

T.C.
GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÖNETİM VE BİLİŞİM SİSTEMLERİ ANA BİLİM DALI

**BULUT BİLİŞİM ALANINDA YAPILAN ÇALIŞMALARIN LATENT
DIRICHLET ALLOCATION (LDA) MODELİ İLE ANALİZİ**

YÜKSEK LİSANS

Emrah ELBASDI

HAZİRAN-2025
GÜMÜŞHANE



**T.C.
GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÖNETİM VE BİLİŞİM SİSTEMLERİ ANA BİLİM DALI

**BULUT BİLİŞİM ALANINDA YAPILAN ÇALIŞMALARIN LATENT
DIRICHLET ALLOCATION (LDA) MODELİ İLE ANALİZİ**

**ANALYSIS OF STUDIES IN THE FIELD OF CLOUD COMPUTING USING
THE LATENT DIRICHLET ALLOCATION (LDA) MODEL**

YÜKSEK LİSANS

Emrah ELBASDI

**HAZİRAN -2025
GÜMÜŞHANE**



**T.C.
GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÖNETİM VE BİLİŞİM SİSTEMLERİ ANA BİLİM DALI

**BULUT BİLİŞİM ALANINDA YAPILAN ÇALIŞMALARIN LATENT
DIRICHLET ALLOCATION (LDA) MODELİ İLE ANALİZİ**

**ANALYSIS OF STUDIES IN THE FIELD OF CLOUD COMPUTING USING
THE LATENT DIRICHLET ALLOCATION (LDA) MODEL**

YÜKSEK LİSANS

Emrah ELBASDI

Danışman: Prof. Dr. Handan ÇAM

**HAZİRAN-2025
GÜMÜŞHANE**

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Yüksek Lisans Tezi olarak hazırlamış olduğum “**Bulut Bilişim Alanında Yapılan Çalışmaların Latent Dirichlet Allocation (LDA) Modeli İle Analizi**” isimli bu tezimin, tamamen kendi çalışmam olduğunu, her alıntıya kaynak gösterdiğimi, alıntı yaptığım tüm çalışmaları kaynakçada belirttiğimi ve Gümüşhane Üniversitesi’nin lisanslı kullanıcısı olduğum intihal yazılım programı ile Lisansüstü Eğitim Enstitüsü’nün belirlediği kıstaslara uygun olarak raporladığımı taahhüt ederim. Tezimin kâğıt ve elektronik kopyalarının Gümüşhane Üniversitesi Lisansüstü Eğitim Enstitüsü arşivinde saklanmasına izin verdiğimi onaylarım.

Lisansüstü Eğitim ve Öğretim Yönetmeliği’nin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

18/06/2025

Emrah ELBASDI

TEŞEKKÜR

Bu tez çalışmasının her aşamasında bilgi ve tecrübesiyle bana yol gösteren, cömertçe paylaştığı rehberlik ve akademik birikimiyle bu çalışmanın başarısında kilit rol oynayan tez danışmanım Prof. Dr. Handan ÇAM'a en derin saygılarımı ve teşekkürlerimi sunarım. Kendisinin bilimsel rehberliği, düşünce dünyamda eşsiz bir vizyon kazanmamı sağlamıştır.

Araştırmamın tamamlanmasında sunduğu kaynaklar ve desteklerle önemli bir rol oynayan Gümüşhane Üniversitesi yönetimine, destek personeline ve bu çalışmanın gerçekleşmesine katkı sağlayan tüm akademisyenlere teşekkür ederim. Üniversitenin sağladığı akademik ortam ve araştırma olanakları, bu çalışmanın niteliğini önemli ölçüde artırmıştır.

Son olarak, bu tez sürecinde sabır, anlayış ve teşvikleriyle her zaman yanımda olan aileme gönülden teşekkür ediyorum. Onların manevi desteği, bu uzun ve zorlu akademik yolculuğun en değerli parçası olmuştur.

Emrah ELBASDI
GÜMÜŞHANE – 2025

ÖZET

Bulut bilişim (BB), iş dünyası, akademik araştırmalar ve bireysel kullanım alanlarında kritik bir öneme sahiptir. Bu tez çalışmasında, BB literatüründe öne çıkan temaların belirlenmesi ve alanın zamansal eğilimlerinin analiz edilmesi amaçlanmıştır. Bu doğrultuda, Scopus veri tabanında 2014–2024 yılları arasında yayımlanan 18.689 makaleden oluşan kapsamlı bir veri seti kullanılmıştır.

Araştırmanın metodolojisi, öncelikle veri toplama ve ön işleme süreçlerini içermekte, ardından metin madenciliği ve konu modelleme aşamalarını kapsamaktadır. Latent Dirichlet Allocation (LDA) algoritması kullanılarak gerçekleştirilen analizlerde, bulut bilişim literatüründeki tematik yapılar, kavramsal kümelenmeler ve eğilim değişimleri detaylı biçimde ortaya konmuştur.

Çalışmanın bulguları, BB alanındaki baskın temaların (örneğin, yeni teknolojiler, dağıtık uygulamalar) yanı sıra henüz gelişmekte olan stratejik araştırma alanlarını (örneğin, sağlık sistemleri, blockchain çözümleri) da tespit etmeye olanak sağlamıştır. Ayrıca yıllara göre değişen tematik yoğunluklar analiz edilerek, BB'nin hangi yönlerde evrildiği ortaya çıkarılmıştır.

Bu tez çalışmasının, BB alanında içerik temelli bir kavramsal haritalama sunarak, mevcut literatürdeki boşlukların tespitine ve gelecekteki akademik çalışmalara yön verme potansiyeline katkı sağladığı düşünülmektedir.

Anahtar Kelimeler: Bulut bilişim, Konu modelleme, LDA, Metin madenciliği, Bibliyometrik analiz

SUMMARY

Cloud computing plays a critical role across business sectors, academic research, and personal applications. This thesis aims to identify prominent themes in the cloud computing literature and to analyze temporal trends within the field. For this purpose, a comprehensive dataset comprising 18,689 articles published between 2014 and 2024 was retrieved from the Scopus database.

The methodology of the study encompasses data collection, preprocessing, and subsequently, text mining and topic modeling processes. Analyses were conducted using the Latent Dirichlet Allocation (LDA) algorithm to uncover thematic structures, conceptual clusters, and the evolution of research trends in cloud computing.

The findings reveal dominant research themes such as emerging technologies and distributed applications, alongside the identification of strategic and developing areas like healthcare systems and blockchain-based solutions. Additionally, the analysis of thematic intensities across years illustrates the evolving directions of cloud computing research.

By offering a content-driven conceptual mapping of the field, this thesis contributes to identifying current research gaps and provides strategic guidance for future academic studies in cloud computing.

Keywords: Cloud computing, Topic modeling, LDA, Text mining, Bibliometric analysis

İÇİNDEKİLER

KABUL VE ONAY	III
BİLİMSEL ETİĞE UYGUNLUK BEYANI	IV
TEŞEKKÜR	V
ÖZET	VI
SUMMARY	VII
İÇİNDEKİLER	VIII
TABLolar DİZİNİ	XI
ŞEKİLLER DİZİNİ	XII
SİMGELER VE KISALTMALAR DİZİNİ	XIII
1.GİRİŞ	1
2.BULUT BİLİŞİM VE KONU MODELLEMESİ	3
2.1.Kavramsal Çerçeve ve Analitik Altyapı	3
2.2.Bulut Bilişim Nedir?	3
2.3.Bulut Bilişimin Tarihçesi	4
2.4.Bulut Bilişim Mimarisi	6
2.4.1.Servis Modelleri	7
2.4.2.Dağıtım Modelleri	11
2.5.Bulut Bilişimin Farklı Sektörlerdeki Etkileri	14
2.6.Bulut Bilişimin Özellikleri	16
2.6.1.Çoklu Kiracılık (Kaynak Paylaşımı)	16
2.6.2.Etkili Ölçeklenebilirlik	17
2.6.3.Esneklik	17
2.6.4.Kullandığın Kadar Öde Modeli	18
2.6.5.Kaynakların Kendi Kendini Tedarik Etmesi	18
2.7.Bulut Bilişimde Güvenlik ve Veri Gizliliği	19
2.7.1.Paylaşılan Sorumluluk Modeli	20
2.7.2.Verİ Şİfreleme ve Koruma	20
2.7.3.Kimlik ve Erişim Yönetimi (IAM)	20
2.7.4.Uyumluluk ve Yasal Zorluklar	21
2.7.5.Yeni Nesil Tehditler	21
2.8.Bulut Bilişimde Güvenlik Önlemleri ve Stratejiler	21
2.9.Bulut Bilişim Sistemlerinin Avantajları ve Dezavantajları	23

2.9.1.Bulut Bilişim Sistemlerinin Avantajları	23
2.9.2.Bulut Bilişim Sistemlerinin Dezavantajları	25
2.10.Bulut Bilişimin Geleceği ve Yeni Teknolojik Eğilimler	27
2.11.Metin Madenciliği ve Konu Modellemesi	31
2.11.1.Metin Madenciliği	31
2.11.2.Metin Madenciliğinin Uygulama Alanları	32
2.11.3.Metin Madenciliğinin Aşamaları	33
2.12.Konu Modelleme (Topic Modelling)	36
2.12.1.Konu Modelleme: Tanımı, Önemi ve Uygulama Alanları	36
2.12.2.Konu Modelleme Yöntemleri	37
3.LATENT DIRICHLET ALLOCATION (LDA) MODELİ VE BULUT BİLİŞİM YAYINLARININ ANALİZİNE YÖNELİK LİTERATÜR ÖZETİ	44
3.1.Akademik Yayınların Konu Modellemesinde LDA'nın Rolü	44
3.2.LDA Modeli ve Genel Yapısı	45
3.3.LDA'nın Zaman İçindeki Evrimi	47
3.4.LDA Modelinin Temel Bileşenleri	49
3.5.LDA Modelinin Avantajları ve Kısıtları	52
3.6.Bulut Bilişim Literatüründe Yöntemsel İnceleme	55
3.7.Literatürdeki Önemli Çalışmalar:	57
4.LDA MODELLEME SÜRECİ VE TEMATİK YAPININ BELİRLENMESİ	45
4.1.LDA Tabanlı Yöntemsel Yaklaşım ve Analiz Süreci	45
4.2.Veritoplama	46
4.2.1.Verit Kaynağı ve Formatı	46
4.2.2.Verit Setinin Oluşturulması	46
4.3.Ön İşleme	47
4.3.1.Kütüphaneleri Kullanımı	47
4.3.2.Verit Temizleme	48
4.3.3.Stopwords Listesi	49
4.3.4.Tokenizasyon ve N-gram Modelleri	49
4.3.5.Bigram ve Trigram Modellerinin Oluşturulması	51
4.3.6.Stopwords ve Boş Cümlelerin Kaldırılması	52
4.3.7.Korpus ve Sözlük Oluşturma	53
4.4.LDA Modelleme Süreci	54
4.4.1.LDA Modelinin Eğitilmesi	54
4.4.2.LDA Modeli Sonuçları ve Görselleştirilmesi	55

4.5.Bulut Bilişim Literatüründe Tematik ve İstatistiksel Eğilimler	56
4.5.1.Yıllara Göre Makale Sayısının Analizi	56
4.5.2.En Çok Yayın Yapan Dergilerin Analizi	57
4.5.3.Bulut Bilişim Araştırmalarının Coğrafi Dağılımı	58
4.5.4.BB Yayınlarının LDA ile Analizi	59
4.5.5.LDA Modelinden Türetilen Genel Kelime Bulutu	60
4.5.6.LDA Tabanlı Tematik Yapılar ve Frekans Analizi	62
4.5.7.BB Konuları için Kelime Kümesi Analizi	64
4.5.8.Konu Dağılımları ve Trend Analizi	65
4.5.9.BB Taksonomi Görselleştirmesi	67
5.SONUÇLAR VE DEĞERLENDİRME	69
KAYNAKÇA	79
ÖZGEÇMİŞ	100

TABLÖLAR DİZİNİ

Tablo 1. Bulut bilişim servis modellerinin karşılaştırılması (Younis vd., 2024'ten derlenmiştir).	8
Tablo 2. Bulut bilişim dağıtım modellerinin avantaj ve dezavantajları (Goyal, 2014'ten derlenmiştir)	13
Tablo 3. Bulut bilişimin farklı sektörlerdeki avantajları ve zorlukları (Birden çok kaynaktan derlenmiştir).	14
Tablo 4. Konu modelleme yöntemlerinin özellikleri ve sınırlamaları (George ve Birla, 2018, derlenmiştir).	38
Tablo 5. LDA yönteminin farklı alanlardaki akademik uygulamaları (Birden fazla kaynaktan derlenmiştir).	48
Tablo 6. LDA modelinin avantaj ve dezavantajları (birden fazla kaynaktan derlenmiştir).	53
Tablo 7. Bulut bilişim literatüründe kullanılan analiz yaklaşımları.....	56
Tablo 8. Literatür özeti (birden fazla kaynaktan derlenmiştir).	64
Tablo 9. Veri toplama sürecinde uygulanan kriterler.....	46
Tablo 10. Analiz sürecinde kullanılan python kütüphaneleri ve işlevleri.....	47
Tablo 11. Stopwords kategorileri ve örnekler.....	49
Tablo 12. LDA model parametreleri ve işlevleri.	55
Tablo 13. LDA modelinden türetilen 20 ana konu ve anahtar kelimeleri.....	62
Tablo 14. Bulut bilişim konularının yıllık frekans değerleri ve trend skorları	67

