

## Akıllı şehirlerde büyük coğrafi veri yönetimi ve analizi: hava kalitesi örneği

Arif Çağdaş Aydınoglu<sup>\*1</sup>, Rabia Bovkır<sup>1</sup>, Muzafer Bulut<sup>2</sup>

<sup>1</sup>Gebze Teknik Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği, Kocaeli, Türkiye

<sup>2</sup>Nil Proje Müh. İnş. Tur. Mad. Dış Tic. Ltd. Şti., Ankara, Türkiye

### Anahtar Kelimeler

Hava Kalitesi İndeksi (HKİ)  
NoSQL  
Büyük Coğrafi Veri Yönetimi  
MongoDB  
Büyük Veri

### ÖZ

Bilişim teknolojilerinin gelişmesiyle, veri üretim teknikleri ve toplanan veri hacmi artmıştır. Akıllı şehir uygulamaları ile sensörler, IoT, internet, giyilebilir teknolojiler gibi farklı veri kaynaklarından akan verilerin yönetimi ve bu verilerden değer yaratmak mümkün hale gelmiştir. Günümüzde toplanan büyük hacimli ve karmaşık verinin yönetimi için geleneksel veri depolama ve yönetim yaklaşımları yetersiz kalmış ve büyük verinin hacim, hız ve çeşitlilik gibi karakteristik özellikleri kapsamında yeni bir yaklaşım ihtiyacı doğmuştur. SQL tabanlı yapısal veri tabanlarının yanı sıra, bu ihtiyaca cevap olarak yapılandırılmamış veriyi yönetmede esnek ve ölçeklenebilir bir çözüm olarak NoSQL veri tabanı sistemleri geliştirilmiştir. Bu çalışmada, akıllı şehirlerde örnek teknolojiler değerlendirilmiş, coğrafi büyük verinin CBS ile entegrasyonu kapsamında hava izleme istasyonlarından elde edilen anlık sensör ölçme verileri kullanılarak NoSQL veri tabanı ortamı olan MongoDB’ de Hava Kalitesi İndeksi (HKİ) hesaplanmıştır. CBS ortamında hava izleme istasyonlarına yakın olan trafik sensörlerinden elde edilen veriler ile ortalama trafik yoğunlukları hesaplanmıştır. Elde edilen sonuçlara göre hava kalitesinin trafik ile ilişkisi belirlenmiştir.

## Geographical big data management and analysis in smart cities: the example of air quality

### Keywords

Air Quality Index (AQI)  
NoSQL  
Big Geographic Data  
Management  
MongoDB  
Big data

### ABSTRACT

With the development of information technologies, the data production style and the collected data volume have increased. With smart city applications, the importance of managing flowing from different data sources such as sensors, IoT, internet, wearable technologies and creating value from these data has increased. Today, traditional data storage and management approaches are insufficient for the management of large volumes and complex data collected, and a new approach has been born with the characteristics of big data such as volume, speed and diversity. Besides SQL-based databases, the NoSQL database provides flexible and scalable solution to manage unstructured data in response to this need. Such sample technologies were evaluated and displayed on the sensor obtained from air monitoring stations in the integration of geographic big data with GIS. Air Quality Index (AQI) was calculated in MongoDB on the NoSQL monitor. The average traffic density was calculated with the data obtained from traffic sensors close to air monitoring stations in GIS environment. According to the results, its relationship with air traffic has been determined.

## 1. Giriş

Bilişim teknolojilerinin gelişmesi ile özellikle son on yılda teknolojilerin günlük yaşamda kullanılması ve sürekli veri üretilmesi söz konusu olmuştur (George ve ark., 2014). 21. yüzyılın petrolü olarak tanımlanan büyük veri, bilgisayar ağlarının büyüüp karmaşıklaşması, internet, IoT, giyilebilir teknolojiler, sosyal medya ve sensörler gibi farklı ortamlardan elde edilen büyük hacimli veri kümelerini ifade eder (Doğan ve Arslantekin, 2016). Günümüz teknolojisinde sadece internet yoluyla çok küçük zaman dilimlerinde terabyte boyutlarına varan veriler üretilmekte ve dijitalleşmiş verilerin %90'a yakınının yalnızca son iki yılda toplanmış olduğu belirtilmektedir (Al Nuaimi ve ark., 2015; Aydın, 2017; Şahin ve ark., 2020).

Büyük veri kavramı ile verilerin elde edilme yollarının farklılaşması ve bunu takiben depolanabilir veri miktarında meydana gelen artış, bu büyük hacimdeki karmaşık verinin modellenerek saklanması için yeni arayışlara yönelmeyi zorunlu hale getirmiştir (Öztürk ve Atmaca, 2017; Şahin ve ark., 2020). Büyük veri, farklı kaynak ve ortamlardan gerçek zamanlı olarak elde edildiği için genellikle yapılandırılmamış olarak üretilmektedir (Aydın, 2017; Aydınoglu ve ark., 2020) ve son yıllarda dünyada üretilen tüm veri hacminin %80'den fazlasının yapısal olmayan formatta üretildiği belirtilmektedir (Wei ve ark., 2017; Taleb ve ark., 2018; Onay, 2020). Geleneksel veritabanı modellerinde bir şemaya bağlılıktan dolayı kaynak ve format çeşitliliği fazla olan büyük verinin yönetiminde daha pratik yaklaşımlara ihtiyaç vardır (Daşdemir ve Kara, 2019). Bu gereksinime çözüm olarak geliştirilen NoSQL (Not Only SQL-SQL'den Fazlası) veritabanı yaklaşımları, büyük hacimli ve yapısal olmayan veriyi performanstan feragat etmeden kullanmaya imkân sağlayan sistemlerdir (JRC, 2014; Daşdemir ve Kara, 2019; Aydınoglu ve ark., 2020). NoSQL veritabanları, büyük veri yönetiminde klasik ilişkisel veritabanı yaklaşımı olan SQL veritabanlarından çok daha iyi performans göstermekte ve büyük veri uygulamalarında özellikle esneklik ve yatay ölçeklenebilirlik avantajları sebebiyle sıklıkla tercih edilmektedirler (Schönberger ve Cukier, 2013; Aydın, 2017; Baralis ve ark., 2017; Aydınoglu ve ark., 2020; Anbaroğlu, 2021).

Akıllı şehirler, artan nüfus ile birlikte gittikçe yoğunlaşan şehir alanlarındaki trafik, hava ve çevre kirliliği, ulaşım, enerji, su ve atık yönetimi gibi altyapı sorunlarına çözüm olarak ortaya atılmış yeni nesil şehir yönetim anlayışıdır (Neirotti ve ark., 2014; Albino ve ark., 2015; Aydınoglu ve ark., 2020). En genel ifadesi ile akıllı şehirler, vatandaşlarına daha iyi hizmetler sunmak için teknoloji tabanlı yaşam ve planlama çözümlerini üreten ve şehir kaynaklarını etkin, ölçülebilir ve sürdürülebilir olarak kullanan akıllı sistemler bütünü olarak tanımlanabilir (Hollands, 2008; O'Grady ve O'Hare, 2012; Ballas, 2013; Terzi ve Ocakçı, 2017). Akıllı şehirlerde başarı için bilişim teknolojilerindeki gelişmelerin izlenmesi ve büyük veri işleme kabiliyetleri oldukça önemlidir. Bilgi ve İletişim Teknolojilerinin (BİT) yardımı ile trafik, enerji, ulaşım, hareketlilik, güvenlik, arazi ve taşınmaz yönetimi gibi kritik şehir sorunlarını çözmek için bilgisayar, bulut ve ulaşım ağlarının Kent Bilgi

Sistemleri (KBS) uygulamaları ile birlikte kullanılabileceği öngörülmektedir (Aydınoglu ve ark., 2020).

Şehirlerde yaşanabilirlik ve sürdürülebilirlik kapsamında çevre ve hava kalitesi önemli bir yere sahiptir. Dünya Sağlık Örgütü, hava kirliliğinin insan sağlığını etkileyen en önemli faktörlerden biri olduğunu ve her yıl yaklaşık 7 milyon insanın kirli havaya maruz kaldığı için yaşamını yitirdiğini ifade etmektedir (WHO, 2021). Yoğunlaşan şehirlerde nüfustaki hızlı ve kontrolsüz artış nedeniyle, mega şehirlerin neredeyse tamamı artan sıcaklık olaylarından ve yüksek hava kirliliği seviyelerinden etkilenmiştir (Sabrin ve ark., 2020). Bu nedenle şehirlerde hava kalitesi endeksinin (HKİ) belirlenmesi, hava kalitesi koşullarını ölçmek ve değerlendirmek kapsamında yetkililer için temel bir ihtiyaç haline gelmiştir (Tan ve ark., 2021). Bu anlamda HKİ değerlerinin, şehirlerde hareketliliğin yoğun olduğu alanlar için hesaplanması ve değer farklılıklarının yorumlanması oldukça önemlidir.

Bu çalışmada, akıllı şehir mimari bileşenleri, büyük coğrafi veri teknolojileri ve altyapıları ile akıllı şehir uygulamaları irdelenmiştir. Özellikle yapısal ve yapısal olmayan veri kavramları açıklanarak, yapısal olmayan veri yönetimi kapsamında ortaya çıkan veri altyapı çözümü olan NoSQL kavramı üzerinde durulmuştur. Akıllı şehirlerde büyük veri yönetimi kapsamında NoSQL veritabanlarını yaklaşımları olan MongoDB ve açık kaynaklı CBS veritabanı çözümü olan PostGIS kullanılarak, günümüzün en büyük problemlerinden olan trafik yoğunluğunun hava kalitesine olan etkisinin irdelenmesi amacı ile örnek uygulama gerçekleştirilmiştir. İstanbul ilinde trafik ve hava kalitesi ölçüm sensörlerinden gelen büyük coğrafi verinin MongoDB'nin etkin veri yönetim imkânları ile birlikte analizi sonucu, akıllı şehirler için örnek bir CBS uygulama yaklaşımı gerçekleştirilmiştir.

