsaslanaraq həkimlərə müalicə kursu və xəstəliklərin profilaktikasını məsləhət görməyə imkan verir. Fərdiləşdirilmiş tibbdə əsas problem elmi-praktiki məqsədlərlə mümkün dərəcədə çox sayda insanların genetik verilənlərinin müqayisəsi üçün proqram təminatının işlənməsidir. Proqram təminatının reallaşması üçün müntəzəm tibbi baxışdan keçən pasiyentlərin verilənlərinin toplanması əsas şərtidir. Bu insanların arasında həm müxtəlif xəstəliklərə tutulmuş pasiyentlər, həm də könüllülərdən ibarət sağlam şəxslər ola bilər. Milyonlarla yazıların müqayisəli analizi patoloji proseslərin inkişafının detallarını və onların formalaşmasında genetik pozulmaların rolunu aşkar etməyə kömək edə bilər. Korrelyasiya modelləri həm müxtəlif genom yazılarının müqayisəsi, həm də onların pasiyentin sağlamlığının vəziyyəti haqqında əhəmiyyətli verilənlərlə birlikdə geniş analizi üçün istifadə oluna bilər. Mütəxəssislər hesab edirlər ki, hər bir insanın xarakteristikasına genetik verilənlər əlavə etməklə onun müxtəlif vəziyyətlərdə davranışını uğurla proqnozlaşdırmaq və xəstələndikdə ona daha effektiv tibbi yardım göstərmək olar.

Aydındır ki, bu metod ictimai səhiyyənin məsələlərini həll etmək üçün də yararlıdır. Tibbi genetik və BV-nin analizi xərçəngin müxtəlif formaları və insulin asılılıqlı şəkər diabeti və s. kimi xəstəliklərin müalicəsində mühüm rol oynayacaqdır. İlk dəfə insan genomunun praktiki məqsədlərlə tam sekvenləşməsi ABŞ-ın Arizona ştatında Genomika Tədqiqatları İnstitutunda (*ing. Translational Genomics Research Institute*) yerinə yetirilmişdir [31]. Belə imkan yaranana qədər mürəkkəb klinik hallarda müəyyən mutasiyaların axtarışı ilə kifayətləndirilirdi. Genomun deşifrənməsi prosedurunun dəyəri azaldıqca, metod daha əlçatan olur, bioinformatika isə səhiyyəyə daha dərin nüfuz edir. Əvvəllər insan geninin dekodlaşmasının emalı 10 il vaxt aparırdısa, indi bu bir həftə ərzində əldə oluna bilər.

Fərdiləşdirilmiş tibbin ilk tərəfdaşlarından genetik verilənlərin bulud servisi vasitəsilə birgə istifadəsini irəli sürən ABŞ-ın Kaliforniya ştatında yerləşən DNAnexus şirkətini və genom tədqiqatları İnstitutunun banisi, genomun sekvensiyası sahəsində ilk pioner Kreyq Venterini göstərmək olar. O, həmçinin Human Longevity biotexnoloji şirkətinin yaradıcısıdır. DNAnexus şirkətinin əməkdaşları belə onlayn-servisləri genomikanın sonrakı inkişafı üçün əsas və fərdiləşdirilmiş tibbin ehtiyacları üçün daha effektiv adaptasiya üsulu hesab edirlər [32]. Tədqiqatçılar, həmçinin DNAnexusun bulud servisini tətbiq edərkən verilənlərə çıxış hüququnu etibarlı bölüşdürür və universitetlər, klinik laboratoriyalar, xəstəxanalar və statistik mərkəzlər kimi üçüncü şəxslərlə təhlükəsiz əməkdaşlıq edirlər. O, tamamilə həkim sirri və fərdi verilənlərin qorunması hüququnu təmin edən HIPAA (*ing. Health Insurance Portability and Accountability Act*) – ABŞ-ın qanunvericilik aktının tələblərini ödəyir [33–35].

## Nəticə

BV-nin analitikasının potensialı böyükdür. Tibb işçiləri mürəkkəb texnologiyalar vasitəsilə klinik və digər verilənlər xəzinəsindən istifadə edərək əsaslandırılmış qərar qəbul edirlər. Gələcəkdə BV-nin analitikasının səhiyyə sisteminin bütün sahələrində sürətli, geniş tətbiqi və istifadəsi gözlənilir. Bu məqsədlə bu sahədə mövcud olan bir sıra problemlər həll olunmalıdır. BV-nin analizində bir çox məsələlər, məsələn, şəxsi həyatın toxunulmazlığı, təhlükəsizliyin təminatı, standartların və idarəetmə sistemlərinin qurulması, həmçinin fasiləsiz təkmilləşən alət və texnologiyaların istifadəsi daim diqqəti özünə cəlb edir. Hazırda səhiyyə sahəsində BV-nin analizi və tətbiqləri ilkin inkişaf mərhələsindədirlər, lakin platforma və alətlərin sürətli tərəqqisi təkmilləşmə prosesini tezləşdirə bilər.