ŞEKİLLER DİZİNİ

Şekil 1. Bulut bilişim mimarisi (Yiğit ve Ayata, 2024).	7
Şekil 2. Hizmet modelleri (IaaS, PaaS, SaaS) ve bileşenleri (Younis 2024).	10
Şekil 3. Bulut bilişim dağıtım modelleri (Butt vd., 2020).	11
Şekil 4. Metin madenciliği sürecinde veri işleme ve yapılandırma adımları (Şeker, 2015).	34
Şekil 5. LDA için grafiksel model (Genç vd., 2024).	46
Şekil 6. 2014–2024 yılları arasında BB konulu yayın sayılarının yıllık dağılımı.	57
Şekil 7. 2014–2024 yılları arasında BB alanında en çok yayın yapan dergiler	58
Şekil 8. 2014–2024 yılları arasında Bulut bilişim alanında en çok yayın yapan 10 ülkenin coğrafi dağılımı.	59
Şekil 9. LDA modeline dayalı tematik dağılım haritası (Intertopic Distance Map) ve en baskın anahtar kelimeler grafiği (Top-30 Most Salient Terms)	60
Şekil 10. LDA modelinden elde edilen temalara dayalı genel kelime bulutu.	61
Şekil 11. 2010-2014 yılları arasında yayınlanan bulut bilişim konularının frekans dağılımı.	63
Şekil 12. Bulut bilişim literatüründe LDA tabanlı tematik kelime kümeleme ağı.	65
Şekil 13. Bulut bilişim araştırma konularının yıllık eğilim analizi (2014–2024)	66
Şekil 14. Bulut bilişim temalarının taksonomik dağılımı	68

SİMGELER VE KISALTMALAR DİZİNİ

AI	: Yapay Zekâ (Artificial Intelligence)
API	: Uygulama Programlama Arayüzü (Application Programming Interface)
AWS	: Amazon Web Services (Amazon Web Hizmetleri)
BB	: Bulut Bilişim (Cloud Computing)
BoW	: Bag of Words (Kelime Torbası)
CI/CD	: Sürekli Entegrasyon / Sürekli Dağıtım (Continuous Integration / Continuous Deployment)
CRM	: Müşteri İlişkileri Yönetimi (Customer Relationship Management)
CSV	: Comma-Separated Values (Virgülle Ayrılmış Değerler)
CTM	: İlişkili Konu Modeli (Correlated Topic Model)
DL	: Derin Öğrenme (Deep Learning)
ESG	: Environmental, Social and Governance (Çevresel, Sosyal ve Yönetişim)
FinOps	: Bulut Finans Operasyonları (Financial Operations)
GDPR	: Genel Veri Koruma Tüzüğü (General Data Protection Regulation)
HDP	: Hiyerarşik Dirichlet Süreci (Hierarchical Dirichlet Process)
HIPAA	: Sağlık Sigortası Taşınabilirlik ve Sorumluluk Yasası
IAM	: Kimlik ve Erişim Yönetimi (Identity and Access Management)
IaaS	: Hizmet Olarak Altyapı (Infrastructure as a Service)
IEEE	: IEEE Xplore (Elektrik ve Bilgisayar Mühendisliği Yayın Veritabanı)
IoT	: Nesnelerin İnterneti (Internet of Things)
IT	: Bilgi Teknolojisi (Information Technology)
LDA	: Gizli Dirichlet Dağılımı (Latent Dirichlet Allocation)
LMS	: Learning Management System (Öğrenme Yönetim Sistemi)
MFA	: Çok Faktörlü Kimlik Doğrulama (Multi-Factor Authentication)
MCMC	: Markov Chain Monte Carlo (Markov Zinciri Monte Carlo Yöntemi)
ML	: Makine Öğrenmesi (Machine Learning)
NIST	: National Institute of Standards and Technology
NLTK	: Natural Language Toolkit (Doğal Dil İşleme Araç Takımı)
NLP	: Doğal Dil İşleme (Natural Language Processing)
OPA	: Açık Politika Aracısı (Open Policy Agent)
OpEx	: Operational Expenditure (Operasyonel Harcama)

PaaS	: Hizmet Olarak Platform (Platform as a Service)
PCI DSS	: Payment Card Industry Data Security Standard (Ödeme Kartı Endüstrisi Veri Güvenliği Standardı)
PLSA	: Olasılıksal Gizli Anlamsal Analiz (Probabilistic Latent Semantic Analysis)
RBAC	: Rol Tabanlı Erişim Kontrolü (Role-Based Access Control)
SaaS	: Hizmet Olarak Yazılım (Software as a Service)
SLA	: Hizmet Seviyesi Anlaşması (Service Level Agreement)
SVD	: Tekil Değer Ayrıştırması (Singular Value Decomposition)
UBYS	: Üniversite Bilgi Yönetim Sistemi
VPN	: Sanal Özel Ağ (Virtual Private Network)
v.b.	: Ve benzeri
v.d.	: Ve diğerleri/ve devamı
y.y.	: Yayıncı bilinmiyor



1. GİRİŞ

Teknolojinin hızlı gelişimi, sosyal, ekonomik ve kurumsal süreçlerde önemli değişimlere yol açmıştır. Dijitalleşmenin etkisiyle, bireyler ve işletmeler bilgiye erişim, veri işleme ve hizmet sunma biçimlerini köklü şekilde dönüştürmüştür. Bu dönüşümün merkezinde yer alan bulut bilişim (BB), esnekliği, maliyet etkinliği ve ölçeklenebilir yapısıyla modern teknoloji altyapısının temel bileşeni hâline gelmiştir.

BB, yalnızca veri depolama ya da işlem kapasitesi sunan bir teknolojik yenilik değil; aynı zamanda dijital dönüşüm süreçlerinin kritik bir bileşeni olarak değerlendirilmektedir. Büyük veri analitiği, yapay zeka (AI) uygulamaları ve uzaktan çalışma sistemleri gibi günümüzün stratejik dijital dönüşüm alanları, doğrudan BB altyapısına dayanarak geliştirilmektedir.

Özellikle sağlık, eğitim ve finans gibi sektörlerde BB, veri güvenliğinden hizmet erişimine kadar geniş bir yelpazede yenilikçi çözümler sunmaktadır. Örneğin, sağlık alanında hasta verilerinin hızlı erişimi ve güvenli analizi mümkün hâle gelirken; eğitimde uzaktan eğitim platformlarının geliştirilmesine olanak sağlamıştır. Bununla birlikte, BB'nin hızlı gelişimi, güvenlik açıkları, veri gizliliği ihlalleri ve platform bağımlılığı gibi çeşitli sorunları da beraberinde getirmiştir. Bu nedenle, BB'nin sunduğu fırsatlar kadar oluşturduğu yeni risklerin de sistematik olarak ele alınması gereklidir.

Bu çalışmanın temel amacı, BB alanında gerçekleştirilen akademik yayınları Latent Dirichlet Allocation (LDA) modeli ile analiz ederek, literatürdeki temel temaları, kavramsal kümeleri ve zamansal eğilimleri sistematik biçimde ortaya koymaktır. LDA, metin madenciliği alanında geniş veri kümelerinde tematik yapıyı ortaya çıkarmada güçlü bir yöntem olarak kabul edilmekte olup, çoklu konu analizi yapabilmesi ve ölçeklenebilirliği sayesinde bu çalışmada tercih edilmiştir.

Bu kapsamda, BB literatürüne yönelik klasik bibliyometrik analizlerin ötesine geçilerek, içerik temelli bir modelleme gerçekleştirilmiş; literatürdeki mevcut tematik boşluklar, kavramsal yoğunluklar ve gelişen araştırma alanları detaylı biçimde haritalandırılmıştır. Mevcut literatürde BB özelinde konu modelleme yöntemlerinin kullanıldığı çalışmaların sınırlı olması, bu çalışmanın özgün katkısını oluşturmaktadır. Böylece, BB alanında konu modelleme tabanlı yeni nesil analizlerin uygulanabilirliğine yönelik önemli bir metodolojik çerçeve sunulmaktadır.

Çalışma dört ana bölümden oluşmaktadır. Birinci bölümde BB kavramı ve dijital dönüşüm sürecindeki rolü ele alınmış; ikinci bölümde konu modelleme ve LDA

yönteminin teorik temelleri açıklanmıştır. Üçüncü bölümde yöntem ve analiz süreci detaylandırılmış, dördüncü bölümde ise elde edilen bulgular değerlendirilmiş ve literatüre katkıları tartışılmıştır. Bu yönüyle, çalışmanın sonuçlarının hem BB alanındaki akademik gelişime hem de konu modelleme tabanlı metin madenciliği uygulamalarına özgün katkılar sunması hedeflenmektedir.



2. BULUT BİLİŞİM VE KONU MODELLEMESİ

2.1. Kavramsal Çerçeve ve Analitik Altyapı

Bu bölümde, çalışmanın teorik temelini oluşturan Bulut Bilişim (BB), konu modelleme ve metin madenciliği kavramları ele alınacaktır. BB'nin temel hizmet (IaaS, PaaS, SaaS) ve dağıtım (kamu, özel, topluluk, hibrit) modelleri açıklanarak, veri güvenliği, gizlilik ve erişim kontrolü gibi kritik konulara değinilecektir. Ardından, konu modelleme kapsamında Latent Dirichlet Allocation (LDA) yöntemine giriş yapılacak ve bu yöntemin BB literatürünü analiz etmedeki rolü tartışılacaktır. Son olarak, metin madenciliği çerçevesinde büyük veri ortamlarında bilgi çıkarımı süreçleri özetlenecektir. Bu teorik çerçeve, ilerleyen bölümlerde yapılacak uygulamalara zemin oluşturacaktır.

2.2. Bulut Bilişim Nedir?

Bilgi teknolojileri alanında devrim niteliğinde bir gelişme olarak kabul edilen BB, donanım ve yazılım kaynaklarını sanallaştırarak internet üzerinden esnek ve ölçeklenebilir bir hizmet modeli sunmaktadır. Bu teknoloji, kullanıcıların yalnızca ihtiyaç duydukları kadar kaynak tüketmesine ve buna bağlı olarak kullandıkları kadar ödeme yapmasına imkân sağlayarak maliyet etkin çözümler üretmektedir (Mell ve Grance, 2011).

ABD Ulusal Standartlar ve Teknoloji Enstitüsü (NIST, 2011), BB'yi şu şekilde tanımlamaktadır: “Bulut bilişim, her yerden, isteğe bağlı olarak, bir ağ üzerinden kolayca erişilebilen bilgi işlem kaynaklarının (örneğin, ağlar, sunucular, depolama birimleri, uygulamalar ve hizmetler) paylaşılmasını mümkün kılan bir modeldir.” Bu tanım, BB'nin sunduğu temel özellikleri, özellikle erişim kolaylığı ve ölçeklenebilirlik avantajlarını açıkça ortaya koymaktadır.

İşletmelerin tüm verilerini fiziksel depolama birimleri yerine sanal sunucularda saklamasına ve internet üzerinden bu verilere erişmesine olanak tanıyan BB sistemleri, yazılım ve bilgi kaynaklarını merkezi bir yapıda toplayarak bilgisayarlar ve diğer dijital cihazlar aracılığıyla kesintisiz veri paylaşımını mümkün kılar. Aynı zamanda, elektrik dağıtım sistemlerine benzer bir mantıkla çalışan bu teknoloji, bilgi akışını düzenli ve erişilebilir hale getirir (Göv ve Erdoğan, 2020, s. 302).

Literatürde BB, farklı açılardan değerlendirilmiştir. Buyya vd. (2009), bu teknolojiyi kullanıcı taleplerine göre ölçeklenebilen ve ekonomik avantajlar sunan bir model olarak tanımlamış ve özellikle maliyet etkinliği ile esnekliği vurgulamıştır. Armbrust vd. (2010) ise BB sistemlerinin işletmeler için operasyonel maliyetleri önemli ölçüde düşürme ve geniş kapsamlı bilgi işlem kaynaklarına erişim sağlama potansiyeline dikkat çekmiştir.

Ayrıca, Falah vd. (2021), yalnızca veri depolama aracı olmanın ötesinde, geliştiricilere yenilikçi fırsatlar sunan bir platform olan BB'nin IoT uygulamaları için kritik bir altyapı sunduğunu ve büyük veri yönetimini kolaylaştırdığını belirtmektedir. Bu doğrultuda, büyük hacimli verilerin etkin işlenmesi ve yönetilmesi için bulut tabanlı çözümler giderek daha fazla önem kazanmaktadır.

Hamutoğlu (2018), BB'nin eğitimde erişim imkânlarını genişleterek sürdürülebilir öğrenme süreçlerini desteklediğini ifade etmektedir. BB altyapıları, öğretim süreçlerini iyileştirirken eğitimde fırsat eşitliğini de teşvik etmektedir.