## 2. Akıllı şehirler

Akıllı şehir (smart city) kavramı, özellikle son 20 yılda BİT gelişiminde meydana gelen hızlı gelişim ile birlikte literatürde yoğun olarak tartışılmaya başlamıştır (Aydınoglu ve ark., 2020). Literatürde akıllı şehirler ile ilgili birçok tanım yer almakta "akıllı-smart" kavramı yerine "zeki-intelligent", "dijital-digital" ya da "algılayan-sensing" kavramları kullanılabilmektedir (O'Grady ve O'Hare, 2012). Bu kavramların hepsi aynı anlama gelmekle birlikte akıllı şehirler en temelde vatandaşları için yaşam kalitesini artırmayı, teknolojiyi yaşama entegre etmeyi ve daha iyi kentsel hizmetler sunmayı hedefler (Ballas, 2013; Demirci ve Köseoglu, 2018; Aydınoglu ve ark., 2020). Akıllı şehirler için birçok farklı tanım yapılmıştır. Örneğin OGC (2015)'de akıllı şehir; vatandaşları için daha yaşanabilir ve sürdürülebilir hizmetler sağlayan, yerel düzeyde paydaşlar arasındaki iş birliğini destekleyen, sistemlerin yeni teknolojiler ve yenilikçi yaklaşımlarla birlikte çalışabilirliğini sağlayan, verilerin etkin yönetimi ile şehrin sorunlarını ve ihtiyaçlarını öngören şehir olarak tanımlanmıştır.

Boyd Cohen'e göre (USDN, 2014) akıllı şehirler; akıllı ekonomi, akıllı çevre, akıllı insanlar, akıllı hareketlilik, akıllı yaşam ve akıllı devlet olmak üzere 6 ana bileşenden oluşur. KBS'nin de temel bileşenleri

olarak kabul edilebilen bu bileşenler için farklı uygulamalar ve farklı gereksinimlere ihtiyaç duyulabilir. Uygulama alanından bağımsız olarak değerlendirildiğinde, akıllı şehirler temel olarak teknoloji, veri ve insan faktörlerini içermektedir. Teknoloji açısından bir akıllı şehir uygulaması; İletişim Arayüzü (Communication Interfaces), Entegre Operasyon ve Kontrol Merkezleri (Integrated Operation and Control Centers), Sensörler ve Bağlantı Altyapısı (Connectivity Infrastructure) olarak 4 temel bileşeni içerir (IDB, 2016). Sürdürülebilir bir akıllı şehir mimarisinde, Algılama (Sensing), İletişim (Communication), Veri (Data) ve Hizmet/Uygulama (Service/Application) olmak üzere 4 ana katman (Aydınöglü ve ark., 2020; URL-1) mevcuttur (Şekil 1).

- Algılama katmanı, çeşitli tematik şehir aksiyonları hakkında fiziksel veri toplayan ve bu verileri veri katmanındaki merkezlere gönderen, şehir alanına yayılmış sensörler ve diğer veri algılama cihazlarını içerir. Bu çalışmada ulaşım uygulamalarına yönelik RTMS (Remote Traffic Microwave Sensor) gibi aktif sensörlerden ve diğer pasif veri sağlayıcılarından gelen verinin uygulamalarda kullanılması sağlanacaktır.

- İletişim katmanı, sensörler ve veri işleme platformu arasındaki iletişimi sağlar. WiFi, optik fiber, Ethernet gibi geleneksel LAN ve WAN iletişim teknolojileri ile UMTS ve LTE gibi mobil iletişim teknolojilerinden faydalanmaktadır.

- Veri katmanı, akıllı şehir mimarisinin veri altyapısının yapılandırıldığı yer olarak kabul edilebilir. Bu katman güvenilir ve birlikte çalışabilir verinin varlığıyla etkin bir şekilde çalışabilir. Bu katmanda veri dönüşüm araçları ile verilerin kullanılabilir hale gelmesi ve ileri istatistiksel modeller ile işlenmesi olanaklı hale

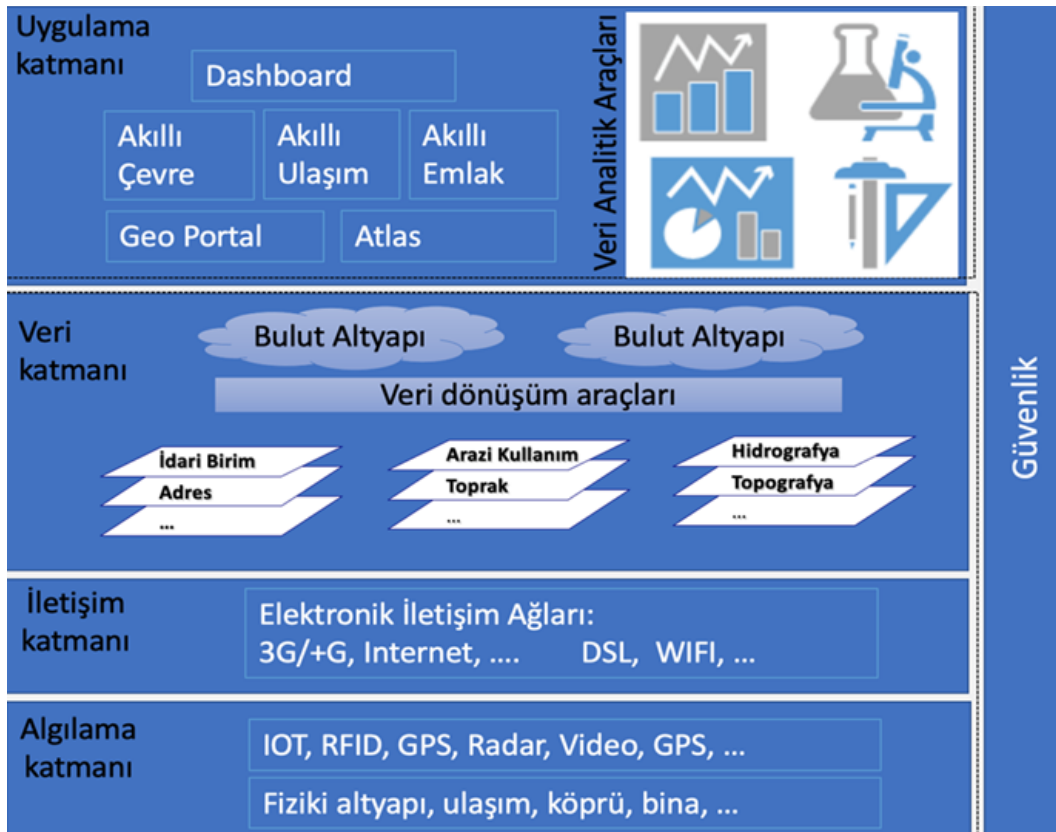
gelmektedir. Veri sunucuları ve bulut altyapısı bu katmanda yapılandırılmaktadır.

- Uygulama katmanı, farklı kurumlar ve birimler arası bir operasyon merkezi olarak çalışır. Kamu/belediye kurumları, bu katmanda yer alan web portalları ve mobil uygulamalar aracılığıyla bilgi paylaşmaktadır. Vatandaşlar da bu bilgilere kurulan portal ve uygulamalar aracılığıyla erişebilmektedir. Çalışmada geliştirilecek uygulamalara yönelik, mevcut donanımlar ile birlikte web ve veri tabanı arayüzlerinde veri erişimi için uygulama oluşturulmakta ve test edilmektedir. Büyük veri işleme ve karar destek için geliştirilen veri analitik araçları da bu katmanda yapılandırılmaktadır.

KBS'nin de parçası olarak kabul edilen akıllı şehir uygulama alanları, BIT'lerin de kilit etkinleştirici bileşen olarak önemli rol oynadığı "hard domains" ve "soft domains" olarak sınıflandırılabilir (Aydınöglü ve ark., 2020).

Hard domains, konut veya ticari amaçlı kullanılan binalar, enerji ve su yönetimi, doğal kaynaklar, çevre, ulaştırma ağları, atık yönetimi, sağlık, kamu güvenliği, hareketlilik ve lojistik alanlardaki akıllılaştırma süreç ve uygulamalarını kapsar. Büyük verilerin yönetimi için akıllı sensörler, kablosuz teknolojiler ve yazılım çözümleri ile daha etkin akıllı şehir uygulamalarına olanak sağlar (McKinsey Global Institute, 2011; McAfee ve Brynjolfsson, 2012).

Soft domains, eğitim, kültür, sosyal, ekonomi, arazi ve taşınmaz yönetimi ve e-devlet gibi uygulama alanlarında etkin ve akıllı yönetim çözümlerini kapsar (Washburn ve ark., 2010; Correia ve Wünstel, 2011; Nam ve Pardo, 2011; Chourabi ve ark., 2012; Accenture, 2015).



Şekil 1. Akıllı şehir mimari bileşenleri (Aydınöglü ve ark., 2020)

### 3. Büyük coğrafi veri kavramı ve teknolojileri

İnsan yaşamının parçası olarak BİT'lerin kullanımı, daha hareketli şehirlerin oluşmasına ve şehirlerde karmaşıklığın artmasına sebep olarak şehir yönetimini zorlaştırmıştır (Örselli ve Akbay, 2019). İnternet ve kablosuz ağ teknolojilerinin gelişmesi ile birlikte, şehirlerde başta sensörler olmak üzere her an veri üretilebilir hale gelmiş, akıllı cihazlar ve telefonlar ile her bir birey de veri kaynağı haline gelmiştir. Büyük veri yönetimi kavramı, akıllı şehir uygulamaları kapsamında çok çeşitli veri kaynaklardan anlık ve büyük hacimde akan verinin hızlı, etkin ve gerçek zamanlı yönetilerek bilgilerin konum bazlı toplanması ve coğrafi özellikleri ile analizi anlamında oldukça önemlidir (Yue ve Jiang, 2014; Al Nuaimi ve ark., 2015; Hashem ve ark., 2015; Duarte ve Teodoro, 2021; Goodchild, 2021).

Büyük veri için de çeşitli tanımlamalar yapılmıştır. Büyük veri, genel olarak geleneksel veritabanı sistemleri ile depolanamayan, yönetilemeyen ve analiz edilemeyen devasa hacimlere sahip yapılandırılmış ve yapılandırılmamış veri kümelerini (Li ve ark., 2016), bütüncül bir bilgi yönetimi stratejisini tanımlamaktadır (Oracle, 2017). Bilgi elde edilmesi, depolanması, dağıtılması, yönetimi ve analizini gerçekleştirmek için gelişmiş ve eşsiz veri depolama, yönetim, analiz ve görselleştirme teknolojilerine ihtiyaç duyacak kadar büyük ve karmaşık olan veri setlerini tanımlayan terimdir (NESSI, 2012; TechAmerica, 2012; Microsoft, 2013).

Büyük veriler; 3V olarak da adlandırılan hız (velocity), hacim (volume) ve çeşitlilik (variety) özellikleri ile birlikte büyük veriyi anlamlı ve kullanılabilir hale getiren değer (value) ve doğrulama (veracity) özelliklerine sahiptir (White 2012; Demchenko ve Membrey, 2014; Jin ve ark., 2015; Yin ve Kaynak, 2015; Viceconti ve ark., 2015; Aydınoğlu ve ark., 2020; Şahin ve ark., 2020). Bunların yanında literatürde bazı kaynaklarda bu özellikler 7V'ye (Khan ve ark., 2014; Landmark Solutions, 2015; Agraphari ve Rao, 2017) ve 10V'ye (BalaAnand ve ark., 2017; Firican 2017) çıkabilmektedir (Şahin ve ark., 2020). Temel anlamda benzer özellikleri ifade eden bu bileşenler büyük verinin hem zorluklarını hem de sağlayacağı avantajları anlamak için büyük öneme sahiptir (Firican, 2017).