Səhiyyə sahəsində analitik platformalar verilənlərin emalı üçün əsas funksiyaları dəstəkləməlidirlər. Platformanın qiymətləndirilməsi meyarlarına əlçatanlıq, fasiləsizlik, istifadənin sadəliyi, əhatəliliyi, müxtəlif detallaşdırma səviyyələrində manipulyasiya qabiliyyəti, konfidensiallıq və təhlükəsizlik, keyfiyyət təminatı aiddir. Hazırda tətbiq edilən açıq mənbəli platformaların əksəriyyətinin üstünlükləri və çatışmazlıqları vardır. Səhiyyə sahəsində BV-nin analizi elə təşkil edilməlidir ki, idarəetmə rahat və şəffaf olsun. Real vaxt rejimində BV-nin analizi səhiyyə sahəsində əsas tələb sayılır. Yəni verilənlərin toplanması ilə emalı arasındakı problem həll edilməlidir. Verilənlərin fasiləsiz toplanması, təmizlənməsi məsələlərinə baxılmalı, eyni zamanda idarəetmə və standartlar kimi mühüm idarəetmə məsələləri də nəzərə alınmalıdır. Səhiyyə sahəsindəki verilənlər nadir hallarda standartlaşdırılmış olur, bir çox hallarda fraqmentləşdirilir və köhnəlmiş İT-sistemlərində generasiya edilir. Bu böyük problemi həll etmək lazımdır.

## Ədəbiyyat

1. Laney D. 3D Data Management: Controlling Data Volume, Velocity and Variety, Technical report, META Group, Inc (now Gartner, Inc.), February 2001. <http://blogs.gartner.com>
2. Laura B. Madsen Data-Driven Healthcare: How Analytics and BI are Transforming the Industry, Publisher: John Wiley & Sons, Inc., 2014, 224 p.
3. Wullianallur R. Data Mining in Health Care. Healthcare Informatics: Improving Efficiency and Productivity, CRC Press, 2010, Taylor & Francis, pp.211–224. [www.crcnetbase.com](http://www.crcnetbase.com).
4. Wullianallur R., Viju R. Big data analytics in healthcare: promise and potential, Raghupathi and Raghupathi; licensee BioMed Central Ltd. Health Information Science and Systems, 2014, vol.2, no.3, pp.2–10. online resource, [www.hissjournal.com](http://www.hissjournal.com).
5. Билл Ф. Укрощение больших данных. Как извлекать знания из массивов информации с помощью глубокой аналитики, пер. с англ. Андрея Баранова, М.: Манн, Иванов и Фербер, 2014, 352 с.
6. Bian J., Topaloglu U., Yu F. Towards Large-scale Twitter Mining for Drug-related Adverse Events / Proceedings of the 2012 international workshop on Smart health and wellbeing (SHB'12), New York, USA, 2012, pp.25–32.
7. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим, пер. с англ. Инны Гайдюк, М.: Манн, Иванов и Фербер, 2013, 240 с.
8. Gudivada V.N., Rao D., Raghavan V.V. NoSQL Systems for Big Data Management / Proceedings of the 2014 IEEE World Congress on Services (SERVICES '14), USA, 2014, pp.190–197.
9. Institute for Health Technology Transformation (IHTT): Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry, 2013. [http://c4fd63cb482ce6861463-bc6183f1c18e748a49b87a25911a0555.r93.cf2.rackcdn.com/iHT2\\_BigData\\_2013.pdf](http://c4fd63cb482ce6861463-bc6183f1c18e748a49b87a25911a0555.r93.cf2.rackcdn.com/iHT2_BigData_2013.pdf)
10. Ayankoya K., Calitz A., Greyling J. Intrinsic Relations between Data Science, Big Data, Business Analytics and Datafication / Proceedings of the Southern African Institute for Computer Scientist and Information Technologists Annual Conference (SAICSIT 2014), New York, USA, 2014, pp.192.
11. Qasimova R.T. Milli domen adları ilə bağlı biliklər bazasının yaradılmasının konseptual əsasları haqqında // Bakı Universitetinin Xəbərləri, Fizika-Riyaziyyat Elmləri Seriyası, № 4, 2010, s.95–102.
12. Qasimova R.T. Big data analitikası: mövcud yanaşmalar, problemlər və həllər // İnformasiya Texnologiyaları Problemləri, 2016, №1, s.75–93.
13. Clifford L. Big data: How do your data grow? // Nature, 2008, vol.455, pp.28–29.
14. Əliquliyev R.M., Hacırahimova M.Ş. Big data fenomeni: problemlər və imkanlar // İnformasiya Texnologiyaları Problemləri, 2014, №2, s.3–16.