Diğer yandan, Aydın (2022), BB teknolojilerinin e-ticaret sektöründe işletmeler için kritik bir öneme sahip olduğunu belirtmiştir. Bu sistemlerin, işletmelere verilerini güvenli bir şekilde depolama imkânı sunduğunu ve siber saldırılara karşı koruma sağlayarak dayanıklılığı artırdığını ifade etmiştir.

Sonuç olarak BB, yalnızca bir bilgi işlem modeli olarak değil, aynı zamanda veri yönetiminde sağladığı verimlilik, ekonomik avantajlar, erişim kolaylığı ve inovasyon kapasitesiyle dijital dönüşümün temel taşlarından biri olarak değerlendirilmektedir. Eğitimden e-ticarete, IoT uygulamalarından büyük veri analitiğine kadar uzanan geniş bir yelpazede sunduğu çözümlerle, BB modern toplumların teknolojik altyapısında vazgeçilmez bir model haline gelmiştir. Bu doğrultuda, BB'nin tanımı ve temel özellikleri incelendikten sonra, bir sonraki bölümde bu teknolojinin tarihsel gelişim süreci detaylı şekilde ele alınacaktır.

2.3. Bulut Bilişimin Tarihçesi

Bilgisayar kaynaklarının bir hizmet olarak sunulması fikriyle ortaya çıkan BB, ilk olarak 1961 yılında MIT'de John McCarthy tarafından şu sözlerle dile getirilmiştir: “Bilgisayarlar bir gün, tıpkı telefon sisteminin bir kamu hizmeti olarak düzenlenmesi gibi organize edilebilir” (Surbiryala ve Rong, 2019). McCarthy'nin bu öngörüsü, günümüzde BB'nin sunduğu hizmet odaklı mimariyle büyük ölçüde örtüşmektedir.

Öte yandan, Mell ve Grance (2011) BB'yi, her yerden erişilebilir, paylaşımlı ve ölçeklenebilir bir bilişim modeli olarak tanımlamaktadır. Modern BB altyapıları,

kullanıcıların yalnızca ihtiyaç duydukları kadar kaynak tüketmesini ve bu sayede verimli bir bilgi işlem süreci yürütmesini sağlamaktadır.

1970’li yıllarda, geliştirdiği dağıtık iletişim sistemleri ile günümüzde ağ tabanlı hizmetlerin gelişimi açısından önemli bir temel oluşturan ARPANET (Advanced Research Projects Agency Network), internetin erken dönem altyapısını oluşturarak veri iletişiminde devrim yaratmıştır (Leiner vd., 2009).

1980’lerde, fiziksel bir sunucunun birden fazla sanal sunucuya dönüştürülmesine olanak tanıyarak, bilgi işlem kaynaklarının daha etkin ve verimli bir şekilde kullanılmasını sağlayan sanallaştırma (virtualization) modelleri BB teknolojisinin gelişiminde temel bir rol oynamıştır. IBM’in geliştirdiği “utility computing” (kullanıma dayalı bilgi işlem) modeli, kullanıcıların yalnızca kullandıkları kaynaklar için ödeme yapmalarını sağlayarak esnek bir bilgi işlem ortamı sunmuştur (Buyya vd., 2009). Bu yaklaşım, günümüzde BB’nin maliyet etkinliği ve ölçeklenebilirlik özelliklerini belirleyen temel unsurlardan biri olmuştur.

1990’lı yıllar, internetin yaygınlaşması ve servis tabanlı modellerin gelişmesiyle BB kavramının daha somut hale gelmesine katkı sağlamıştır (Sultan, 2014). Bu dönemde, işletmeler için yazılım hizmetlerini internet üzerinden sunmayı amaçlayan yeni modeller ortaya çıkmıştır. Özellikle 1999 yılında Salesforce tarafından sunulan ilk Software as a Service (SaaS) uygulaması, bulut tabanlı hizmetlerin ticari kullanımına önemli bir örnek teşkil etmiştir (Irshad, 2014). Bu gelişme, geleneksel yazılım lisanslama yöntemlerinden farklı olarak, yazılımların kullanımına dayalı bir modelle sunulmasına olanak tanımıştır.

2000’li yıllarda BB, giderek genişleyen bir kullanıcı kitlesine hizmet veren bir model olarak öne çıkmaya başlamıştır. Buyya, Yeo ve Venugopal (2009), bu teknolojinin geleceğini şekillendirirken, onun “*hesaplama alanında beşinci temel kamu hizmeti*” (su, elektrik, gaz ve telefonla eşdeğer) olarak ele alınması gerektiğini savunmuş ve bu görüşlerini akademik çalışmalarında detaylandırmıştır (s.600). BB’nin esneklik, ölçeklenebilirlik ve maliyet etkinliği açısından sunduğu avantajlar, işletmelerin bilgi işlem altyapılarını dönüştürmesine olanak sağlamıştır.

Bu perspektiften hareketle, Amazon Web Services (AWS) gibi küresel sağlayıcılar, BB’nin ticarileşme sürecine öncülük etmiştir. AWS’nin 2006’da başlattığı Elastic Compute Cloud (EC2) hizmeti, dinamik olarak ölçeklenebilen sunucu kapasitesiyle işletmelerin esnek ihtiyaçlarını karşılayarak bulut altyapısının benimsenmesinde kritik bir dönüm noktası oluşturmuştur (Voorsluys, Broberg ve Buyya, 2011). EC2’nin sunduğu kullandıkça öde (pay-as-you-go) modeli, BB

hizmetlerini küçük ölçekli girişimlerden büyük kurumsal yapılara kadar geniş bir kullanıcı kitlesi için erişilebilir hale getirmiştir.

Günümüzde ise BB, yalnızca bir bilgi işlem modeli olmanın ötesine geçerek, büyük veri analitiği, Nesnelerin İnterneti (IoT), yapay zeka ve kenar bilişim gibi ileri teknolojilerle bütünleşik bir yapı haline gelmiştir. Bu entegrasyon, veri işleme süreçlerinin daha hızlı ve verimli hale gelmesini sağlarken, özellikle gerçek zamanlı analiz ve karar destek sistemlerinde önemli avantajlar sunmaktadır (Prasad, Puthal ve Gupta, 2017). IoT cihazlarından toplanan büyük veri, yapay zekâ destekli analitik sistemler tarafından işlenerek anlamlı bilgilere dönüştürülmekte ve BB altyapısı sayesinde merkezi veya dağıtık sistemler üzerinde etkin şekilde yönetilmektedir.

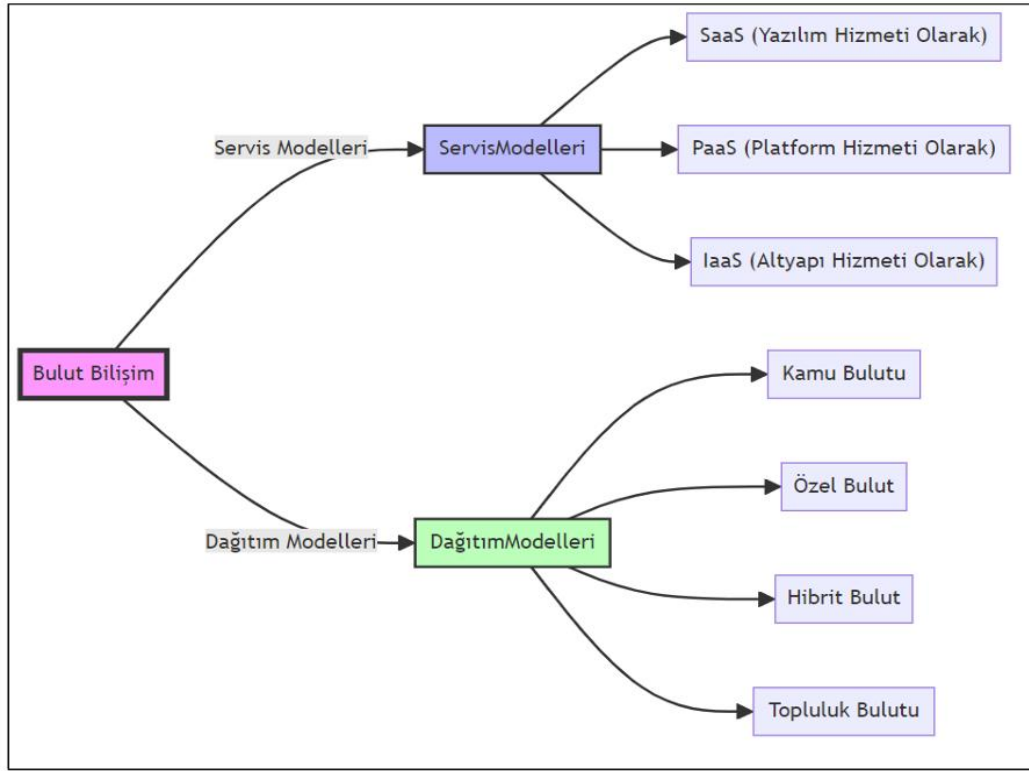
Bu teknolojik dönüşüm, BB'nin günümüzde eğitim, sağlık, üretim ve finans gibi farklı sektörlerde geniş çapta benimsenmesini sağlamıştır. Örneğin, akıllı şehir uygulamalarında IoT sensörleri aracılığıyla veri toplanarak, BB tabanlı yapay zekâ sistemleri tarafından işlenmekte ve şehir yönetimlerine optimizasyon imkânı sunmaktadır (Minh vd., 2021). Benzer şekilde, sağlık sektöründe hasta verileri BB üzerinde depolanıp analiz edilerek kişiselleştirilmiş tıbbi çözümler geliştirilmektedir. Tüm bu gelişmeler, BB'nin modern teknolojilerin vazgeçilmez bir bileşeni haline gelmesine katkıda bulunmuştur.

Tarihsel gelişim süreci, BB'nin günümüzdeki mimari yapısının şekillenmesinde kritik bir rol oynamıştır. Bu bağlamda, bir sonraki bölümde BB mimarisinin bileşenleri ve işleyişi detaylandırılacaktır.

2.4. Bulut Bilişim Mimarisi

BB mimarisi, kullanıcıların internet aracılığıyla altyapı (IaaS), platform (PaaS) ve yazılım (SaaS) hizmetlerinden faydalanmasını sağlayan bir sistemdir. Bu yapı, fiziksel donanım gereksinimini azaltırken, esneklik, ölçeklenebilirlik ve erişim kolaylığı gibi avantajlar sunmaktadır (Mell ve Grance, 2011; Yiğit ve Ayata, 2024). Ancak, BB aynı zamanda güvenlik tehditleri de içermektedir ve veri ihlalleri, yetkisiz erişim gibi riskler yönetilmelidir (Yiğit ve Ayata, 2024).

BB mimarisi, kullanıcı cihazları ile bulut hizmet sağlayıcıları arasındaki etkileşimi organize eder. Şekil 1, bu yapının hem servis modelleri (IaaS, PaaS, SaaS) hem de dağıtım modelleri (Kamu, Özel, Hibrit, Topluluk) üzerinden nasıl çalıştığını görselleştirmektedir. (Yiğit ve Ayata, 2024).



Şekil 1. Bulut bilişim mimarisi (Yiğit ve Ayata, 2024).

Bu mimari, modern bilgi teknolojilerinin temelini oluşturarak veri transferi, kaynak yönetimi ve hizmet sunumu gibi kritik alanlarda güçlü bir altyapı sağlar. Yiğit ve Ayata (2024), BB'nin servis ve dağıtım modellerini bir arada sunarak, bireylerden kurumsal kullanıcılara kadar geniş bir kitleye hitap ettiğini ve dijital dönüşüm süreçlerinde işletmelerin stratejik gereksinimlerini karşıladığını ifade etmektedir.

BB'nin yapısı, yalnızca kaynakların paylaşımıyla sınırlı kalmayıp, uygulama geliştirme, veri analizi ve güvenlik süreçlerinde de önemli bir rol oynamaktadır. Özellikle hibrit bulut modeli, kamu ve özel bulutların avantajlarını birleştirerek hassas verilerin korunmasını sağlarken, daha az kritik işlemler için ekonomik çözümler sunmaktadır (Sultan, 2014; Yiğit ve Ayata, 2024).

BB mimarisi, servis modelleri ve dağıtım modelleri olarak iki ana unsurdan oluşur. Bu unsurlar, mimarinin işlevselliğini artırarak kullanıcılar için özelleştirilmiş çözümler sunar.

2.4.1. Servis Modelleri

BB servis modelleri, kullanıcıların farklı ihtiyaçlarını karşılayacak şekilde Yazılım Hizmeti (SaaS), Platform Hizmeti (PaaS) ve Altyapı Hizmeti (IaaS) olmak üzere üç temel kategoriye ayrılmaktadır. Tablo 1, bu modellerin temel özelliklerini ve kullanım alanlarını detaylı bir şekilde karşılaştırmaktadır. Şekil 2 ise, her bir bulut

modeli için kullanıcı ve sağlayıcı arasındaki sorumluluk dağılımını açıkça göstermektedir (Younis vd., 2024).