**Hız (Velocity):** Büyük verinin üretilme, analiz edilme, saklama ve işleme hızını ifade eder (Al Nuaimi ve ark., 2015; Firican, 2017; Khan ve ark., 2014). Mobil cihazlar ve sosyal medya ağlarının yaygınlaşması ile veri üretim hızı doğanın çok üzerine çıkmıştır. Bu durum verileri depolamak ve işlemek için mevcut yöntemlerin yetersiz kalmasına sebep olmuştur.

**Hacim (Volume):** Büyük verinin en iyi bilinen özelliklerinden biridir. Mevcut tüm kaynaklardan üretilen verinin hacmini ifade eder (Al Nuaimi ve ark., 2015; Firican, 2017; Khan ve ark., 2014).

**Çeşitlilik (Variety):** Büyük veri ile yapılandırılmış veriden çok yapılandırılmamış veri ele alınmaktadır. Veriler, ses, görüntü, metin ve video gibi farklı ortamlardan ele edilebilir. Büyük veri çeşitliliği 3 temel

formda ifade edilebilir (Al Nuaimi ve ark., 2015; Firican, 2017; Khan ve ark., 2014):

- Yapısal veri: Tanımlanmış bir veri tabanı şemasında sabit sütunlar altında kaydolarak tutulabilen ve işlenebilen verileri ifade eder (Ming ve ark., 2014; Seveli, 2016). Veriler belirli tür ve boyutlarda olduğundan, yapısal verileri modellemek, depolamak, sorgulamak, işlemek ve görselleştirmek kolaydır. Sorgular, yapısal sorgu dili olan SQL ile gerçekleştirilir.

- Yarı yapısal veri: Yarı yapısal veriler, yapısal verilere benzer şekilde kolay şekilde ilişkilendirilebilir ve işlenebilir, ancak yapısal veriler gibi sağlam bir modele sahip değildir (Rusu ve ark., 2013; Aktan, 2018). XML (Extensible Markup Language) ve JSON (JavaScript Object Notation) yapıdaki veriler yarı-yapısal verilere örnektir.

- Yapısal olmayan veri: Önceden tanımlanmış bir yapı veya biçime sahip olmayan veriler yapısal olmayan verilerdir (Ghazal ve ark., 2013; Seveli, 2016). Yapısal olmayan verilere örnek; basit metin dosyaları, e-posta, sosyal medya verisi, ses ve video gibi çoklu-ortam dosyalarıdır.

**Değer (Value):** Büyük verinin en önemli bileşeni ve istenen sonucudur. Büyük veriden çıkarılacak faydalı bilgiyi ifade eder. Eğer büyük veriden değer kavramı üretilmezse diğer özellikler anlamsızdır (Al Nuaimi ve ark., 2015; Firican, 2017; Khan ve ark., 2014).

**Doğrulama (Veracity):** Büyük verinin en önemli bileşeni, analizler sonucunda veri çöplüğünden çıkarılacak olan değerdir. Ancak doğru bilgi çıkarabilmek için verilerin doğru tutulması gereklidir. Doğrulama kavramı verinin yapılacak analiz için ne kadar anlamlı olduğunun ölçüsüdür (Firican, 2017; Khan ve ark., 2014).

Ayrıca Değişkenlik (Variability), Geçerlik (Validity), Güvenlik (Vulnerability), Dalgalanma (Volatility) ve Görselleştirme (Visualization) kavramları ile de büyük veri tanımlanmaktadır.

En çok bilinen ve büyük veri teknolojilerinde öncü olan Apache Hadoop, popüler arama motoru kütüphanesi olan Apache Lucene projesinin de geliştiricisi olan Doug Cutting tarafından Java dili ile geliştirilmiş açık kaynaklı bir yazılım kütüphanesidir. Apache Hadoop, basit programlama modellerini kullanarak, bilgisayar kümeleri arasında, büyük veri setlerinin dağıtık işlemlerine izin verir ve bir sunucudan binlerce makineye, hesaplama ve depolama olanağı yaratır. Hadoop, tek bir sunucudan yerel hesaplama ve depolama sağlayan binlerce makineye kadar ölçeklendirmeye olanak sağlar. Ayrıca sağladığı dağıtık dosya sistemi ve desteklediği paralel işlem gücü ile önemli performans artışı sağlamaktadır (URL-2). Apache Hadoop ekosistemi; Hadoop Common, Hadoop Dağıtılmış Dosya Sistemi (HDFS), Hadoop YARN ve Hadoop MapReduce dört temel modülden meydana gelmektedir. Apache Hadoop ekosisteminde; küme ve veri yönetimi altyapıları, veritabanı ve veri ambarı altyapıları ve analiz altyapıları için teknolojiler tanımlanmıştır.

### 3.1. Büyük veri altyapıları

Büyük veri yönetiminde, NoSQL ve NewSQL olarak ifade edilen iki veritabanı türü kullanılmaktadır (JRC, 2014).

**NoSQL – Not Only SQL** (SQL ve daha fazlası), ilişkisel veri tabanı yönetim sistemlerine alternatif olarak ortaya çıkmıştır. SQL tabanlı ilişkisel veritabanı, yenilikler devam etse de verilerin orantısız büyümesi, sensör verilerinin kullanılmaya başlaması, sosyal medya ve büyük ölçekli web uygulamalarının yaygınlaşması ile farklı kanallardan gelen büyük veriyi yönetmede yetersiz kalmıştır (Aydınöğlü ve ark., 2020). Geleneksel ilişkisel veritabanları, şema öncelikli bir tasarıma sahip olduğu için büyük veri yönetiminde yazılımsal ve donanımsal kaynak tüketiminin artmasına sebep olmakta (Aydın, 2017), büyük verileri hızlı ve etkin bir şekilde yönetebilmek hantal ve maliyetli olmaktadır.

NoSQL veritabanları, önceden belirlenmiş bir kayıt yapısı gerektirmeyen, farklı tip ve büyüklüklerde veri içerisinden arama yapılmasına olanak sağlar. Bu veritabanı yapısı, yapısal dağınıklığa izin vermesinin karşılığında daha fazla işlem ve depolama alanı gerektirse de maliyetleri düşüğü ve ölçeklenebilirlik, esneklik ve performans açısından sağladığı avantajlar doğrultusunda büyük veri uygulamalarının hemen hepsinde kullanılmaktadır (Schönberger ve Cukier, 2013; Daşdemir ve Kara, 2019; Anbaroğlu, 2021).

NoSQL 'in büyük veride dağıtık kurgu mimarisinde daha esnek, yüksek performans ve güvenilirlik sağlar (Gökşen ve Aşan, 2015; Aydın, 2017; Aydınöğlü ve ark., 2020). NoSQL sistemler, ilişkisel veritabanlarının aksine kullanıcıya şemasız veri saklayabilme olanağı sağlar ve ölçeklenebilirler. Ancak NoSQL ve klasik SQL ilişkisel veritabanı sistemleri, birbirlerinin alternatifi olmamakla birlikte, bu iki sistemin farklı ihtiyaçların çözümüne yönelik geliştirilmiş teknolojiler olduğu söylenebilir.

Daha iyi performans ve ölçeklenebilirlik özellikleri beraberinde tutarlılık açısından bir takım fedakârlık yapılmasını gerektirir. İlişkisel veritabanları, **ACID** olarak adlandırılan özelliklere sahiptir. Bu özellikler kısaca; "hepsi ya da hiçbiri" anlamına gelen **Atomicity (Bölünmezlik)**; "Her bir işlemin sonucunun tutarlı olması" anlamında **Consistency (Tutarlılık)**; "tüm işlemlerin bağımsızlığı" anlamına gelen **Isolation (Yalıtım)** ve "veritabanı sistem hatalarından arınma" anlamına gelen **Durability (Dayanıklılık)** özellikleridir (Pokorny, 2013). NoSQL veritabanlarında, ACID özelliklerini eşzamanlı olarak sağlamak zordur (Aydın, 2017), tutarlılık duygusu, uzun süreli bir güncelleme olmadığında, tüm güncellemelerin sistem üzerinden yayılacağı ve tüm düğümlerin tutarlı olacağı anlamına gelen nihai tutarlılığı ifade eder (JRC, 2014). Bu anlamda NoSQL veritabanları için ACID ilkelerine karşılık olarak **BASE (Basicly Available-Kolay Ulaşılabilir, Soft state-Esnek, Eventuallyly consistent-Eninde sonunda tutarlı)** özellikleri tanımlanmıştır (Gökşen ve Aşan, 2015):

- Kolay Ulaşılabilirlik: Verileri dağıtık bir ortamda kopyaları ile tutar ve veri erişim sorunlarını ortadan kaldırır. Veriler farklı sunucuda depolanır ve veri güvenliği sağlanmış olur.

- Esnek Durum: NoSQL sistemler SQL sistemlerden farklı olarak tutarsız verilerin barınmasına olanak tanır.

- Eninde Sonunda Tutarlı: NoSQL sistemler anlık değil gelecekte bir zamanda tutarlılığı garanti eder.

NoSQL sistemler veri saklama ve veri okuma yöntemleri bakımında çeşitlilik gösterir ve temel anlamda beş çeşitten oluşur (JRC, 2014):

- Anahtar-Değer (Key-value store) Veritabanları: En basit ve kolay NoSQL veritabanlarıdır. Veritabanında saklanan her bir veri bir anahtar (key) ve bu anahtara karşılık gelen değer (value) ikilisi olarak saklanır. En popüler örnekleri olarak redis, Voldemort ve riak verilebilir.

- Sütun Bazlı (Wide-column store) Veritabanları: Veriler satır değil sütun tabanlı gruplanarak tutulur. Bu tarz veritabanları büyük verileri dağıtık olarak saklayabilmek için oluşturulmuştur. Popüler sütun veritabanlarına örnek olarak Hadoop ekosistemi içerisindeki HBase ve Cassandra, Google'un kullandığı veritabanı yapısı olan BigTable ve BigQuery verilebilir.

- Döküman Bazlı (Document store) Veritabanları: Anahtar-Değer veri tabanlarına benzemektedir. Ancak burada anahtarlara karşılık gelen değerler versiyonlanmış dökümanlardan oluşur ve daha karmaşıktır. Bu veritabanlarına örnek olarak MongoDB, CouchBase, Apache CouchDB ve Xist verilebilir.