15. Andreas H. Biomedical Informatics 2014: Discovering Knowledge in Big Data, Publisher: Springer International Publishing AG., 1st. Edition, 2014, 606 p.
16. Babu S., Herodotou H. Massively Parallel Databases and MapReduce Systems, Foundations and Trends in Databases, 2013, vol.5, no.1, pp.1–104,.
17. Deng X., Donghui W. Big data and predictive modeling topics in healthcare / Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '15), New York, USA, 2015, p.677.
18. Intel: Big Data Analytics, 2012, [www.intel.com/content/dam/www/public/us/en/documents/reports/data-insights-peer-research-report.pdf](http://www.intel.com/content/dam/www/public/us/en/documents/reports/data-insights-peer-research-report.pdf)
19. Jimeng S., Chandan K. Big data analytics for healthcare / Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '13), New York, USA, 2013, pp.15–25.
20. Varun C., Sukumar Sreenivas R., Schryver Jack C. Knowledge discovery from massive healthcare claims data / Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '13), New York, USA, 2013, pp.1312–1320.
21. Davenport T.H., Dwight McNeill. Analytics in Healthcare and the Life Sciences: Strategies, Implementation Methods, and Best Practices, Publisher: Pearson Education, USA, 2013, 352 p.
22. Nambiar U., Niranjana T. Data management & analytics for healthcare (DARE 2013) / Proceedings of the 22nd ACM international conference on Information & Knowledge Management (CIKM '13) New York, USA, 2013, pp.2565–2566.
23. Королюк И.П. Медицинская информатика. Учебник, 2 изд., перераб. и доп., Самара: ООО «Офорт», ГБОУ ВПО «СамГМУ», 2012, 244 с.  
[www.samsmu.ru/files/smu/chairs/radiology/med\\_inf.pdf](http://www.samsmu.ru/files/smu/chairs/radiology/med_inf.pdf)
24. Yamakami T. Inter-service revisit analysis of three user groups using intra-day behavior in the mobile clickstream / Proceedings of the 2009 International Conference on Hybrid Information Technology (ICHIT '09), New York, USA, 2009, pp.340–344.
25. Колесниченко О.Ю., Смородин Г.Н. Большие данные: социальные вызовы / Тезисы V социологической Грушинской конференции «Большая социология: расширение пространства данных», Материалы конференции, М: ВЦИОМ, 2015, с.26–29.
26. Schmarzo B. Big Data MBA: Driving Business Strategies with Data Science, Publisher: John Wiley & Sons, Inc. 1st. Edition, 2015, 312 p.
27. Hadoop Distributed File System. <http://hadoop.apache.org/docs>
28. Vignesh P. Big Data Analytics with R and Hadoop, Publisher: Packt Publishing Ltd, 2013, pp.238.
29. Чак Л. Hadoop в действии, Издательство: ДМК Пресс, 2012, 424 с.
30. Ohlhorst Frank J. Big Data Analytics: Turning Big Data into Big Money, Publisher: John Wiley & Sons Inc, 2013, 176 p.
31. Translational Genomics Research Institute, Arizona, USA, [www.tgen.org/](http://www.tgen.org/)
32. DNAnexus, Providing cloud solutions for the global genomics industry, USA. [www.dnanexus.com](http://www.dnanexus.com)
33. Human Longevity, Inc., San Diego, California, USA. [www.humanlongevity.com/about/j-craig-venter](http://www.humanlongevity.com/about/j-craig-venter)
34. Department of Health & Human Services (HHS) USA. [www.hhs.gov/about/index.html](http://www.hhs.gov/about/index.html)
35. Assunção M.D., Rodrigo N., Bianchi S., Netto Marco A.S., Rajkumar B. Big Data computing and clouds // Journal of Parallel and Distributed Computing, 2015, vol.79, pp.3–15.

**УДК: 004.89**

**Касумова Рена Т.**

Институт Информационных Технологий НАНА, Баку, Азербайджан

[rena.gasimova@science.az](mailto:rena.gasimova@science.az)

**Аналитика больших объемов медицинских данных: текущие проблемы и перспективы**

Увеличение объема данных и потребность их анализа в реальном времени привели к необходимости создания одной из основных проблем аналитики больших данных. В статье рассматриваются актуальные проблемы большого анализа данных, созданных в области здравоохранения, исследуются основные характеристики данных. В то же время определяются различные возможности, преимущество и характеристики данных в области здравоохранения и дан ряд рекомендаций.

**Ключевые слова:** *хранилища данных, Big Data, Big Data analytics, биометрические данные, генетика, доказательная медицина, аналитика генома, MapReduce, Hadoop.*

**Rena T. Gasimova**

Institute of Information Technology of ANAS, Baku, Azerbaijan

[rena.gasimova@science.az](mailto:rena.gasimova@science.az)

**Analysis big volume medical data: current problems and prospects**

Big Data is becoming a promising area in the field of healthcare. It is used to improve the results of the analysis of large volumes of data sets and to reduce the costs. Increase in data volume and demand for ad hoc analysis of data created one of the biggest problems of Big Data called Big Data analysis. The article deals with actual problems of a large analysis of data generated in the field of health, and explores the main characteristics of the data. At the same time, it defines the various opportunities, advantages and characteristics of the date in the area of health, and provides a number of recommendations.

**Keyword:** *data warehouse, big data, big data analytics, biometric data, evidence-based medicine, genomic analytics, MapReduce, Hadoop.*