Tablo 1. Bulut bilişim servis modellerinin karşılaştırılması (Younis vd., 2024'ten derlenmiştir).¹

Kategori	IaaS	PaaS	SaaS
Tanım	Sanallaştırılmış altyapı kaynakları (ör. Sunucular, depolama) internet üzerinden sağlanır.	Altyapıyı yönetmeden uygulama geliştirme, test etme ve dağıtma platformu.	Yazılım uygulamaları internet üzerinden sunulur ve abonelikte erişilebilir.
Temel Özellikler	Sanal makineler, depolama, ağ yönetimi, self-servis yönetimi.	Geliştirme araçları, entegre ortamlar, veri tabanları ve ara yazılımlar.	Kullanıma hazır yazılımlar, çoklu kiracı mimarisi, otomatik güncellemeler.
Sağlayıcılar	AWS, Microsoft Azure, Google Cloud.	Google App Engine, Microsoft Azure, Heroku.	Salesforce, Google Workspace, Microsoft 365.
Yönetim	Kullanıcı sanal makineleri, depolamayı ve ağ yönetir; altyapı sağlayıcı tarafından yönetilir.	Sağlayıcı altyapıyı yönetir; kullanıcı uygulamaları yönetir.	Sağlayıcı her şeyi, yazılım güncellemelerini de dahil olmak üzere, yönetir.
Kullanıcı Kontrolü	Altyapı ve ağ kaynakları üzerinde yüksek kontrol.	Orta düzeyde kontrol, uygulama geliştirme odaklı.	Düşük kontrol, yazılım tamamen sağlayıcı tarafından yönetilir.
Özelleştirme Seviyesi	Yüksek: Kullanıcılar altyapıyı kontrol eder ve ihtiyaçlarına göre yapılandırır.	Kullanıcılar uygulama geliştirmeye odaklanırken altyapıyı sağlayıcı yönetir.	Düşük: Kullanıcılar yazılımı sağlayıcının sunduğu şekilde kullanır.
Maliyet Yapısı	Kullanılan altyapı kaynakları için kullandıkça öde.	Platform kullanımı ve geliştirme kaynakları için ödeme yapılır.	Yazılım hizmetleri için abonelik tabanlı fiyatlandırma.

¹ Younis, R., Iqbal, M., Munir, K., ve Javed, M. A. (2024). A comprehensive analysis of Cloud service models: IaaS, PaaS, and SaaS in the context of emerging technologies and trend. *IEEE Transactions on Computers*. <https://doi.org/10.1109/ICECCE63537.2024.10823401>

Tablo 1. (Devamı)

Kategori	IaaS	PaaS	SaaS
Genel Kullanım Alanları	Web sitelerinin barındırılması, kurumsal uygulamaların çalıştırılması, büyük veri işleme.	Web uygulamaları, mobil uygulamalar ve mikro hizmetlerin geliştirilmesi ve dağıtımı.	E-posta hizmetleri, ofis verimlilik araçları, müşteri ilişkileri yönetimi (CRM).
Avantajları	Altyapı üzerinde tam kontrol, esnek ölçekleme, maliyet optimizasyonu.	Daha hızlı geliştirme döngüleri, kolaylaştırılmış uygulama yönetimi ve ölçeklenebilirlik.	Kullanım kolaylığı, altyapı bakımı gerektirmez, her cihazdan erişilebilirlik.
Dezavantajları	Altyapıyı yönetmek için IT uzmanlığı gerektirir.	Altyapı üzerinde sınırlı kontrol, sağlayıcının platformuna bağımlılık.	Sınırlı özelleştirme, çoklu kiracı mimarisi nedeniyle veri güvenliği endişesi.

BB servis modelleri, farklı kullanım senaryolarına yönelik sunmuş olduğu avantajlar Tablo 1’de gösterilmiştir. *IaaS modeli*, fiziksel donanım gereksinimi olmaksızın sanal sunucular, depolama alanları ve ağ bileşenlerine erişim imkânı sağladığından büyük veri analitiği, web barındırma ve yüksek ölçekli bilgi işlem süreçlerinde yaygın olarak tercih edilmektedir (Gandhi ve Kumbharana, 2014).

WS Elastic Compute Cloud (EC2), Microsoft Azure sanal makineler ve Google Cloud Compute Engine gibi hizmet sağlayıcıları, IaaS modelinin sunduğu esnek kaynak kullanımına ve ölçeklenebilir yapıya örnek olarak sunulabilir. Bu hizmetler sayesinde kullanıcılar, fiziksel donanım altyapısına yatırım yapmadan yüksek işlem gücüne erişebilmektedir. Ancak, IaaS kullanıcılarının altyapı yönetiminde aktif rol oynaması, belirli teknik bilgi düzeyini gerekli kılmakta ve bu durum modelin önemli sınırlılıklarından biri olarak karşımıza çıkmaktadır (Mathur, 2024).

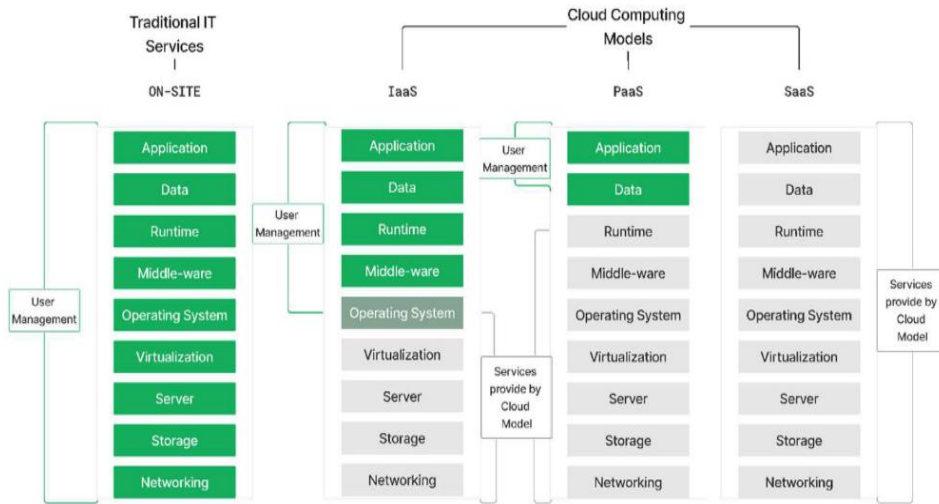
Uygulama geliştirme, test etme ve dağıtma süreçlerinde yazılımcıların ihtiyaç duydukları araç ve platformları sunan *PaaS modeli*, hızlı geliştirme süreçleri ve entegrasyon kolaylığı gibi avantajlar sağlayarak yazılım geliştirme maliyetlerini önemli ölçüde düşürmekte ve yazılım süreçlerinin hızlanmasında önemli bir rol oynamaktadır (Varma ve Choi, 2016).

Örneğin, Google App Engine ve Microsoft Azure App Services, altyapı yönetimini ortadan kaldırarak geliştiricilere entegre araçlarla hızlı geliştirme imkânı tanımaktadır (Bajaj vd., 2020). Ancak, Tablo 1’de gösterildiği üzere platform

kontrolünün sağlayıcıda olması ve sağlayıcılar arası özellik farkları nedeniyle yazılım geliştiricilerin esnekliği sınırlanmakta, ayrıca farklı PaaS'ler arasında geçişte veri taşınabilirliği sorunları yaşanabilmektedir (Ravi, 2017).

SaaS modeli, kullanıcıların yazılımlara internet üzerinden erişmesini sağlayarak yerel kurulum ve bakım ihtiyacını ortadan kaldırmaktadır (Bilur ve Salunkhe, 2023). Özellikle son kullanıcılar ve küçük işletmeler için ideal bir çözüm sunan bu modelde, Google Workspace, Salesforce ve Microsoft 365 gibi SaaS çözümleri yaygın olarak kullanılmaktadır.

Ancak, SaaS sağlayıcılarının merkezi sunucularda veri saklama stratejisi, güvenlik ve gizlilik açısından endişelere yol açabilmektedir (Jia vd., 2021). Çoklu kiracı mimarisi, maliyetleri optimize ederken veri güvenliği ve gizlilik endişelerini de beraberinde getirmektedir. Ayrıca, Şekil 2'de gösterilen SaaS modelinde kullanıcıların yönetim rolü minimum düzeyde olup tüm sorumluluk sağlayıcıya aittir.



Şekil 2. Hizmet modelleri (IaaS, PaaS, SaaS) ve bileşenleri (Younis 2024).

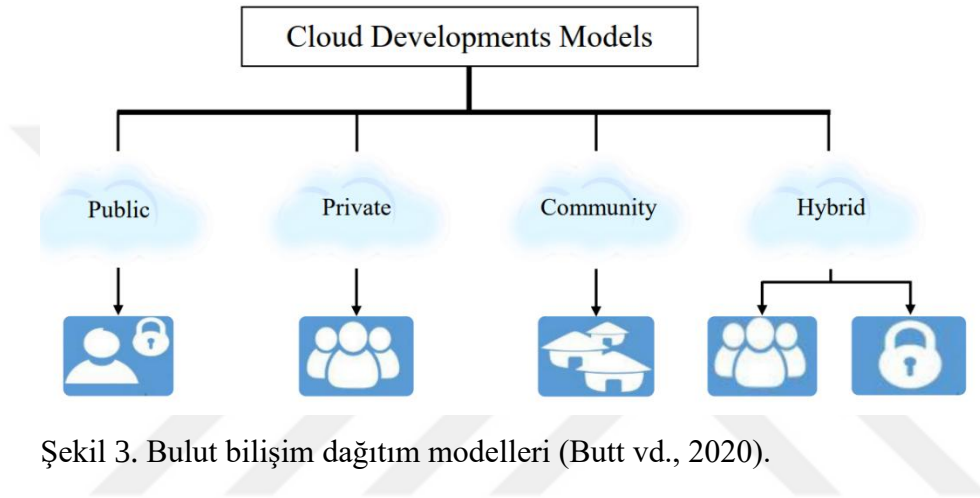
Gelecekte, AI ve makine öğrenimi (ML) entegrasyonu, BB servislerinin önemini daha da artıracaktır. Özellikle SaaS platformları, AI algoritmalarını kullanarak kullanıcı deneyimlerini kişiselleştirme potansiyeline sahiptir ve bu da bireysel ve kurumsal kullanımda önemli fırsatlar sunmaktadır (Gandhi ve Jain, 2025).

Şekil 2'de gösterildiği gibi, servis modelleri arasındaki kullanıcı ve sağlayıcı sorumluluk dağılımı, AI destekli otomasyon süreçleriyle daha esnek hale gelebilir. Örneğin, SaaS modelinde sağlayıcının üstlendiği yönetim sorumluluğu, AI algoritmalarının katkısıyla daha verimli hale getirilebilir. Ayrıca, büyük veri teknolojilerinin bulut modelleriyle entegrasyonu, bu servislerin farklı sektörlerde daha geniş kullanım alanlarına ulaşmasını sağlayacaktır.

Servis modellerinin karşılaştırmalı analizi, BB'nin esnekliğini ve sektörel uyumunu anlamada önemli bir adımdır. Bu bilgiler ışığında, bir sonraki alt başlıkta dağıtım modellerinin avantaj ve dezavantajları incelenecektir.

2.4.2. Dağıtım Modelleri

BB sistemleri, kullanıcıların ihtiyaçlarına göre şekillendirilebilen farklı dağıtım modelleri sunmaktadır. Şekil 3'te görüldüğü gibi, bu modeller Özel Bulut, Kamu Bulut, Hibrit Bulut ve Topluluk Bulut olarak dört ana kategoriye ayrılmaktadır (Kothapalli, 2018).



Şekil 3. Bulut bilişim dağıtım modelleri (Butt vd., 2020).

Özel Bulutlar, yalnızca belirli bir organizasyona veya kullanıcı grubuna tahsis edilen bir modeldir. Özellikle yüksek güvenlik ve özelleştirme gerektiren durumlar için ideal olan bu sistemler, sağlık sektörü gibi hassas veri içeren alanlarda, hasta bilgilerini düzenlemelere uygun bir şekilde depolamak için yaygın olarak kullanılmaktadır (Kratzke, 2018; Rimal vd., 2011).

Örneğin, bir sağlık kuruluşu, hasta verilerini saklamak için özel bir bulut altyapısını tercih ederek veri gizliliğini güvence altına alabilir. Ayrıca, finans sektörü müşteri bilgilerinin korunması ve kritik işlemlerin güvenli bir şekilde yönetimi için; savunma sektörü ise operasyonel verilerin gizliliğini sağlamak amacıyla özel bulut çözümlerine yönelmektedir. Tablo 2'de görüldüğü gibi, Özel Bulutlar, yüksek veri güvenliği ve özelleştirme gereksinimlerini karşılamak için ideal bir çözüm sunarken, yüksek maliyet ve teknik uzmanlık gereksinimi gibi dezavantajlara da sahiptir.

Kamu Bulutu, genel kullanıma açık bir altyapı sunarak her yerden erişim kolaylığı sağlar. “Kullandıkça öde” modeli sayesinde küçük ve orta ölçekli işletmeler için maliyet avantajı yaratmaktadır (Buyya vd., 2010).