- Graf (Graph store) Veritabanları: Sütun ve satır kullanım yerine daha esnek olan graf modelini temel alarak kenarlar (edge) ve düğüm noktaları (node) yardımıyla ağ (network) hakkında bilgi tutar. Birden çok bilgisayara ölçeklenebilir. Popüler graf veritabanı örnekleri olarak Neo4j, InfoGrid, Oracle Spatial ve Graph veritabanı verilebilir.

- Dizi (Array) Veritabanları: Özellikle raster veri olarak da adlandırılan diziler için etkin veritabanı hizmetleri sağlar. Dizi yapısı, genellikle sensör, simülasyon, görüntü veya istatistik verileri temsil etmek için kullanılır. Dizi veritabanları, bu verilere esnek ve ölçeklenebilir depolama ve erişim imkânı sunmayı amaçlar. Dizi veritabanları; coğrafi-bazlı 1B'lu sensör veri dizileri, 2B uydu görüntüleri, 3B x, y, t görüntü-zaman serileri ve x,y,z jeofizik verileri ile 4B'lu x,y,z,t iklim ve oşinografi verilerinin yönetiminde etkin şekilde kullanılır. Popüler örnekleri arasında Oracle Spatial ve Graph, MonetDB, PostGIS, Rasdaman ve SciDB bulunur.

**NewSQL** veritabanları, modern ilişkisel veritabanı sistemleri olarak kabul edilebilir. Geleneksel veritabanı sistemleri ACID özelliklerini korurken, NoSQL sistemleri hızlı veri işleme ve ölçeklenebilirlik performansını sağlamayı hedeflemektedir. NewSQL veritabanları şu şekilde sınıflandırılabilir:

- Yeni mimariler: Her düğümün verinin bir alt kümesine sahip olduğu dağıtık kümelendirilmiş ve ortak bir düğümleri olmayan yeni veritabanı platformlarıdır. NewSQL veritabanlarına örnek olarak; Google Spanner, Clustrix, VoltDB, MemSQL, FoundationDB ve Nuodb verilebilir.

- SQL Motorları: Bu kategoride yer alan sistemler, SQL programlama arabirimini sağlayan ve SQL için yüksek düzeyde optimize edilmiş depolama motorları ile donatılmıştır. Örnekleri arasında TokuDB ve InfiniDB bulunmaktadır.

- Transparan parçalama: Bu sistemler, veritabanlarını otomatik olarak birden fazla düğüme bölmek için veritabanı veya arama motoru içerisinde yer alan veri kümelerinin yatay bölünmesini sağlar. Bu sistemlere örnek olarak dbShards, Scalearc ve ScaleBase verilebilir.

#### 4. Örnek uygulama

Şehir yaşamında, üretilenden daha fazla kaynak, enerji ve zaman bilinçsiz tüketim eğilimine girmiştir (Demirarslan ve Akıncı, 2018). Söz konusu bilinçsiz ve aşırı tüketim, şehirlerde hava, su, toprak kirliliği ve yeşil alanların yok olması başta olmak üzere, trafik, gürültü ve aşırı ısınma gibi çok çeşitli problemlere sebep olmuştur (Cui ve Shi 2012; Lu ve Cao 2015). Hava kirliliği, Dünya Sağlık Örgütü (WHO) tarafından insan sağlığını etkileyen en önemli faktörlerden biri olarak kabul edilmekte ve her yıl yaklaşık 7 milyon insanın havada yer alan kirletici partiküllere bağlı hastalıklar nedeniyle hayatlarını kaybettiği belirtilmiştir (WHO, 2021). Öldürücü etkisinin yanında uzun süre kirli havaya maruz kalındığında, kanser, bronşit ve kronik öksürük gibi akciğer hastalıkları başta olmak üzere çeşitli kalp ve damar yolu hastalıkları ile depresyon, kronik yorgunluk gibi ruh sağlığı hastalıklarına neden olduğu belirtilmektedir (Çetin ve Demirci 2015; Özşahin ve ark., 2016).

Hava kirliliği, doğal veya insan kaynaklı bir şekilde ortaya çıkabildiği bilinmekte, günümüzde ulaşılan ciddi seviyenin temel sorumlusunun insan kaynaklı faaliyetler olduğu belirtilmektedir (Özşahin ve ark., 2016). Söz konusu bu kirletici faaliyetlerin temel kaynağı, kentleşme ve endüstrileşme ile birlikte artan trafik, sanayi faaliyetleri ve atık gazlar oluşturmaktadır. Bu çalışmada, hava kalite endeksi belirlenen saat dilimlerinde hesaplanarak kirliliğin en önemli sebeplerinden olan trafik ile olan ilişkisi irdelenmektedir.

##### 4.1. Çalışma alanı seçimi ve büyük veri yönetimi

Çalışma alanı olarak Türkiye'nin en yoğun kentleşme ve nüfus yoğunluğuna sahip şehri İstanbul seçilmiştir. İstanbul yalnızca ülkemizin en kalabalık şehri değil, aynı zamanda belediye sınırları dikkate alınarak yapılan sıralamada Avrupa'da birinci, dünyada ise altıncı sıradadır (URL-3). İstanbul 5.461 km<sup>2</sup> lik yüzölçümüne ve 39 ilçeye sahiptir. 2020 yılı sayımlarına göre ilde 15.462.452 kişi ikamet etmektedir (www.tuik.gov.tr). İstanbul'un sahip olduğu bu yoğun nüfus ve ileri derecedeki sanayileşme, hava, su ve toprak kalitesi üzerinde olumsuz sonuçlara neden olmuştur. Ayrıca denetimsiz ve plansız kentsel alanlar, trafik yoğunluğu ve gürültü kirliliği gibi sorunlara neden olmaktadır.

Uygulama kapsamında öncelikli veri ihtiyacı hava kirleticilere ve trafik yoğunluğuna ait verilerdir. Çalışma

alanına ait hava kirletici verileri, Çevre ve Şehircilik Bakanlığı tarafından 2005 yılında kurulmuş olan 355 izleme istasyonundan gelen hava kalitesi izleme verilerini paylaşan Ulusal Hava Kalitesi İzleme Ağı (www.havaizleme.gov.tr) veri portalından temin edilmiştir. İstanbul ili çalışma alanına ait Şekil 2'deki haritada görülen 39 adet istasyon verisine ait CO, NO<sub>2</sub>, NO<sub>x</sub>, O<sub>3</sub> SO<sub>2</sub>, PM<sub>10</sub> ve PM<sub>2,5</sub> konsantrasyonları 1 Ocak – 31 Aralık 2020 tarihleri arasında saatlik olarak CSV formatında temin edilmiştir.

Trafik yoğunluğunu temsil eden trafik sensör verileri ise İBB'nin belediye ve çevre kuruluşları tarafından üretilen verileri kullanıcılara açık olarak paylaştığı İBB Açık Veri Portalı (<https://data.ibb.gov.tr/>)'ndan CSV formatında temin edilmiştir. Veri setinde İstanbul'un coğrafi olarak eşit bölünmüş parçalarına ait saatlik araç sayısı ile minimum, maksimum ve ortalama hız bilgilerini içermektedir. Kullanılan veri seti 1 Ocak – 31 Aralık 2020 tarihleri arasındaki trafik bilgilerini içermektedir.

Temin edilen trafik yoğunluğuna ait CSV formatındaki sensör verisi konum bilgisi (enlem-boyun) içerdiğinden daha etkin bir yönetim için PostgreSQL/PostGIS ortamında, hava izleme sensorlerinden elde edilen CSV formattaki ölçümler ise HKİ hesaplaması için daha esnek ve hızlı yönetim imkânı sunan MongoDB ortamında depolanmıştır. MongoDB yapısal olmayan büyük veri altlıklarının yönetiminde geliştirme ve ölçekleme kolaylığı ile yaygın olarak kullanılan belge tabanlı bir NoSQL veritabanı sistemidir (Daşdemir ve Kara, 2019). Eklenen her kayıt bir belge olarak ifade edilir ve BSON, JSON ve GeoJSON gibi veri formatlarını destekler. Ayrıca temel coğrafi analizlere imkân sağlar (URL-4). Çalışma kapsamında temin edilen hava kalitesi parametreleri MongoDB ortamına aktarıldı. Veriler, \$match fonksiyonu ile 15-19 saatleri arasında olacak şekilde filtre edildi. Daha sonra \$group fonksiyonu kullanılarak veriler sensör bilgisi ve tarih alanlarına göre gruplandı ve günlük ortalama değerler bulundu. \$switch fonksiyonu ile ortalama değerlerin hangi kırılım aralığına düştüğü belirlendi ve HKİ hesaplandı. Ardından gün bazlı etkin kirleticiler tespit edildi. PostgreSQL büyük hacimli yapısal verinin yönetiminde kullanılan güçlü, açık kaynak kodlu ve nesne ilişkisel veri tabanı yönetim sistemidir. Coğrafi veri yönetiminde PostGIS ile binden fazla mekansal fonksiyonu desteklemektedir (Agarwal ve Rajan, 2016). Yarı-yapısal olan CSV, JSON ve XML gibi NoSQL veri formatlarını destekler ve MongoDB gibi çeşitli veritabanlarına erişim imkânı sağlar (URL-4). Çalışma kapsamında elde edilen gün bazlı etkin kirletici tablosu PostGIS ortamına aktarıldı ve hava kalitesi sensörlerinin geometrileri ile birleştirildi. Ardından trafik yoğunluğu verisi 15:00-19:00 saatleri arasında filtre edildi ve günlük ortalamaları bulundu. PostGIS'in st\_dwithin fonksiyonu kullanılarak hava kalitesi sensörlerine 250 m mesafede bulunan trafik yoğunluğu noktalarının değerleri bulunmuştur ve etkin kirletici tablosu ile tarih ve st\_dwithin fonksiyonunun sonucuna göre birleştirilerek sonuç tablo elde edilmiştir.



Şekil 2. Hava izleme istasyonlarının haritası

#### 4.2. Hava kalitesi hesaplanması

Hava kirleticiler gazlar ( $CO$ ,  $SO_2$ ,  $NO_x$ ,  $CO_2$ ,  $O_3$ ) ve partiküller ( $PM_{10}$ ,  $PM_{2,5}$ ) olmak üzere iki kısımda incelenirler.

- *Karbonmonoksit ( $CO$ )*; yanma işlemi sırasında oksijen miktarı yetersiz olduğunda karbonun tamamı oksitlenemez ve çok kuvvetli bir zehirli gaz olarak meydana getirir. Solunan havada konsantrasyonu artarsa, kanla dokulara oksijen taşınmaz ve hücre ölümü meydana gelir. Uzun süreli maruz kalındığında ölüme sebep olabilir.