Birçok start-up, başlangıç aşamasında büyük yatırımlar yapmak yerine, kamu bulut hizmetlerini tercih ederek düşük maliyetli ve esnek altyapılardan yararlanmaktadır. Özellikle Amazon Web Services (AWS) gibi sağlayıcılar, start-up'lara ölçeklenebilir çözümler sunarak, maliyet tasarrufu sağlamalarına ve enerjilerini ürün geliştirme ve kullanıcı deneyimini iyileştirme gibi kritik iş süreçlerine odaklamalarına imkân tanımaktadır.

Bu doğrultuda, Abell ve arkadaşları (2021), AWS'nin düşük maliyet, ölçeklenebilirlik ve kolay erişim gibi özelliklerle Asya-Pasifik bölgesindeki start-up'ların büyümesine büyük katkı sağladığını vurgulamaktadır. Ancak, verilerin üçüncü taraf bir sağlayıcı tarafından yönetilmesi güvenlik açısından risk teşkil edebilir (Buyya vd., 2009).

Tablo 2'de de vurgulandığı gibi, *Kamu Bulutu* modeli, düşük başlangıç maliyetleri ve yüksek ölçeklenebilirlik avantajı sağlarken, veri güvenliği ve kontrol eksikliği gibi önemli riskler içermektedir.

Hibrit Bulut, Özel ve Kamu Bulut altyapılarının bir arada kullanıldığı bir model olup, kurumlara esneklik ve maliyet kontrolü açısından dengeli bir çözüm sunmaktadır. Bu yapı sayesinde şirketler, hassas verilerini özel bulut üzerinde saklarken; daha az kritik işlemleri genel bulut üzerinden yürüterek kaynaklarını verimli biçimde kullanabilmektedir (Dokuz ve Çelik, 2017). Örneğin, bir e-ticaret firması müşteri verilerini özel bulutta korurken, kampanya dönemlerinde genel bulutun sağladığı esnekliği kullanarak trafik yoğunluğunu yönetebilir.

Ancak hibrit mimarinin uygulanması bazı teknik ve operasyonel zorlukları da beraberinde getirir. Özellikle veri tutarlılığı, sistem uyumluluğu ve hizmetlerin entegrasyonu gibi konular hibrit yapının karmaşıklığını artıran temel etkenler arasında yer almaktadır. Ayrıca, veri transferlerinde yaşanan gecikmeler (latency), merkezi yönetim eksikliği ve güvenlik politikalarının bütünleşik şekilde yönetilmesi gibi faktörler, güçlü entegrasyon stratejileri ve araçları gerektirmektedir (Shrivastava ve Agrawal, 2024).

Topluluk Bulutları, belirli bir kullanıcı grubu veya sektöre yönelik olarak oluşturulan, altyapı ve kaynak paylaşımına dayalı BB modelidir. Bu model; üniversiteler, hastaneler, belediyeler veya benzer iş alanlarına sahip kuruluşların ortak ihtiyaçlarına yanıt vermek üzere yapılandırılır. Özellikle kaynak paylaşımını teşvik eden ve maliyeti düşürmeye yönelik uygulamalarda etkin bir çözüm sunmaktadır (Kropczynski vd., 2024).

Topluluk Bulutları, yalnızca akademik alanlarda değil, aynı zamanda akıllı şehir projelerinde de yaygın biçimde kullanılmaktadır. Bu altyapılar; trafik akışı, enerji tüketimi ve çevresel veriler gibi bilgilerin güvenli ve uyumlu biçimde paylaşılmasını sağlayarak, kamu hizmetlerinin dijitalleşmesini ve veriye dayalı yönetim sistemlerinin oluşumunu desteklemektedir (Hagerbaumer vd., 2024).

Ancak bu modelin etkinliği, yüksek düzeyde iş birliği ve ortak yönetim gerektirir. Geniş kapsamlı projelerde, farklı kurumların aynı altyapıyı kullanması; veri yönetimi, güvenlik politikaları ve teknik entegrasyon açısından çeşitli zorluklar doğurabilir. Bu bağlamda, AI destekli IoT sistemlerinin bulut tabanlı entegrasyonu ve kullanıcılar arasında uyumlu bir yönetim anlayışı kritik önem taşımaktadır (Priyanka vd., 2024).

Tablo 2’de de gösterildiği gibi, topluluk bulut modeli; maliyet paylaşımı ve özelleştirme avantajlarının yanında sürdürülebilirlik, yönetsel karmaşıklık ve veri sahipliği gibi bazı sınırlamalar da içermektedir.

Tablo 2. Bulut bilişim dağıtım modellerinin avantaj ve dezavantajları (Goyal, 2014’ten derlenmiştir) ²

Model	Avantajlar	Dezavantajlar
Özel Bulut	- Yüksek veri güvenliği ve gizlilik.	- Yüksek kurulum ve yönetim maliyetleri.
	- Tam kontrol ve özelleştirme.	- Teknik bilgi ihtiyacı vardır.
	- Yasal düzenlemelere kolay uyum.	
Kamu Bulut	- Düşük başlangıç maliyetleri.	- Veri güvenliği ve gizliliğiyle ilgili endişeler.
	- Kullandıkça ödeme modeli.	- Veri kontrolünün üçüncü taraf sağlayıcıda olması.
	- Yüksek ölçeklenebilirlik.	- Yoğun trafikte performans sorunları.
Hibrit Bulut	- Esneklik ve maliyet etkinliği sağlar.	- Entegrasyon ve yönetim karmaşıklığı.
	- Hem özel hem kamu bulutlarının avantajlarını birleştirir.	- Veri taşınabilirliği sorunları.
Topluluk Bulut	- Kaynak paylaşımı sayesinde maliyet avantajı sağlar.	- Yönetim ve iş birliği gereksinimleri.
	- Belirli sektörler veya kullanıcı gruplarına özelleştirilebilir.	- Sürdürülebilirlik ile ilgili zorluklar.

² Goyal, S. (2014). Public vs private vs hybrid vs community-cloud computing: a critical review. *International Journal of Computer Network and Information Security*, 6(3), 20–29. <https://www.mecspress.org/ijcnis/ijcnis-v6-n3/IJCNIS-V6-N3-3.pdf>

Farklı bulut dağıtım modellerinin sektörel gereksinimlere uygun olarak nasıl avantajlar ve dezavantajlar sunduğu Tablo 2’ de açıkça görülmektedir. Bu dağıtım modellerinin çeşitli sektörlerdeki kullanım senaryoları aşağıda detaylandırılmıştır.

2.5. Bulut Bilişimin Farklı Sektörlerdeki Etkileri

BB, farklı sektörlerde geniş kapsamlı dönüşümler sağlayarak veri yönetimi, operasyonel verimlilik ve maliyet tasarrufu konularında önemli katkılar sunmaktadır. Sağlık, finans, perakende, üretim, eğitim ve akıllı şehirler gibi çeşitli alanlarda BB’nin sunduğu avantajlar, büyük veri analitiği ve AI ile desteklenerek süreçleri optimize etmektedir.

Bu avantajların yanında, veri güvenliği, regülasyonlara uyumluluk ve yüksek altyapı maliyetleri gibi bazı zorluklar da sektörel düzeyde farklı etkiler yaratmaktadır. Bu farklılıkları daha detaylı ele almak amacıyla Tablo 3’te BB’nin farklı sektörlerde sağladığı avantajlar ve karşılaşılan zorluklar özetlenmiştir.

Tablo 3. Bulut bilişimin farklı sektörlerdeki avantajları ve zorlukları (Birden çok kaynaktan derlenmiştir).

Sektör	Ana Avantajlar	Ana Zorluklar
Sağlık	BB teknolojilerinin sağlık hizmetlerinde kullanımı, hasta verilerine hızlı erişim sağlayarak klinik karar destek sistemlerinin etkinliğini artırmakta ve teşhis süreçlerini desteklemektedir (Ali vd., 2018).	Sağlık verilerinin güvenli biçimde yönetilmesi, HIPAA/GDPR gibi regülasyonlara uyum ve altyapı güvenliği gibi zorluklar sağlık sektöründe BB kullanımını sınırlamaktadır (Devarajan, 2020).
Finans	BB tabanlı büyük veri analitiği, müşteri davranışlarının analizini kolaylaştırmakta ve dolandırıcılık tespitinde finansal kurumlara hız ve doğruluk kazandırmaktadır (Ravi ve Kamaruddin, 2017).	BB kullanımıyla birlikte siber saldırı riski artmakta ve finansal regülasyonlara uyum zorunluluğu, güvenlik politikalarının daha karmaşık hale gelmesine neden olmaktadır (Mustafa ve Bukhari, 2023).
Perakende	Bulut tabanlı müşteri analitiği uygulamaları, perakende sektöründe müşteri davranışlarının daha etkili analiz edilmesini sağlayarak hizmet kalitesini artırmakta; aynı zamanda stok yönetimi gibi operasyonel süreçlerde önemli verimlilik kazançları sunmaktadır (Lisowska ve Pamula, 2020).	BB’nin perakende sektöründe yaygınlaşması, yüksek rekabet nedeniyle büyük ölçekli veri yönetimi ihtiyacı yaratırken, veri gizliliği ihlalleri önemli bir risk oluşturmaktadır (Lisowska ve Pamula, 2020).

Tablo 3. (Devamı)

Sektör	Ana Avantajlar	Ana Zorluklar
Üretim	IoT entegrasyonu, bulut tabanlı üretim sistemlerinin verimliliğini artırmakta ve tedarik zinciri yönetiminde daha etkin ve esnek çözümler sunmaktadır. Bu tür sistemlerin üretim süreçlerini gerçek zamanlı verilerle optimize ederek operasyonel maliyetleri düşürdüğünü ve karar destek süreçlerini hızlandırdığını belirtmektedir. (Shen vd. 2024).	BB tabanlı üretim sistemlerinde veri analitiği altyapısına yatırım gereksinimi, özellikle küçük ve orta ölçekli işletmeler (KOBİ'ler) için maliyet ve teknik uzmanlık açısından önemli bir engel teşkil etmektedir. Hassan (2021), bu tür işletmelerin BB hizmetlerine geçişte en büyük zorluğu, yüksek ilk yatırım maliyetleri ve teknik bilgi eksikliği olarak belirlemiştir.
Eğitim	BB tabanlı çevrimiçi öğrenme sistemleri, uzaktan eğitimi yaygınlaştırmış; kullanıcı dostu platformlarla öğrenci erişimini ve sistem verimliliğini artırmıştır (Almaiah vd., 2020).	Bu sistemlerin benimsenmesinde teknik yeterlilik, güvenlik ve sistem kalitesinin belirleyici olduğunu vurgularken; dijital eşitsizlik, içerik uyumsuzluğu ve güvenlik açıklarının ise önemli zorluklar arasında yer aldığını ifade edilmektedir Almaiah vd. (2020).
Akıllı Şehirler	BB tabanlı çözümler, özellikle akıllı şehir projelerinde enerji verimliliği ve kamu hizmetlerinin optimizasyonu açısından önemli katkılar sağlamaktadır. Priyanka ve Prabhakaran (2024), IoT destekli bulut sistemlerinin metro hatları ve altyapı risk tahmininde kullanılmasının şehirlerde kaynak verimliliğini artırdığını vurgulamaktadır.	Büyük veri yönetimi ve siber güvenlik tehditleri, akıllı şehir altyapıları için ciddi zorluklardır. Priyanka ve Prabhakaran (2024), IoT tabanlı sistemlerin yüksek veri işleme taleplerine dikkat çekerken, Kropczynski vd. (2024) güvenlik ve standart eksikliklerinin tehditleri artırdığını ifade etmektedir.

Perakende sektöründe, müşteri analitiği ve envanter yönetimi gibi operasyonel süreçlerde BB'nin katkıları öne çıkarken; büyük veri yönetimi ve veri gizliliği gibi konular önemli riskler teşkil etmektedir (bkz. Tablo 3). Üretim sektöründe, IoT ile entegre BB çözümleri üretim optimizasyonu sağlamakta, ancak özellikle KOBİ'ler için maliyet ve teknik yeterlilik temel engeller olarak öne çıkmaktadır (Shen vd., 2024; Hassan, 2021).

Eğitimde, çevrimiçi öğrenme sistemlerinin yaygınlaşması BB ile altyapı verimliliğini artırsa da dijital eşitsizlik ve sistem güvenliği ciddi sorunlar oluşturmaktadır (Almaiah vd., 2020). Akıllı şehirlerde ise trafik ve enerji yönetimi gibi alanlarda sağlanan kazanımlar, büyük ölçekli veri yönetimi ve güvenlik tehditleriyle dengelenmektedir (Priyanka ve Prabhakaran, 2024; Kropczynski vd., 2024).

BB'nin farklı sektörlerde sunduğu avantajları maksimize etmek ve zorlukları en aza indirmek için güvenlik politikalarının güçlendirilmesi, sektöre özgü regülasyonlara uyum sağlanması ve altyapı yatırımlarının artırılması gerekmektedir.

Sektörel avantajlar ve zorlukların incelenmesi, BB'nin genel özelliklerini ve bu özelliklerin nasıl optimize edilebileceğini anlamak için kritik bir temel oluşturur. Bu bağlamda, bir sonraki bölümde BB'nin temel özellikleri detaylandırılacaktır.

2.6. Bulut Bilişimin Özellikleri

BB tanımları ile ilgili yapılan araştırma ve incelemeler sonucunda, “Çoklu Kiracılık (Kaynakların Paylaşımı), Etkili Ölçeklenebilirlik, Esneklik, Kullandığın Kadar Öde, Kaynakların Kendi Kendini Tedarik Etmesi” başlıkları altında uluslararası düzeyde kabul gören beş temel özelliği ön plana çıkmaktadır (Çam, 2012).

2.6.1. Çoklu Kiracılık (Kaynak Paylaşımı)

BB'nin öne çıkan temel özelliklerinden biri olan çoklu kiracılık, aynı fiziksel altyapı kaynaklarının (örneğin sunucu, depolama, ağ) birden fazla kullanıcı tarafından esnek bir şekilde paylaşılmasına imkân tanır. Bu yapı, özellikle küçük ve orta ölçekli işletmelere (KOBİ'ler), yüksek donanım maliyetlerine katlanmadan bulut hizmetlerinden faydalanma fırsatı sunar (Mell ve Grance, 2011). AWS ve Azure gibi önde gelen servis sağlayıcılar, bu paylaşımı hipervizör tabanlı sanallaştırma ya da konteyner teknolojileri aracılığıyla yöneterek maliyet verimliliğini en üst düzeye çıkarmaktadır (Armbrust vd., 2010).

Ancak aynı altyapı üzerinde birden fazla kullanıcıya hizmet verilmesi, “gürültülü komşu” etkisi gibi sorunlara yol açabilir; bu durum, bir kullanıcının yoğun kaynak tüketiminin diğer kullanıcıların performansını olumsuz etkilemesiyle sonuçlanır (Varghese ve Buyya, 2018). Ayrıca veri yalıtımı eksikliği, gizlilik ve güvenlik açısından hassas sektörler için büyük bir sorun teşkil edebilir. Bu nedenle sağlık (HIPAA) ya da finans (PCI DSS) gibi düzenlemelere tabi alanlarda, AES-256 gibi güçlü şifreleme yöntemleri, rol tabanlı erişim kontrolü (RBAC) ve OpenStack benzeri yalıtım çözümleri tercih edilmektedir (NIST, 2013; Sefraoui vd., 2012).

Örneğin, Salesforce gibi çoklu kiracılık mimarisi kullanan platformlar, her müşterinin verisini güvenli ve ayrı bir şekilde yönetmek için gelişmiş kaynak ayırma stratejileri ve şifreleme tekniklerinden yararlanmaktadır (Subramanian ve Jeyaraj, 2018). Böylece yüksek erişilebilirlik ve güvenilirlik sağlarken, sistem performansı da korunmuş olur.

2.6.2. Etkili Ölçeklenebilirlik

BB teknolojisinin en öne çıkan avantajlarından biri olan *etkili ölçeklenebilirlik*, sistem kaynaklarının (örneğin işlem gücü, depolama alanı ve bant genişliği) iş yüküne göre dinamik biçimde artırılmasına veya azaltılmasına olanak sağlar (Mell ve Grance, 2011). Bu avantaj, özellikle yoğun trafik dalgalanmalarının yaşandığı uygulamalarda sistemlerin kesintisiz hizmet verebilmesi açısından kritik öneme sahiptir. Örneğin, AWS Auto Scaling ve Google Kubernetes Engine (GKE) gibi araçlar, kaynak talebini otomatik olarak algılayarak uygun şekilde tahsis yapmakta; böylece hem performansı hem de maliyeti optimize etmektedir (Leitner ve Cito, 2016).

Ancak çoklu kiracılı ortamlarda bu esneklik bazı riskleri de beraberinde getirir. Aynı fiziksel altyapının birden fazla kullanıcı tarafından eşzamanlı kullanılması, kaynak rekabetine yol açarak performans dalgalanmalarına neden olabilir (O'Loughlin ve Gillam (2015). Özellikle sağlık ve finans gibi regülasyona tabi sektörlerde, kaynak yönetimi yalnızca performans açısından değil, veri güvenliği bağlamında da yüksek öneme sahiptir. Bu nedenle, katmanlı erişim kontrolleri, gelişmiş izleme sistemleri ve şifreleme teknolojileri gibi koruyucu önlemler yaygın olarak kullanılmaktadır (Subramanian ve Jeyaraj, 2018).

Ölçeklenebilirliğin sistem güvenilirliği üzerindeki etkisini değerlendirmek amacıyla son yıllarda “kaos mühendisliği” yaklaşımı öne çıkmaktadır. Yadav (2024), bu yaklaşımın dağıtık sistemlerin dayanıklılığını test etmek için kontrollü hata senaryoları sunduğunu ve özellikle bulut tabanlı sistemlerin güvenilirliğini artırmada önemli rol oynadığını vurgulamaktadır. Bu tür uygulamalar, sistem zayıflıklarını proaktif olarak tespit ederek altyapının olası kesintilere karşı daha hazırlıklı olmasını sağlar.

2.6.3. Esneklik

BB'nin öne çıkan temel özelliklerinden biri olan esneklik, kullanıcıların mekân ve zaman sınırlamasına bağlı kalmadan sistem kaynaklarına erişimini mümkün kılmaktadır. Bu durum, uzaktan çalışma modelleri ve mobil cihazlardan erişim gibi yeni nesil iş yapış biçimlerine doğrudan katkı sağlamaktadır (Marston vd., 2011). Örneğin, Google Workspace ve Microsoft 365 gibi bulut tabanlı çözümler sayesinde, farklı coğrafyalarda bulunan çalışanlar aynı belgeler üzerinde eş zamanlı çalışabilmekte, bu da ekip verimliliğini artırmaktadır (Al-Hajri vd., 2021).

Kurumsal ölçekte ise esneklik yalnızca bireysel erişimle sınırlı kalmaz; aynı zamanda şirket içi veri merkezleriyle genel bulut hizmetlerinin entegre bir şekilde çalışmasına da olanak tanır. Zhang, Cheng ve Boutaba (2010), hibrit bulut

altyapılarının, özel bulutların güvenliği ile genel bulutların esnekliğini birleştirerek işletmelere daha dengeli ve kontrol edilebilir çözümler sunduğunu belirtmektedir. Bu tür mimariler sayesinde kurumlar, verilerini kendi fiziksel ortamlarında muhafaza ederken, ani işlem yükü artışlarını genel bulut kaynakları aracılığıyla karşılayabilmektedir.

Esneklik bu geniş erişim imkânlarını sunarken, güvenlik açısından bazı tehditleri de beraberinde getirmektedir. Özellikle açık Wi-Fi ağları üzerinden yapılan erişimler, kullanıcıları “Man-in-the-Middle” (MITM) türü saldırılara açık hale getirebilir (Gupta ve Kumar, 2021). Bu tür tehditlerin önüne geçmek için, çok faktörlü kimlik doğrulama (MFA), sanal özel ağ (VPN) bağlantısı, sıfır güven mimarisi ve rol tabanlı erişim kontrolü (RBAC) gibi ileri seviye güvenlik önlemleri yaygın şekilde kullanılmaktadır (NIST, 2020; Subramanian ve Jeyaraj, 2018).

2.6.4. Kullandığın Kadar Öde Modeli

BB’nin öne çıkan avantajlarından bir diğeri olan “kullandığın kadar öde” modeli, işletmelere yalnızca tükettikleri kaynaklar kadar ödeme yapma olanağı tanıyarak hem operasyonel hem de sermaye maliyetlerinde önemli tasarruflar sağlar. Armbrust vd. (2010), bu modelin, özellikle kaynak ihtiyacının dalgalı olduğu durumlarda maliyet verimliliği sunduğunu vurgulamaktadır. Aynı şekilde, Al-Hajri vd. (2021), küçük ölçekli girişimlerin bu model sayesinde yüksek başlangıç yatırımlarına gerek kalmadan faaliyet gösterebildiklerini belirtmektedir.

Ayrıca, Azure Cost Management gibi araçlar, kurumlara kaynak kullanımlarını izleme, harcama sınırları belirleme ve israfı önleme imkânı sunarak finansal kontrolü artırmaktadır (Pakdil ve Çelik, 2023).

Ancak bu esnek modelin bazı riskleri de bulunmaktadır. Özellikle ölçeklenebilir kaynakların kontrolsüz kullanımı “maliyet sürüklenmesi” (cost drift) olarak bilinen olumsuz bir duruma yol açabilir. Yadav vd. (2021), bu gibi durumları önlemek için FinOps (Finansal Operasyonlar) yaklaşımlarının ve AI destekli tahmin sistemlerinin kullanılması gerektiğini önermektedir.

2.6.5. Kaynakların Kendi Kendini Tedarik Etmesi

BB’nin en dikkat çeken özelliklerinden biri olan kendi kendini tedarik etme yeteneği, kullanıcıların yalnızca ihtiyaç duydukları kaynaklar için ödeme yapmasını sağlayarak hem maliyet hem de operasyonel verimlilik açısından önemli avantajlar

sunar (Pakdil ve Çelik, 2023; Morris, 2020). Günümüzde bu model, geleneksel altyapı kurulumlarının yerini giderek daha fazla otomatikleştirilmiş çözümlere bırakmaktadır.

Kullanıcılar, Terraform, AWS CloudFormation gibi Altyapı Kodlama (Infrastructure as Code – IaC) araçları aracılığıyla birkaç dakika içinde sanal makineler, veri depolama birimleri ve ağ bileşenleri gibi kaynakları kolayca oluşturabilir. Morris (2020), bu tür araçların bulut altyapısının esnekliğini artırmakla kalmayıp, aynı zamanda sürekli teslimat ve sürekli entegrasyon (CI/CD) süreçleriyle birlikte kullanıldığında sistemlerin hızla güncellenmesini sağladığını vurgulamaktadır.

Bu avantajlara rağmen, yüksek düzeyde otomasyonun getirdiği riskler de göz ardı edilmemelidir. Özellikle insan denetiminin azalmasıyla birlikte, yapılandırma hataları ya da güvenlik açıkları gibi sorunlar ortaya çıkabilir. Bu tür riskleri en aza indirmek amacıyla, Open Policy Agent (OPA) gibi politika tabanlı denetim sistemleri ve CI/CD boru hatlarına entegre edilen statik kod analiz araçları sıklıkla tercih edilmektedir (Paleyes vd., 2020).

Nitekim Spotify, Kubernetes ve Google Cloud altyapılarını kullanarak, mikroservis mimarisini dinamik bir şekilde yönetmekte ve sistem ihtiyaçlarına anında yanıt verebilen esnek kaynak tedarik süreçleri geliştirmektedir (Younis vd., 2024).

Bu özellikler, BB'nin avantaj ve dezavantajlarının temelini oluşturmaktadır. Özellikle güvenlik ve veri gizliliği gibi kritik konular, bir sonraki bölümde derinlemesine irdelenecektir.

2.7. Bulut Bilişimde Güvenlik ve Veri Gizliliği

BB teknolojilerinin yaygınlaşması, güvenlik ve veri gizliliği konularını hem kurumsal hem de bireysel düzeyde giderek daha önemli bir endişe kaynağı haline getirmiştir. Özellikle çoklu kiracılık mimarisi ve uzaktan erişim gereksinimi, siber tehditlerin artmasına ve veri ihlali risklerinin çoğalmasına zemin hazırlamakta; bu durum, güvenlik protokollerinin sürekli olarak güncellenmesini zorunlu kılmaktadır (Subashini ve Kavitha, 2011; Hashizume vd., 2013).

Bu kapsamda, sağlık, finans ve kamu gibi hassas verilerin işlendiği sektörlerde, veri güvenliği büyük bir öncelik haline gelmiştir. Bu alanlarda gelişmiş şifreleme yöntemleri, çok faktörlü kimlik doğrulama ve uluslararası mevzuatlara uygunluk gibi önlemler, güvenliğin sağlanmasında hayati rol oynamaktadır (Zissis ve Lekkas, 2012).

Bu doğrultuda, BB sistemlerinde güvenliği sağlamak için hem teknik hem de yönetsel çeşitli yaklaşımlar benimsenmektedir. Aşağıda bu yaklaşımlar detaylı şekilde ele alınmıştır.

2.7.1. Paylaşılan Sorumluluk Modeli

Bulut güvenliği, yalnızca hizmet sağlayıcılarının değil, kullanıcıların da sorumluluk üstlendiği bir model çerçevesinde yönetilmektedir. Subashini ve Kavitha (2011), bulut sağlayıcılarının fiziksel ve sanal altyapının güvenliğini sağlamakla yükümlü olduğunu, kullanıcıların ise veri güvenliğini artırmak için erişim yönetimi ve şifreleme politikalarını etkin bir şekilde uygulaması gerektiğini vurgulamaktadır.

Ancak, bu paylaşılan sorumluluk modeli, kullanıcıların güvenlik önlemlerini yetersiz uygulaması durumunda riskleri artırabilir. Hashizume ve arkadaşları (2013), birçok işletmenin güvenlik politikalarını yalnızca sağlayıcıların sunduğu standart çözümlerle sınırlı tuttuğunu ve bu durumun güvenlik açıklarını artırabileceğini belirtmektedir.