- *Kükürtdioksit ( $SO_2$ )*; yakıtların içinde bulunan kükürdün oksitlenmesi sonucu oluşan renksiz ve sağlığa zararlı gazdır. Ana kaynakları, termik santraller ve endüstriyel atıklardır. Uzun süre ile maruz kalındığında solunum ve kalp hastalıklarına sebep olur.

- *Azotoksitler ( $NO_x$ )*; en önemli kaynağı araç egzozu ve yakma/bertaraf tesisleridir. Kötü kokulu bir gaz olan azotoksit bileşiklerinin yüksek konsantrasyonları zehirlenmelere yol açarken, aşırı düzeyde maruz kalındığında ise öldürücü ciğer hastalıklarına ve felce yol açabilir.

- *Karbondioksit ( $CO_2$ )*; renksiz, kokusuz ve zehirsiz olan karbondioksit gazı normalde atmosfer havasının %0.03-0.04'ünü kaplamaktadır. Havada daha yüksek değerlere ulaştığında merkezi sinir sistemi oksijen azlığından dolayı etkilenir. Konsantrasyon değeri %15'lere ulaştığında baş ağrısı, kusma, mide bulantısı ve

baygınlığa yol açabilir. Daha yüksek konsantrasyonlarda, kan dolaşımı yetersizliğine ve ölüme sebep olabilir.

- *Ozon ( $O_3$ )*; atmosferin üst katmanlarında yer alan ozon tehlikeli morötesi ışınlarından koruyucu etkiye sahipken, atmosferin en alt katmanı olan troposferde, ozon insan sağlığına ve doğaya zarar veren önemli bir kirleticidir. Yüksek ozon konsantrasyonları daha az oksijen solumaya bağlı kalp-damar hastalıkları veya astım gibi solunum hastalıklarına neden olabilir.

- *Partikül maddeler ( $PM$ )*; birçok farklı bileşenden oluşmuş olan fiziksel ve kimyasal yapısı sağlık açısından oldukça önemli olan maddelerdir.  $PM_{10}$ , aerodinamik çapları  $10 \mu m$  den daha küçük olan toz, polen ve küf gibi partikülleri;  $PM_{2.5}$  ise aerodinamik çapları  $2,5 \mu m$ 'den küçük olan yanma ve metal parçacıkları, organik bileşikler gibi ince partiküllerdir. Havadaki partikül maddeler insan sağlığını etkileyen en zehirli ve kanser yapıcı kirleticilerden biridir.

HKİ (Air Quality Index-AQI), hava kalitesinin değerlendirilmesi için tüm dünyada yaygın olarak kullanılan sınıflama sistemidir. HKİ ile havadaki kirleticilerin değer aralıklarına göre hava kalitesini iyi, orta, hassas, sağlıksız ve tehlikeli gibi derecelendirmeler yapılarak sağlık endişe seviyeleri özetlenmektedir (Tablo 1). Ulusal Hava Kalitesi İndeksi, Amerika Birleşik Devletleri Çevre Koruma Ajansı (EPA)'nın geliştirdiği hava kalitesi indeksine göre hesaplanmaktadır (ÇŞB, 2017). HKİ, temel kirleticilerin konsantrasyonlarını birleşik bir endekle özetleyerek hava kalitesi hakkında

mevcut durum ve ilerisi için tahminler yapmaya olanak sağlar (Mirabelli ve ark., 2020).

Endeks formülüne göre bir coğrafi konum için HKİ değeri, her bir kirletici için ayrı ayrı hesaplanır ve en yüksek değere sahip kirletici etkin kirletici kabul edilerek ona ait endeks ilgili konumdaki HKİ değeri olarak kullanılır (U.S. EPA, 2018). EPA tarafından geliştirilen HKİ formülü;

$$I_p = \frac{I_{Hi} - I_{Lo}}{BP_{Hi} - BP_{Lo}} * (C_p - BP_{Lo}) + I_{Lo}$$

şeklinde. Burada IHi ölçümün düştüğü kesme noktası aralığının maksimum değerini, ILo ise minimum değerini, BPHi aralığın üst sınırını, BPLo aralığın alt sınırını, Cp ise o an için ölçülen kirletici değerini ifade etmektedir.

Formül hesaplamalarında yer alan sabitler U.S. EPA (2018) standardında yer alan kesme noktaları tablosundan gerekli dönüşümler yapılarak kullanılmıştır. Hesaplama aşamasında O3 verisi virgülden sonra 3 hane, PM2.5 ve CO verisi virgülden sonra 1 hane, PM10, SO2 ve NO2 verileri tamsayı olacak şekilde kesilmiştir (U.S. EPA, 2018).

**Tablo 1.** HKİ sınıf değerleri ve anlamları (URL-5)

Hava Kalitesi İndeksi (HKİ) Değerler	Sağlık Endişe Seviyeleri	Renkler	Açıklaması
0- 50	İyi	Yeşil	Hava kalitesi memnun edici ve hava kirliliği az riskli veya hiç risk teşkil etmiyor.
51- 100	Orta	Sarı	Hava kalitesi uygun fakat alışılmadık şekilde hava kirliliğine hassas olan çok az sayıda insan için bazı kirleticiler açısından orta düzeyde sağlık endişesi oluşabilir.
101- 150	Hassas	Turuncu	Hassas gruplar için sağlık etkileri oluşabilir. Genel olarak kamunun etkilenmesi olası değildir.
151- 200	Sağlıksız	Kırmızı	Herkes sağlık etkileri yaşamaya başlayabilir, hassas gruplar için ciddi sağlık etkileri söz konusu olabilir.
201- 300	Kötü	Mor	Sağlık açısından acil durum oluşturabilir. Nüfusun tamamının etkilenmesi olasıdır.
301- 500	Tehlikeli	Kahverengi	Sağlık alarmı: Herkes daha ciddi sağlık etkileri ile karşılaşabilir.

HKİ hesaplamaları için mongoshell'de yazılan kodda kullanılan fonksiyonlar ve işlevleri sırası aşağıdaki gibi listelenmiştir.

- **\$addField:** Girdi olarak aldığı belgelerdeki tüm alanlara ek olarak belgeye yeni alan eklemek için kullanılır.
- **\$dateFromString:** String olarak aldığı alanı bir date/datetime objesine döndürür. Parametre olarak string bilgi içeren alanı (dateString), alanın dönüştürüleceği tarih formatını (format) ve saat dilimini (timezone) alır.
- **\$hour:** Bir tarih objesinin saat bölümünü 0-23 arasında bir integer sayı olarak döndürür.
- **\$dateToString:** Bir tarih nesnesini kullanıcının tanımlayacağı formatta bir string nesnesine dönüştürür.
- **\$match:** Belgeleri kullanıcının tanımladığı parametrelere göre filtreleyerek bir sonraki aşamaya sadece kullanıcının istediği belgelerin geçmesini sağlar.
- **\$and:** Belgede üzerinde filtreleme yaparken sağlanması istenen bir veya birden fazla koşulun kontrol edilmesini sağlar. Aldığı koşullardan biri sağlanmazsa diğer koşullar kontrol edilmez ve belge sonuçlara eklenmez.
- **\$ne:** Belirtilen alanın, belirtilen değere eşit olmadığı belgeleri ve alanı içermeyen belgeleri döndürür.
- **\$gte:** Bir alanın belirtilen değerden büyük ve eşit olmasını kontrol eder ve koşulu sağlayan belgeleri seçer.
- **\$lte:** Bir alanın belirtilen değerden küçük ve eşit olmasını kontrol eder ve koşulu sağlayan belgeleri seçer.
- **\$group:** Girdi belgelerini \_id olarak aldığı parametredeki alanlara göre gruplar ve yeni belgeler

çıkartır. Yeni belgeler gruplama aşamasında hesaplanmış ortalama, toplam gibi değerler içerebilir.

- **\$avg:** Gruplama aşamasında aynı \_id'ye sahip her bir belgenin parametre olarak verilen alana ait sayısal değerlerin ortalamasını döndürür. Sayısal olmayan değerleri yok sayar.
- **\$multiply:** Sayıları ve alanları çarpar ve sonucu döndürür.
- **\$trunc:** Bir ondalıklı sayıyı tamsayıya çevirmeye ya da belirli bir basamaktan kesmeye yarar.
- **\$project:** Sorgulama aşamasında istenilen alanları bir sonraki aşamaya geçirmek için kullanılır. Bir sonraki aşamaya geçen belgeler önceki alanları içerebileceği gibi yeni hesaplanmış alanlar da içerebilir.
- **\$switch:** Farklı koşullara sahip bir dizi durum ifadesini işletir. Koşullar branches yapısının altında tutulan case yapılarında sıralanır ve case bloğu true değer üretirse then bloğu çalışır. Eğer hiçbir şart sağlanmazsa default bloğunda belirtilen değer alana eklenir.
- **\$round:** Bir ondalık sayıyı ne yakın tamsayıya yuvarlar.
- **\$sort:** Tüm girdi belgeleri sıralar ve sıralı bir şekilde geri döndürür. Parametre olarak sıralanacak alanlar ve ilgili sıralama düzenini belirten değeri (1: artan, -1: azalan) alır.

\$dateFromString fonksiyonu ile string olarak tutulan tarih verisi, tarih objesine çevirilerek \$addField fonksiyonu ile belgeye yeni alan olarak eklendi. \$hour fonksiyonu ile tarih objesinden saat bilgisi, \$dateToString fonksiyonu ile tarih bilgisi çekilerek



\$addFields fonksiyonu ile belgeye eklendi. \$match fonksiyonu ile veri NaN olmayacak (\$ne: NaN) ve sıfırdan büyük veya eşit (\$gte:0) olacak şekilde filtrelendi. \$group fonksiyonu ile tarih ve sensöre göre gruplanan verilerin birim dönüşümü için belirlenen katsayı ile \$multiply fonksiyonu yardımıyla çarpıldı ve ardından \$avg fonksiyonu ile ortalaması alındı. Bulunan ortalama değerler \$trunc fonksiyonu ile kesildi ve \$switch fonksiyonu ile ait olduğu kesme aralığının katsayıları ile AQI hesaplandı. \$round fonksiyonu ile en yakın tamsayı değerine yuvarlandı ve ilgili kirleticiye ait HKİ değeri hesaplandı. Benzer şekilde her bir kirletici için hesaplamalar yapılarak her bir istasyon için her kirleticinin 8 saatlik ortalama 1 Ocak 2020 – 31 Aralık 2020 tarihleri için benzer şekilde hesaplanmıştır. U.S. EPA (2018) standardında belirttiği üzere bir noktadaki esas kirletici, ilgili noktadaki etkin tüm kirleticiler içinde en yüksek HKİ değerine sahip olan ve onun HKİ değeri ilgili konumun HKİ değeri olarak kabul edilir. Toplam 39 istasyonda günlük her bir kirletici için hesaplanan HKİ değerleri arasından etkin kirletici ve ona ait HKİ değerinin sorgulaması için \$sort ve \$group fonksiyonları kullanılmıştır.