2.7.2. Veri Şifreleme ve Koruma

BB'de veri güvenliğini sağlamanın en etkili yöntemlerinden biri şifreleme teknolojileridir. Fernandes ve arkadaşları (2014), özellikle AES-256, homomorfik şifreleme ve şifreli veri paylaşım protokollerinin yetkisiz erişime karşı önemli bir koruma sağladığını vurgulamaktadır. Ayrıca, sıfır bilgi kanıtı (Zero-Knowledge Proof) gibi doğrulama teknikleri, kullanıcıların verilerini üçüncü taraflarla paylaşmadan kimlik doğrulaması yapabilmesine olanak tanıyarak veri güvenliğini artırmaktadır (Zissis ve Lekkas, 2012).

Bunun yanı sıra, veri anonimleştirme yöntemleri, kişisel bilgilerin korunmasına katkıda bulunarak yetkisiz erişim ve veri sızıntısı risklerini azaltmada etkili bir rol oynamaktadır. Fernandes ve arkadaşlarının (2014) çalışmasında bu tekniklerin güvenli veri işleme süreçlerinde önemli katkılar sağladığı ifade edilmektedir. Ancak, anonimleştirme yöntemlerinin hatalı uygulanması, belirli durumlarda verilerin yeniden tanımlanmasına ve kimlik tespitine yol açabilmektedir (Modi vd., 2013).

2.7.3. Kimlik ve Erişim Yönetimi (IAM)

Bulut sistemlerinde kimlik ve erişim yönetimi, güvenlik mimarisinin en kritik unsurlarından biridir. Hashizume ve arkadaşları (2013), IAM politikalarının eksik ya da hatalı tanımlanmasının veri ihlallerine zemin hazırlayabileceğini belirtmektedir. Bu durum, kullanıcı erişim düzeylerinin doğru şekilde yapılandırılmasının ve yönetilmesinin ne kadar önemli olduğunu göstermektedir.

Modi ve arkadaşları (2013), çok faktörlü kimlik doğrulama (MFA) ve rol tabanlı erişim kontrolü (RBAC) mekanizmalarının yetkisiz erişim risklerini önemli ölçüde

azalttığını vurgulamaktadır. Buna ek olarak, Almorsy ve arkadaşları (2016), bulut mimarisindeki çok katmanlı bağımlılıkların IAM üzerindeki etkisini ele alarak, erişim kontrolünün sadece ağ seviyesinde değil, sistem genelinde uygulanmasının gerekli olduğunu ifade etmektedir.

2.7.4. Uyumluluk ve Yasal Zorluklar

Küresel ölçekte BB kullanımının yaygınlaşması, farklı ülkelerde geçerli veri koruma mevzuatlarına uyumu zorunlu hale getirmiştir. Avrupa Birliği Genel Veri Koruma Tüzüğü (GDPR), Amerika Birleşik Devletleri'nde yürürlükte olan Sağlık Sigortası Taşınabilirlik ve Sorumluluk Yasası (HIPAA) ve ISO 27001 gibi uluslararası güvenlik standartları, bulut ortamlarında veri yönetimini daha karmaşık hale getirmektedir (Zissis ve Lekkas, 2012).

Çapraz sınır veri akışları, yasal düzenlemeler nedeniyle riskleri artırmakta ve işletmelerin belirli coğrafi bölgelerde veri merkezleri kurmasını gerektirmektedir (Fernandes vd., 2014). Bu doğrultuda, birçok büyük ölçekli işletme veri yerelleştirme stratejileri benimseyerek regülasyonlara uyum sağlamak ve verilerini ilgili yasal çerçeveye uygun şekilde depolamaktadır (Modi vd., 2013).

2.7.5. Yeni Nesil Tehditler

BB altyapısının gelişmesiyle birlikte siber tehditler de daha karmaşık hale gelmiştir. Ristenpart vd. (2009), sanal makineler arasında “co-residency” üzerinden yürütülen yan kanal saldırılarıyla veri sızıntısı gerçekleştirilebildiğini göstermiştir. Bu durum, kullanıcılar arasında tam izolasyon sağlanamaması halinde ciddi güvenlik açıklarının oluşabileceğine işaret etmektedir.

Fernandes vd. (2014), mikro hizmet mimarileri ve konteyner tabanlı sistemlerin, saldırganlara yeni giriş noktaları sunduğunu belirtmektedir. Aynı şekilde, API'lerdeki güvenlik açıkları da yetkisiz erişim riskini artırarak, bulut sistemlerinin zayıf noktaları arasında yer almaktadır (Modi vd., 2013).

2.8. Bulut Bilişimde Güvenlik Önlemleri ve Stratejiler

BB'de güvenliğin sağlanması, artan siber tehditlere karşı çok katmanlı ve bütüncül yaklaşımları gerekli kılmaktadır. Geleneksel yöntemlerin yetersiz kaldığı bu ortamda, sağlayıcılar ve kullanıcılar arasında paylaşılan sorumluluklar doğrultusunda geliştirilen stratejiler kritik önem taşımaktadır. Bu bölümde, BB altyapılarında sıkça uygulanan temel güvenlik önlemleri ve yaklaşımlar incelenmektedir.

1. **Sıfır Güven Mimarisi (Zero Trust):** Bu yaklaşım, sistemdeki hiçbir varlığa otomatik olarak güven duyulmaması ve her erişim talebinin sürekli doğrulanması esasına dayanır. Böylece, yetkisiz erişim ihtimali büyük ölçüde azaltılmış olur (Modi vd., 2013). Örneğin, Microsoft, Azure altyapısında sıfır güven modelini benimseyerek, kimlik doğrulama ve cihaz uyumluluğunu her erişim talebinde yeniden kontrol ederek saldırı yüzeyini daraltmıştır (Gupta ve Miryala, 2022).
2. **Makine Öğrenimi Tabanlı İzleme:** Kullanıcı davranışlarını analiz ederek olası tehditleri proaktif bir şekilde tanımlama kapasitesine sahip olan bu sistemlerin güvenlik altyapılarına entegre edilmesi, yalnızca reaktif güvenlik önlemleriyle yetinmeyip, potansiyel saldırıların önceden tespit edilmesini de mümkün kılmaktadır. Ogendi'nin (2022) çalışması, bu teknolojilerin sıfır güven mimarileriyle bütünleşik olarak kullanıldığında, özellikle anormal etkinliklerin erken aşamada saptanmasında etkili bir güvenlik katmanı sunduğunu ortaya koymaktadır.
3. **Şeffaf Veri Yönetimi:** Veri erişimi ve kontrolünün tamamen kullanıcıya bırakılması, BB'de güveni artıran bir diğer stratejidir. Bu sayede veri şeffaflığı sağlanmakta ve kullanıcılar verilerinin kimler tarafından ne zaman ve ne şekilde kullanıldığını izleyebilmektedir (Zissis ve Lekkas, 2012). Örneğin, Google Cloud Platform, kullanıcılarına erişim günlüklerini ve IAM panellerini detaylı bir şekilde sunarak, hangi verilere ne zaman erişildiğini izlenebilir hale getirmekte ve GDPR gibi regülasyonlara uyum sağlamaktadır (Sarkar vd., 2022).
4. **Uyumluluk Tabanlı Güvenlik Politikaları:** Ulusal ve uluslararası regülasyonlara uyum sağlayabilmek amacıyla, sistemlerin tasarımı esnasında veri yerelleştirme, erişim denetimi ve kayıt izleme gibi ilkelere dayalı güvenlik stratejileri uygulanmaktadır. Örneğin, ABD'de federal kurumlar, CISA yönergeleri kapsamında Zero Trust ilkelerine dayalı mimarilere geçerek hem HIPAA hem de FISMA regülasyonlarına uygun veri işleme ortamları oluşturmuştur (Ofili vd., 2025).
5. **Konteyner ve Mikro Hizmet Güvenliği.** Modern bulut yapılarında kullanılan konteyner ve mikro hizmet mimarilerinde, her bileşen için ayrı politika tanımları ve izolasyon önemlidir. Örneğin, Netflix, Kubernetes üzerinde çalışan mikro servislerinde Open Policy Agent (OPA) kullanarak

erişim denetimini merkezi biçimde uygulamakta, bu sayede hem operasyonel esneklik hem güvenlik sağlamaktadır (Fernando, 2022).

Bu stratejiler, yalnızca güvenlik risklerini azaltmakla kalmaz; aynı zamanda BB sistemlerinin avantajlarından en verimli şekilde yararlanmayı da mümkün kılar. Bir sonraki bölümde bu dengeyin sektörel yansımaları ve BB'nin avantaj-dezavantaj analizi detaylı olarak ele alınacaktır.

2.9. Bulut Bilişim Sistemlerinin Avantajları ve Dezavantajları

BB'de güvenliğin sağlanması, artan siber tehditlere karşı çok katmanlı ve bütüncül yaklaşımları gerekli kılmaktadır. Geleneksel yöntemlerin yetersiz kaldığı bu ortamda, sağlayıcılar ve kullanıcılar arasında paylaşılan sorumluluklar doğrultusunda geliştirilen stratejiler kritik önem taşımaktadır. Bu bölümde, BB altyapılarında sıkça uygulanan temel güvenlik önlemleri ve yaklaşımlar incelenmektedir.

2.9.1. Bulut Bilişim Sistemlerinin Avantajları

BB işletmelerin ve bireylerin bilişim altyapılarını yönetme ve kullanma biçimlerini köklü bir şekilde değiştirmektedir. Bu teknolojinin en önemli avantajları aşağıda ele alınmıştır:

- 1. Maliyet Etkinliği ve Verimlilik:** BB sistemleri, işletmelere fiziksel donanım ve altyapıya büyük yatırımlar yapmaksızın yalnızca tükettikleri kadar kaynak üzerinden ödeme yapma imkânı sunar. Bu model, özellikle bütçesi sınırlı küçük ve orta ölçekli işletmeler (KOBİ'ler) için ciddi bir maliyet avantajı sağlar ve kaynakların daha verimli yönetilmesini mümkün kılar (Marston vd., 2011). Armbrust ve arkadaşları (2010), geleneksel bilgi teknolojisi altyapılarıyla kıyaslandığında bulut sistemlerinin sermaye harcamalarını (CapEx) ortadan kaldırarak, firmaların operasyonel harcamalara (OpEx) yönelmesini kolaylaştırdığını ifade etmektedir. Benzer şekilde, Marston vd. (2011) de, pay-as-you-go modeli sayesinde firmaların yalnızca kullandıkları kaynaklar kadar ödeme yaparak daha esnek bir bütçe planlaması gerçekleştirebildiğini belirtmektedir. Borra (2024), AWS üzerinde gerçekleştirdiği çalışmada, elastik altyapıların şirketlere iş yüküne göre kaynak tahsisini dinamik olarak optimize etme olanağı sunduğunu ortaya koymuştur. Bu yapı, değişken iş hacmine sahip işletmeler için özellikle etkili bir çözüm sunmaktadır.
- 2. Esneklik ve Ölçeklenebilirlik:** BB altyapısı, sistem kaynaklarının dinamik olarak artırılabilmesi ya da azaltılabilmesi sayesinde iş yükündeki ani

değişikliklere hızla yanıt verebilen esnek bir yapı sunar. Bu avantaj, özellikle e-ticaret, medya ve oyun sektörleri gibi yoğun veri trafiğine maruz kalan alanlarda hizmet kesintilerini önleyerek sistem sürekliliğini sağlar (Leitner ve Cito, 2016). Tiwari ve arkadaşları (2021), Black Friday ve Cyber Monday gibi yoğun dönemlerde, AWS Lambda gibi sunucusuz mimarilerin sipariş, stok ve ödeme işlemlerinde yüksek talebi karşılayacak şekilde otomatik olarak ölçeklendirilerek hizmet kalitesini koruduğunu belirtmektedir. Armbrust vd. (2010), bulut yapılarının geleneksel sistemlere göre hem daha hızlı uyarlanabilir hem de yatırım maliyetini düşüren bir yapı sunduğunu ifade etmektedir. Böylece organizasyonlar, kısa vadeli projelerde dahi büyük altyapı yatırımları yapmak zorunda kalmadan ihtiyaç duydukları hizmetlere erişebilmektedir.

3. Veri Yedekleme ve Olağanüstü Durum Kurtarma: BB sistemleri, otomatik yedekleme, esnek kaynak yönetimi ve coğrafi olarak yedekli veri merkezleri sayesinde sistemlerin sürekliliğini desteklemekte, beklenmedik kesintilere ve siber tehditlere karşı güçlü bir güvenlik altyapısı sunmaktadır. Armbrust ve arkadaşları (2010), bulut altyapılarının esnekliği sayesinde felaket kurtarma çözümlerinin daha erişilebilir ve maliyet etkin hale geldiğini ifade etmektedir. Voorsluys, Buyya ve Broberg (2011) de hizmet seviyesi anlaşmalarının (SLA), erişim süresi ve yedekleme sistemleri için minimum hizmet garantisi sunduğunu ve bu mekanizmanın iş sürekliliği açısından kritik olduğunu belirtmektedir. Son olarak, Luo ve arkadaşları (2019), büyük ölçekli bulut sağlayıcılarının sistem erişilebilirliğini SLA kapsamında %99,9 ve üzeri seviyelerde sürdürdüğünü, bu sayede veri yedekleme ve felaket kurtarma süreçlerinde yüksek güvenilirlik sağlandığını ortaya koymuştur.