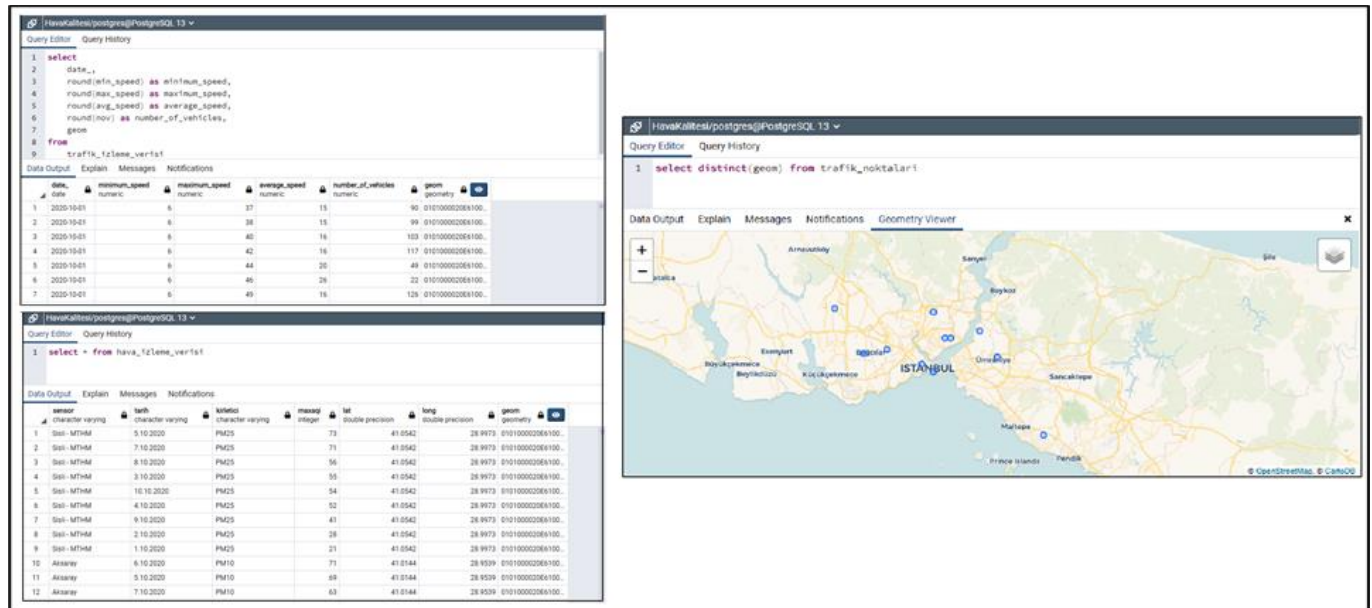
Hesaplanan HKİ değerleri sonucunda MongoCompass arayüzü ile sorgu sonuçları CSV formatında dışarı aktarılmıştır. Hesaplamalar tüm kirleticiler için yapıldıktan sonra raporlar tek bir MongoDB koleksiyonunda toplanmış ve gün bazlı etkin kirletici değerleri bulunarak CSV formatında dışarı aktarılmıştır. Etkin kirletici değerlerinin yer aldığı rapor PostgreSQL veritabanına aktarılmış ve hava izleme istasyonlarının geometrilerinin tablo ile birleştirilmiştir. PostGIS ortamında hava izleme istasyonlarına yapılan analiz ile 250 m yakınlıktaki trafik sensörleri ve hız değerleri elde edilmiştir. Buna göre 11 istasyon için bütünlük HKİ ve trafik bilgisi elde edilmiştir. Trafik yoğunluğu verileri filtrelenmiş, filtrelenen verilerden gün bazlı 15:00-19:00 saatleri arasını kapsayacak şekilde ortalama hızlar ve araç sayısı hesaplanmıştır.

Şekil 3'de uygulama arayüzünde görüldüğü gibi yakın konumda (250m yarıçap içerisinde) bulunan sensörler

PostgreSQL ortamında sorgulanıp belirlenmiş ve 11 sensörden gelen hava kalitesi ve trafik ortalama hız verileri görselleştirilmiştir. Bu şekilde belirlenen her bir konumda trafiğin yoğun saatlerini temsil eden zaman diliminde, trafik yoğunluğundan kaynaklanan hesaplanan hava kalitesi endeks değeri arasındaki ilişki irdelenmiştir. Bu kapsamda elde edilen veriler arasındaki ilişkiyi analiz etmek için Pearson korelasyon analizi yapılmıştır. Pearson korelasyon analizi, veri kümesindeki sürekli değişkenler arasındaki ilişkiyi ölçen en yaygın istatistiksel analiz tekniğidir (Schober ve ark., 2018). Bu analiz tekniğinde hesaplanan Pearson korelasyon katsayısı (r) -1 ile +1 arasında bir değer alır ve bu katsayı tahmin edilen iki değişkenin kapsamını ölçmek için kullanılır. Ayrıca değişkenler arasındaki ilişkinin yönü ve büyüklüğü hakkında da bilgi sağlar. r'nin değeri -1'e yakın değerler alırsa değişkenler arasında negatif, + 1'e yakın değerler alırsa pozitif bir ilişki olduğu söylenebilir (Egghe ve Leydesdorff, 2009; Sensuse ve ark., 2015).

Bu kapsamda veri sorgulamalarından sonra hesaplanan HKİ ile trafik yoğunluğu verileri arasındaki ilişkinin doğası ve anlamlılığı kapsamında yapılan Pearson korelasyon analizinin sonuçları Tablo 2'de verilmiştir. Korelasyon analizi sonucu HKİ ile Ortalama araç sayısı ve Minimum hız arasında anlamlı bir ilişkinin olduğu görülmektedir (Tablo 2). Analiz sonuçları değerlendirildiğinde bir noktadan geçen araç sayısı ne kadar fazla ve trafik yoğunluğu nedeniyle araç hızları ne kadar düşüğe HKİ değeri o noktada o kadar yüksektir.

Korelasyon analizinden sonra 11 istasyon için 1 Ocak – 31 Aralık 2020 tarihleri için 2020 yılı HKİ-Trafik yoğunluğuna ait HKİ, ortalama araç sayısı ve hız verileri Şekil 4 (sol)'da CBS ortamında Ekim ayı için görselleştirilmiştir. Şekil 4 (sağ)'da ise Aksaray istasyonuna ait örnek sensör veri analiz grafikleri görselleştirilmiştir. Üstten alta sırası ile yıllık, aylık ve ekim ayına ait günlük HKİ-Trafik yoğunluğu verileri görselleştirilmiştir.



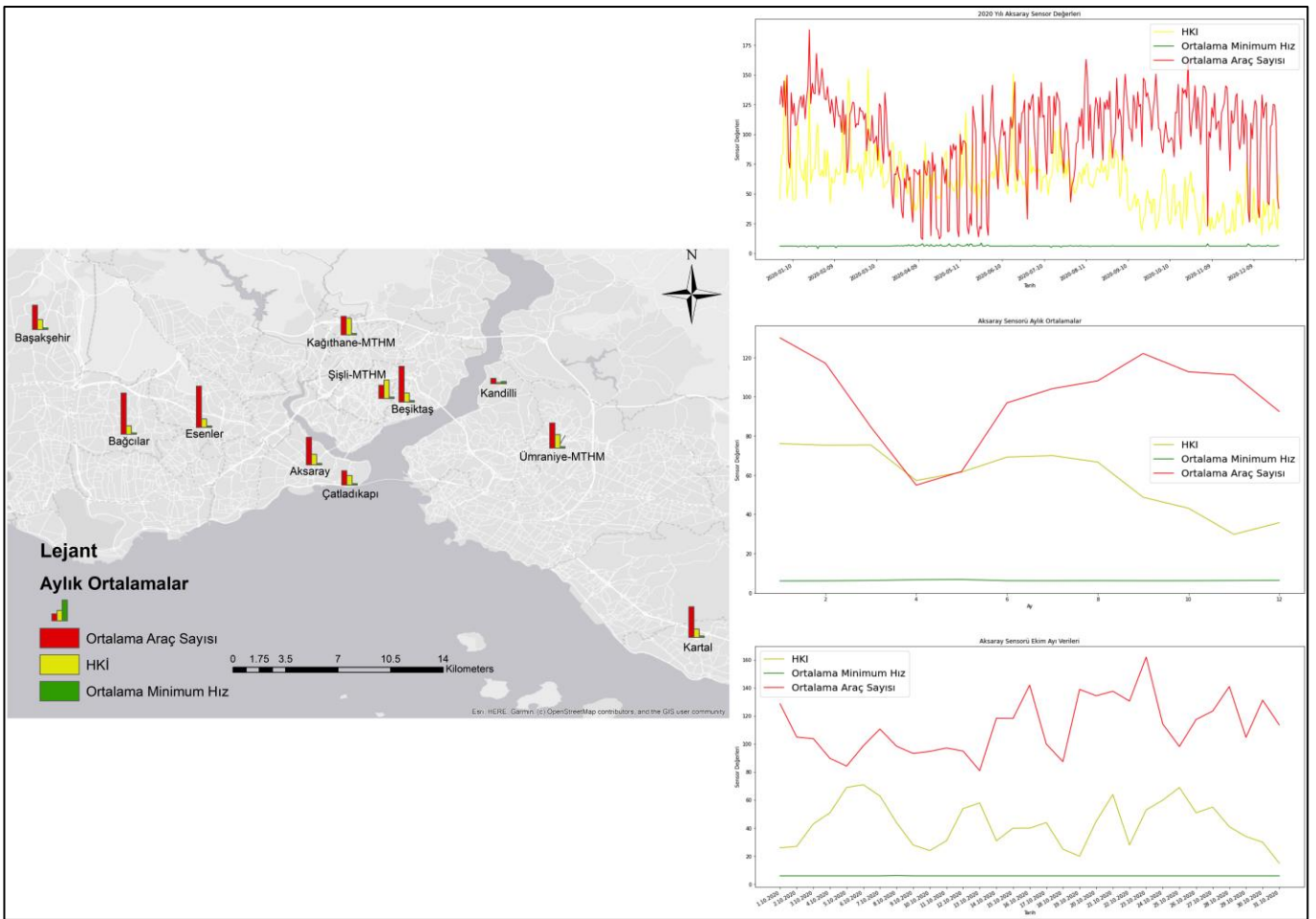
Şekil 3. PostgreSQL / PostGIS sorgu görselleri

**Tablo 2.** Pearson Korelasyon Analizi Sonuçları

		MaxHKİ	MinHiz	MaksHiz	OrtHiz	OrtAracSay
<b>MaxHKİ</b>	Pearson Correlation	1	-,701**	,452**	,039	,613**
	Sig. (2-tailed)		,000	,000	,666	,000
	N	122	122	122	122	122
<b>MinHiz</b>	Pearson Correlation	-,701**	1	-,276**	,194*	-,760**
	Sig. (2-tailed)	,000		,002	,032	,000
	N	122	122	122	122	122
<b>MaksHiz</b>	Pearson Correlation	,452**	-,276**	1	,846**	,461**
	Sig. (2-tailed)	,000	,002		,000	,000
	N	122	122	122	122	122
<b>OrtHiz</b>	Pearson Correlation	,039	,194*	,846**	1	,040
	Sig. (2-tailed)	,666	,032	,000		,661
	N	122	122	122	122	122
<b>OrtAracSay</b>	Pearson Correlation	,613**	-,760**	,461**	,040	1
	Sig. (2-tailed)	,000	,000	,000	,661	
	N	122	122	122	122	122

\*Correlation is significant at the 0.05 level (2-tailed)

\*\*Correlation is significant at the 0.01 level (2-tailed)

**Şekil 4.** Ekim Ayı HKİ-Trafik Yoğunluğu Ortalamaları (sol) Aksaray İstasyonu Örnek Sensör Veri Analiz Grafikleri (sağ)

## 5. Sonuç ve değerlendirme

Şehirlerin nüfusunun hızla artması kentsel ısı adalarının oluşmasına, trafik ve kaynak yönetimi gibi sıkıntılarının artmasına sebep olmuştur. Bu sorunlara çözüm getirebilmek ve yaşam kalitesini artırmak için akıllı şehir uygulamaları kapsamında farklı büyük veri altlıklarından temin edilen ve şehir yaşam kalitesini etkileyen parametrelerin gerçek zamanlı analizi önem arz etmektedir. Akıllı şehir yönetiminde iklimden hava kalitesine, ulaşımdan sosyo-ekonomik parametrelere

kadar büyük veri altlıklarının açık veri portalları ile erişilebilir olması gerekmektedir. Çalışmada bahsedildiği gibi yapılanmış veri yönetiminde kullanılan geleneksel ilişkisel veritabanı ortamlarının yanısıra, yapılanmamış büyük verinin yönetimi için NoSQL olarak ifade edilen veri tabanı ortamlarının kullanımı gerekmektedir.

Gerçekleştirilen örnek uygulamada görüldüğü gibi hava izleme ve trafik izleme sensörlerinin sağladığı veriler akıllı şehir uygulamaları için kullanıma açılmıştır. Bu çalışmada hava izleme sensörlerinin kaydettiği veriler kullanılarak MongoDB ortamında hava kalitesi

endeksi hesaplanmıştır. Yakın konumdaki trafik sensörlerinden elde edilen veriler de kullanılabilir hale getirilmiştir. Sonuç veri analitik raporları, PostGIS ortamında hava izleme istasyonlarının geometrileri ile birleştirilip QGIS ortamına aktarılmıştır. Tamamen açık-kaynaklı yazılım ortamlarında gerçekleştirilen uygulamalarda, trafik yoğunluğu noktalarından da ortalama hız ve araç sayısı değerleri hesaplanmış ve bu değerler karşılaştırılarak trafik yoğunluğunun hava kalitesine etkisi olduğu belirlenmiştir. Gerçekleştirilecek akıllı şehir gösterge paneli uygulamalarında, gerçek zamanlı trafik yoğunluğu ve hava kalitesi ilişkisi analiz edilebilir, risk durumuna göre trafik yönlendirmesi yapılabilir. Bu yaklaşımla, örnek teknolojiler kullanılarak şehir yaşam kalitesini ve yerel yönetimlerin hizmet potansiyelini artıran birçok akıllı şehir uygulaması geliştirilebilir.

### Araştırmacıların katkı oranı

**Arif Çağdaş Aydınoglu:** Düzenleme, Metodoloji, Kontrol; **Rabia Bovkır:** Literatür Çalışması, Veri Analizi, Makale Yazımı; **Muzaffer Bulut:** Veri Analizi ve Uygulama Geliştirme

### Çatışma beyanı

Herhangi bir çıkar çatışması bulunmamaktadır.

### Kaynakça

Accenture (2015). Building and managing an intelligent cities in India.

Agarwal S & Rajan K S (2016). Performance Analysis of MongoDB versus PostGIS/PostgreSQL Databases for Line Intersection and Point Containment Spatial Queries. *Spat. Inf. Res.*, 24:671–677. <https://doi.org/10.1007/s41324-016-0059-1>

Agrahari A & Rao D (2017). Big Data: Technologies, Tools and Trends. *International Research Journal of Engineering and Technology (IRJET)*, 4(10):640-649.

Aktan E (2018). Büyük Veri: Uygulama Alanları, Analitiği ve Güvenlik Boyutu. *Bilgi Yönetimi Dergisi*, 1(1), 1-22.

Albino V, Berardi U & Dangelico R M (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology*, 22(1):3–21.

Al Nuaimi E, Al Neyadi H, Mohamed N & Al-Jaroodi J (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, 6(1), 25. <https://doi.org/10.1186/s13174-015-0041-5>

Anbaroğlu B (2021). Spatial Query Performance Analyses on a Big Taxi Trip Origin–Destination Dataset. In: Mobasher A. (eds) *Open Source Geospatial Science for Urban Studies. Lecture Notes in Intelligent Transportation and Infrastructure*. Springer, Cham. [https://doi.org/10.1007/978-3-030-58232-6\\_3](https://doi.org/10.1007/978-3-030-58232-6_3)

Aydın G (2017). NoSQL Veri Tabanları Üzerinde Bir Metin Madenciliği Uygulaması [Yüksek Lisans Tezi]. İstanbul Aydın Üniversitesi.

Aydınoglu A Ç, Çölkesen İ, Şenbil M, Bovkır R, Yomralıoğlu T (2020). 116Y204 nolu TÜBİTAK Projesi Sonuç Raporu. TÜBİTAK ÇAYDAG, Ankara, 04.2020

BalaAnand M, Karthikeyan N, Karthik S & Sivaparthipan C B (2017). A Survey on BigData with Various V's on comparison of Apache Hadoop and Apache Spark. *Advances in Natural and Applied Sciences*, 11(4):362-369.

Ballas D (2013). What Makes a 'Happy City'? *Cities* 32(1), 39–50.

Baralis E, Dalla Valle A, Garza P, Rossi C & Scullino F (2017). SQL versus NoSQL databases for geospatial applications. 2017 IEEE International Conference on Big Data (Big Data), 3388-3397. <https://doi.org/10.1109/BigData.2017.825834>

Bilici Z & Babahanoğlu V (2018). Akıllı Kent Uygulamaları ve Konya Örneği. *Akademik Yaklaşımlar Dergisi*, 9 (2), 124-139.

Čerešňák R & Kvet M (2019). Comparison of Query Performance in Relational a Non-Relation Databases. *Transportation Research Procedia*, 40, 170-177. <https://doi.org/10.1016/j.trpro.2019.07.027>

Chourabi H, Nam T, Walker S, Gil-Garcia J R, Mellouli S & Nahon K (2012). Understanding Smart City initiatives: An integrative framework. 45. Hawaii international conference on system sciences, 2289-2297, 4–7 Ocak, Maui, HI.

Correia L M & Wüstel K (2011). Smart Cities applications and requirements. White Paper of the Experts Working Group, NetWorks European Technology Platform.

Cui L & Shi J (2020). Urbanization and its environmental effects in Shanghai, China. *Urban Climate*, 2: 1-15. <https://doi.org/10.1016/j.uclim.2012.10.008>

Çetin M & Demirci O K (2016). Erzincan'da Doğal Gaz Kullanımının Hava Kalitesine Etkisi. *Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 9(1):8-18.

Daşdemir Y & Kara B C (2019). Farklı İş Yükleri Altında NoSQL Sistemlerinin Performans Analizi. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 8(4), 1466-1477. <https://doi.org/10.17798/bitlisfen.547532>

Demchenko Y & Membrey P (2014). Defining Architecture Components of the Big Data Ecosystem. *International Conference on Collaboration Technologies and Systems (CTS)*, Minneapolis, MN, USA:104-112, <https://doi.org/10.1109/CTS.2014.6867550>

Demirarslan K O & Akıncı H (2018). CBS ve Hava Kalitesi Verileri Kullanılarak Marmara Bölgesinin Kış Sezonunda Hava Kalitesinin Değerlendirilmesi. *Doğ Afet Çevre Derg.* 4(1):11-27. <https://doi.org/10.21324/dacd.344564>

Demirci Y & Köseoğlu Ö (2018). Akıllı Şehirler ve Yerel Sorunların Çözümünde Yenilikçi Teknolojilerin Kullanımı. 4(2), 40-57.

Doğan K & Arslantekin S (2016). Büyük veri: Önemi, Yapısı ve Günümüzdeki Durum. *Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Dergisi- DTCF Dergisi*, 56(1), 15-36. [https://doi.org/10.1501/Dtcfder\\_0000001461](https://doi.org/10.1501/Dtcfder_0000001461)

Duarte L & Teodoro A C (2021). GIS Open-Source Plugins Development: A 10-Year Bibliometric Analysis on Scientific Literature. *Geomatics* 1(2), 206-245.

Egghe L & Leydesdorff L (2009). The relation between Pearson's correlation coefficient r and Salton's cosine

- measure. *Journal of the American Society for information Science and Technology*, 60(5):1027-1036.
- Firican G (2017). The 10 Vs of Big Data. <https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx>. Erişim: 10.04.2021.
- George G, Haas M R & Pentland A (2014). Big Data and Management. *Academy of Management Journal*, 57(2), 321-326. <https://doi.org/10.5465/amj.2014.4002>
- Ghazal A, Rabl T, Hu M, Raab F, Poess M, Crolotte A & Jacobsen H A (2013). BigBench: Towards an Industry Standard Benchmark for Big Data Analytics. SIGMOD'13: Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data, 1197-1208.
- Goodchild M F (2021). GIS in the Era of Big Data. *European Journal of Geography*.
- Gökşen Y & Aşan H (2015). Veri Büyüklüklerinin Veritabanı Yönetim Sistemlerinde Meydana Getirdiği Değişim: NOSQL. *Bilişim Teknolojileri Dergisi*, 8(3). 125-131 <https://doi.org/10.17671/btd.52374>
- Hashem I A T, Yaqoob I, Anuar N B, Mokhtar S, Gani A & Khan S U (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47:98-115. <https://doi.org/10.1016/j.is.2014.07.006>.
- Harrison G (2015). *Next Generation Databases*. Springer Science+Business Media, New York, ABD.
- Hollands R G (2008). Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? *City*, 12(3):303-319.
- Inter-American Development Bank (IDB) (2016). *The Road toward Smart Cities: Migrating from Traditional City Management to the Smart City*.
- Jin X, Wah B W, Cheng X, Wang Y (2015). Significance and Challenges of Big Data Research. *Big Data Research*, 2, 59-64.
- JRC (European Commission-Joint Research Centre) (2014). *Scientific Pilot Study on Big Data Analysis*.
- Khan M A, Uddin M F, Gupta N (2014). Seven V's of Big Data: Understanding Big Data to extract Value. Proceedings of 2014 Zone 1 Conference of the American Society for Engineering Education.
- Khanduja V, Arora A & Garg S (2017). Applications of big data in real world: It's not what you know. It's what you do with what you know. 2017 International Conference on Computing, Communication and Automation (ICCCA), 159-163. <https://doi.org/10.1109/CCAA.2017.8229792>
- Landmark Solutions (2015). *The 7 pillars of Big Data. White Papers*. Erişim Adresi: [https://www.landmark.solutions/portals/0/lmsdocs/whitepapers/the\\_7\\_pillars\\_of\\_big\\_data\\_whitepaper.pdf](https://www.landmark.solutions/portals/0/lmsdocs/whitepapers/the_7_pillars_of_big_data_whitepaper.pdf). Erişim Tarihi: 19.02.2021
- Li S, Dragicevic S, Castro F A, Sester M, Winter S, Coltekin A, ... & Cheng T (2016). Geospatial big data handling theory and methods: A review and research challenges, *ISPRS Journal of Photogrammetry and Remote Sensing*, 115:119-133. <https://doi.org/10.1016/j.isprsjprs.2015.10.012>.
- Lu J & Cao X (2015). PM2.5 pollution in major cities in China: Pollution status, emission sources and control measures. *Fresenius Environ. Bull.*, 24:1338-1349.
- Lu Y, Xie H, Arshika Zain S, Xu Z (2019). Geographic Information Systems and Big Data Driven Framework for Planning and Design of Smart Cities. 2019 4th International Conference on Information Systems Engineering (ICISE), 6-10. <https://doi.org/10.1109/ICISE.2019.00009>
- McAfee A & Brynjolfsson E (2012). *Big Data's Management revolution*. Harvard Business Review.
- McKinsey Global Institute (2011). *Big Data: The next frontier for innovation, competition and productivity*. McKinsey Global Institute, 1-6.
- Microsoft (2013). *The big bang: how the big data explosion is changing the world*. Erişim Adresi: <https://news.microsoft.com/2013/02/11/the-big-bang-how-the-big-data-explosion-is-changing-the-world/>. Erişim Tarihi: 16.04.2021.
- Ming Z, Luo C, Gao W, Han R, Yang Q, Wang L, Zhan J (2014). BDGS: A Scalable Big Data Generator Suite in Big Data Benchmarking, in: *Information Security and Privacy. Information Security and Privacy*, 138-154.
- Mirabelli M C, Ebel S, Damon S A (2020). Air Quality Index and air quality awareness among adults in the United States. *Environmental Research*, 183, 109185.
- Morcali M H & Akan D S (2017). Monitoring and Determination of Air Pollution Sources in Kahramanmaraş. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 20(2), 105-115. <https://doi.org/10.17780/ksujes.310602>
- Nam T A & Pardo T (2011). Conceptualizing Smart City With Dimensions Of Technology, People, and Institutions. Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, 284-285.
- Neirotti P, De Marco A, Cagliano A, Mangano G, Scorrano F (2014). Current trends in Smart City initiatives: Some stylised facts. *Cities*, 38, 25-36.
- NESSI (2012). *Big Data White Paper*. Erişim Adresi: [http://www.nessi-europe.com/Files/Private/NESSI\\_WhitePaper\\_BigData.pdf](http://www.nessi-europe.com/Files/Private/NESSI_WhitePaper_BigData.pdf). Erişim Tarihi: 16.04.2021.
- O'Grady M & O'Hare G (2012). How Smart Is Your City? *Science* 335(3), 1581-1582.
- Onay A (2020). Büyük Veri Çağında İç Denetimin Dönüşümü. *Muhasebe Bilim Dünyası Dergisi*, 22(1):127-163. <https://doi.org/10.31460/mbdd.620837>
- Oracle (2017). <https://www.oracle.com/big-data/guide/what-is-big-data.html>. Erişim: 13.04.2021.
- Örselli E & Akbay C (2019). Teknoloji ve Kent Yaşamında Dönüşüm: Akıllı Kentler. *Uluslararası Yönetim Akademisi Dergisi*, 2(1):228-241. <https://doi.org/10.33712/mana.544549>
- Özşahin E, Eroğlu İ, Pektezel H (2016). Keşan'da (Edirne) Hava Kirliliği. *Selçuk Ün. Sos. Bil. Ens. Der.* 36, 83-100.
- Öztürk S & Atmaca H E (2017). İlişkisel ve İlişkisel Olmayan (NoSQL) Veri Tabanı Sistemleri Mimari Performansının Yönetim Bilişim Sistemleri Kapsamında İncelenmesi. *Bilişim Teknolojileri*

- Dergisi, 199-199.  
<https://doi.org/10.17671/gazibtd.309303>
- Perkins L, Redmond E & Wilson J R (2018). Seven Databases in Seven Weeks. The Pragmatic Programmers, LLC, ABD.
- Pokorny J (2013). NoSQL databases: a step to database scalability in web environment. *IJWIS* 9(1):69-82.
- Rusu O, Halcu I, Grigoriu O, Neculoiu G, Sandulescu V, Marinescu M & Marinescu V (2013). Converting unstructured and semi-structured data into knowledge. 11th RoEduNet International Conference, Sinaia, 1-4.
- Sabrin S, Karimi M, Fahad Md G R & Nazari R (2020). Quantifying environmental and social vulnerability: Role of urban Heat Island and air quality, a case study of Camden, NJ. *Urban Climate*, 34, 100699.
- Schober P, Boer C, Schwarte L A (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, 126(5):1763-1768. <https://doi.org/10.1213/ANE.0000000000002864>
- Schönberger V M & Cukier K (2013). Büyük Veri- Yaşama, Çalışma ve Düşünme Şeklimizi Dönüştürecek Bir Devrim. Çev. Banu Erol. İstanbul: Paloma Yayınevi.
- Sensuse D I, Cahyaningsih E & Wibowo W C (2015). Identifying Knowledge Management Process of Indonesian Government Human Capital Management using Analytical Hierarchy Process and Pearson Correlation Analysis. *Procedia Computer Science*, 72:233-243.
- Sevli O (2016). A Platform for Processing Non-Structural Data by Big Data Analysis Methods and Associate It With Structural Data: Developing a Recommendation Engine based on Social Media. PhD Thesis, Süleyman Demirel University, Graduate School of Applied and Natural Sciences.
- Şahin E K, Bovkir R & Aydinoglu A Ç (2020). Yeni Teknolojik Gelişmelerin Coğrafi Bilgi Sistemlerine Etkisi. *Harita Dergisi*, 1-16.
- Tan X, Han L, Zhang X, Zhou W, Li W & Qian Y (2021). A review of current air quality indexes and improvements under the multi-contaminant air pollution exposure. *Journal of Environmental Management* 279, 111681.
- Taleb I, Serhani M A, Dssouli R (2018). Big Data Quality Assessment Model for Unstructured Data. 13th International Conference on Innovations in Information Technology (IIT), 69-74.
- T.C. Çevre ve Şehircilik Bakanlığı (ÇŞB), Çevre Yönetimi Genel Müdürlüğü, ÇED, İzin ve Denetim Müdürlüğü (2017). 1. Ulusal Hava Kalitesi Yönetimi Çalıştay Sonuç Bildirgesi ve Raporu.
- TechAmerica Foundation (2012). Demystifying big data: A practical guide to transforming the business of Government. TechAmerica Foundation: Federal Big Data Commission Report.
- Terzi F & Ocaççı M (2017). Kentlerin Geleceği: Akıllı Kentler. *İTÜ Vakfı Dergisi*, 77:10-13.
- Urban Sustainability Directors Network (USDN), (2015). Getting Smart About Smart Cities. USDN Innovation Fund.
- URL 1: <https://historicjournals.itu.int/viewer/699/?return=1&cssname=include#page=17&viewer=picture&o=&n=0&q=> (Erişim Tarihi: 25 Haziran 2021)
- URL 2: <http://hadoop.apache.org/> (Erişim Tarihi: 24 Haziran 2021)
- URL-3: <https://www.statista.com/markets/411/topic/446/demographics/#insights> (Erişim Tarihi: 24 Haziran 2021)
- URL-4: <https://www.mongodb.com/> (Erişim Tarihi: 24 Haziran 2021)
- URL-5: <https://sim.csb.gov.tr/Home/HKI?baslik=HAVZA%20%C4%B0ZLEME%20S%C4%B0STEM%C4%B0> (Erişim Tarihi: 24 Haziran 2021)
- US EPA (2018). Technical Assistance Document for the Reporting of Daily Air Quality AQI Index. U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.
- Viceconti M, Hunter P & Hose R (2015). Big Data, Big Knowledge: Big Data for Personalized Healthcare. *IEEE Journal Of Biomedical And Health Informatics*, 19(4):1209-1215.
- Washburn D, Sindhu U, Balaouras S, Dines R A, Hayes N M & Nelson L E (2010). Helping CIOs understand “Smart City” initiatives: Defining the Smart City, its drivers, and the role of the CIO. Cambridge, MA: Forrester Research, Inc.
- Wei D, Wang B, Lin G, Liu D, Dong Z, Liu H, Liu Y (2017). Research on Unstructured Text Data Mining and Fault Classification Based on RNN-LSTM with Malfunction Inspection Report. *Energies*, 10(406). <https://doi.org/10.3390/en10030406>
- White M (2012). Digital workplaces: vision and reality. *Bus. Inf. Rev.*, 29(4):205-214.
- WHO (2021). Official website of WHO. Erişim Adresi: <https://www.who.int/news-room/air-pollution>. Erişim Tarihi: 02.04.2021.
- Yin S & Kaynak O (2015). Big Data for Modern Industry: Challenges and Trends. *Point of View*, 103(2):143-146.
- Yue P & Jiang L (2014). BigGIS: How big data can shape next-generation GIS. The Third International Conference on Agro-Geoinformatics, 1-6. <https://10.1109/AgroGeoinformatics.2014.6910649>