4. Mobilite ve Erişilebilirlik: Bulut teknolojileri, mekândan bağımsız olarak hizmetlerin kullanıcıya her yerden ve her cihazdan ulaşmasını mümkün kılmaktadır. Bu avantaj, özellikle eğitim ve sağlık gibi temel alanlarda erişilebilirliği artırmakta ve dijital eşitsizliği azaltmaktadır. Sadoughi ve Erfannia (2017), bulut tabanlı sağlık sistemlerinin kırsal bölgelerde yaşayan bireyler için tele-tıp uygulamaları aracılığıyla sağlık hizmetlerine erişimi kolaylaştırdığını göstermiştir. COVID-19 pandemisi, BB altyapılarının bu tür olağanüstü durumlarda sunduğu esnekliğin ne kadar kritik olduğunu bir kez daha ortaya koymuştur. Zhang ve arkadaşları (2020), hibrit bulut mimarilerinin sağlık ve eğitim gibi sektörlerde sistem esnekliğini güçlendirerek hizmet kesintilerini önlediğini vurgulamaktadır. Ayrıca, eğitim kurumları bulut destekli

LMS platformları aracılığıyla çevrim içi öğrenmeyi yaygınlaştırırken; sağlık kuruluşları hasta verilerini bulut ortamında güvenli şekilde yönetebilmiştir. Bu dönüşüm, bulut teknolojisinin sürdürülebilir dijitalleşme için vazgeçilmez bir unsur olduğunu göstermektedir.

5. Güncellemelerin Otomatikleştirilmesi: BB sistemleri, yazılımın kurulumu ve dağıtımı gibi süreçleri otomatikleştirerek hem güvenliği artırmakta hem de sistem sürekliliğini sağlamaktadır. Armbrust vd. (2010), manuel işlemlerin azalması sayesinde operasyonel verimliliğin arttığını ve insan hatalarının minimize edildiğini belirtmiştir. Hafeez vd. (2019), geliştirdikleri Hybrid Cloud Patch Manager (HCPM) sistemiyle yazılımların planlı bakım pencereleri içinde güvenli şekilde yamanabildiğini ve böylece sistem güvenliğinin güçlendirildiğini ifade etmektedir. Zhou vd. (2010), çevrim dışı çalışan ve ölçeklenebilir yama sistemlerinin, sanal makineler üzerinde hizmetlerin her zaman güncel kalmasını sağladığını göstermektedir. Dissanayake vd. (2022) ise, farklı sektörlerde yapılan uygulama örnekleri üzerinden, otomatik yama uygulamalarının siber güvenlik açıklarını azaltmada en etkili yöntemlerden biri olduğunu vurgulamıştır.

6. Çevresel Sürdürülebilirlik: BB teknolojileri, yalnızca teknik ve ekonomik avantajlar sunmakla kalmaz, aynı zamanda çevresel sürdürülebilirliğe de önemli katkılar sağlar. Özellikle enerji tüketiminin merkezi veri merkezlerinde daha verimli biçimde yönetilmesi, kurumların karbon ayak izini azaltmasına yardımcı olmaktadır. Kumar ve Buyya (2012), çoklu kiracılı mimarilerin donanım kaynaklarının daha verimli kullanılmasını sağlayarak enerji tasarrufu oluşturduğunu ifade etmektedir. Armbrust vd. (2010), büyük bulut sağlayıcılarının veri merkezlerinde kullanılan soğutma ve enerji teknolojileri sayesinde geleneksel BT altyapılarına göre daha sürdürülebilir olduğunu belirtmiştir. Buyya vd. (2012) ise Cloudbus Toolkit gibi yönetim araçlarının, kaynak kullanımını optimize ederek hem maliyetleri düşürdüğünü hem de çevresel etkileri azalttığını ortaya koymuştur. Bu gelişmeler, çevre dostu dijital dönüşüm stratejilerinin oluşturulmasında BB'nin kilit rol oynadığını ortaya koymaktadır.

2.9.2. Bulut Bilişim Sistemlerinin Dezavantajları

Her ne kadar BB sistemleri esneklik, maliyet tasarrufu ve ölçeklenebilirlik gibi birçok avantaj sunsa da bu teknolojilerin bazı önemli sınırlamaları göz ardı

edilmemelidir. Özellikle kesintisiz internet bağlantısına duyulan ihtiyaç, veri güvenliği ve gizliliğine ilişkin endişeler, olası hizmet kesintileri ile birlikte farklı ülkelerdeki yasal düzenlemelere uyum zorunluluğu gibi zorluklar, BB'nin uygulama sürecinde sıkça karşılaşılan sorunlardır.

- 1. Sürekli İnternet Bağlantısı İhtiyacı:** Kaynaklara internet üzerinden erişim prensibiyle çalışan BB sistemleri, altyapı eksikliklerinin bulunduğu kırsal veya gelişmemiş bölgelerde hizmet kesintilerine ve performans kayıplarına neden olabilir. Özellikle gerçek zamanlı veri işleme gerektiren sektörlerde, bu durum kritik sonuçlar doğurabilir (Dokuz ve Çelik, 2017). Bu tür sorunları en aza indirmek amacıyla hibrit bulut yapıları yaygın olarak tercih edilmekte; şirket içi altyapı ile bulut hizmetleri entegre edilerek internet erişimi olmasa dahi temel süreçlerin devamlılığı sağlanmaktadır (Leitner ve Cito, 2016). Ayrıca, sık erişilen verilerin yerel sunucularda tutulmasını sağlayan önbellekleme (caching) teknolojileri de hizmet sürekliliğine katkı sunmaktadır. Wang ve arkadaşları (2014), özellikle mobil sistemlerde uygulanan içerik önbellekleme stratejilerinin ağ gecikmelerini azaltarak veri erişim hızını ve hizmet sürekliliğini artırdığını belirtmektedir.
- 2. Güvenlik Riskleri:** Bulut sistemlerinde siber saldırılar, veri ihlalleri ve “gürültülü komşu” etkisi gibi riskler, özellikle çoklu kiracılı yapılarda güvenlik açıklarına yol açabilmektedir. Literatürde, bu ihlallerin önemli bir kısmının yanlış yapılandırılmış güvenlik politikaları, zayıf kimlik doğrulama sistemleri ve yetersiz erişim kontrollerinden kaynaklandığı belirtilmektedir (Dwivedi vd., 2023). Bu doğrultuda, AES-256 gibi güçlü şifreleme algoritmaları ve çok faktörlü kimlik doğrulama (MFA) yöntemleri, veri güvenliğini artırmak amacıyla yaygın olarak kullanılmaktadır (Fernandes vd., 2014; Dwivedi vd., 2023). Son yıllarda yaygınlaşan Sıfır Güven Mimarisi (Zero Trust) yaklaşımı ise, geleneksel ağa bir kez giriş yaptıysan güvendesin anlayışını terk ederek, her erişim isteğinde doğrulama yapılmasını şart koşmaktadır. Bu model, sistem genelinde güvenlik seviyesini yükseltmekte ve sızma girişimlerine karşı korumayı artırmaktadır (Dwivedi vd., 2023).
- 3. Hizmet Kesintileri.** Teknik arızalar, bakım çalışmaları veya altyapı problemleri gibi nedenlerle bulut hizmetlerinde kesintiler yaşanabilmektedir. Özellikle sağlık, finans ve e-ticaret gibi kesintisiz

hizmet gerektiren alanlarda bu tür durumlar ciddi maddi kayıplara ve kullanıcı memnuniyetsizliğine yol açabilir (Dokuz ve Çelik, 2017). Kesinti risklerini azaltmak amacıyla, coğrafi olarak dağıtılmış yedekli sunucu altyapıları ve yüksek erişilebilirlik stratejileri devreye alınmakta, böylece hizmet sürekliliği sağlanmaktadır (Buyya vd., 2013). Ayrıca, hizmet seviyesi anlaşmaları (SLA) sayesinde belirli kesinti süreleri içinde geri dönüş garantisi verilmekte olup bu tür anlaşmalar da kullanıcılar açısından güveni artırmaktadır (Leitner ve Cito, 2016).

- 4. Veri Gizliliği ve Kontrol Zorlukları:** Kullanıcı verilerinin hizmet sağlayıcıların kontrolündeki sunucularda saklanması, veri gizliliği ve verilerin kontrolü açısından bazı endişeleri beraberinde getirmektedir. Özellikle HIPAA, GDPR gibi sıkı regülasyonlara tabi sektörlerde, eksik yetkilendirme mekanizmaları ve yetersiz yapılandırmalar ciddi güvenlik ihlallerine neden olabilmektedir (Sharma vd., 2020; Bai vd., 2022). Bu tür güvenlik risklerini azaltmak amacıyla veri yerelleştirme stratejileri uygulanmakta; verilerin yasal sınırlar içinde depolanması sağlanarak düzenleyici uyumluluk desteklenmektedir (Zissis ve Lekkas, 2012). Erişim kontrolünün etkin yönetimi için otomatik sistemler ve Rol Tabanlı Erişim Kontrolü (RBAC) gibi mekanizmalar kullanılmaktadır. Fernandes ve arkadaşları (2014), RBAC sistemlerinin, kullanıcı yetkilerini dinamik biçimde sınırlandırarak yetkisiz erişim riskini azaltmada etkili olduğunu belirtmektedir.

Avantaj ve dezavantajların analizi, BB'nin gelecekteki yönelimlerini ve etkin kullanımını anlamak için bir çerçeve sunmaktadır. Bu doğrultuda, bir sonraki bölümde yapay zekâ, kenar bilişim, blockchain, sürdürülebilirlik ve kuantum güvenliği gibi yenilikçi yaklaşımların BB'ye etkileri ele alınacaktır.

2.10. Bulut Bilişimin Geleceği ve Yeni Teknolojik Eğilimler

BB, veri işleme, depolama ve dağıtım süreçlerinde kritik rol oynarken, gelişen teknolojiyle birlikte yeni yaklaşımlarla dönüşmektedir. Yapay zeka (AI), kenar bilişim, blokzincir, sürdürülebilir sistemler ve kuantum sonrası güvenlik çözümleri bu dönüşümde öne çıkan başlıca eğilimlerdir.

AI tabanlı yönetim sistemleri verimliliği artırmakta; kenar bilişim gecikmeleri verimliliği düşürmekte, blokzincir ise veri güvenliği ve şeffaflığı güçlendirmektedir (Zhang vd., 2020; Fernandes vd., 2014). Ayrıca, kuantum bilgisayarların gelişimiyle

post-kuantum kriptografi uygulamalarına duyulan ihtiyaç da artmaktadır (Lohachab vd., 2020). Bu doğrultuda, BB'nin geleceğini belirleyen bu eğilimler aşağıda detaylı biçimde ele alınmaktadır.

Yapay Zekâ Destekli Bulut Bilişim. AI ve ML, bulut hizmetlerinin verimliliğini artırmakta, kaynak tahsisi ve sistem optimizasyonu gibi süreçlerde öngörüye dayalı kararlar alınmasını mümkün kılmaktadır. AI tabanlı çözümler, bulut ortamındaki büyük veri kümelerini analiz ederek, dinamik iş yüklerine göre kaynak kullanımını otomatik olarak ayarlayabilmekte ve bu sayede performans kayıplarını önlemektedir (Ramamoorthi, 2024).

Örneğin, Google Cloud AI ve AWS Machine Learning platformları; müşteri davranışlarının analiz edilmesi, otomatik ölçekleme, kaynak planlaması ve hata öngörüsü gibi uygulamalar aracılığıyla operasyonel verimliliği artırmaktadır. Bu sistemler sayesinde şirketler, sistem performansını yükseltirken insan müdahalesine olan ihtiyacı da azaltabilmektedir (Ramamoorthi, 2024).

Ayrıca, AI tabanlı siber güvenlik çözümleri, tehdit algılama ve olay müdahale mekanizmalarında önemli rol oynamaktadır. Fernandes ve arkadaşları (2014), AI destekli analiz sistemlerinin, bulut ortamlarında gerçek zamanlı tehdit tespiti ve risk değerlendirmesi gerçekleştirerek veri güvenliğini güçlendirdiğini ifade etmektedir.

AI destekli sistemlerin bir diğer önemli katkısı da enerji verimliliğidir. Ramamoorthi (2024), AI tabanlı kaynak yöneticilerinin iş yükü profillerini analiz ederek veri merkezi enerji tüketimini optimize edebildiğini ve karbon ayak izini azaltabildiğini belirtmektedir.

AI sistemleri, sağladığı avantajların yanı sıra insan müdahalesinin azalması nedeniyle karar alma süreçlerinde bazı riskleri de beraberinde getirmektedir. Özellikle düşük kaliteli ya da önyargılı veri kümeleriyle eğitilmiş modeller, hatalı kaynak tahsisi veya tehditlerin yanlış sınıflandırılması gibi performans sorunlarına yol açabilir (Ramamoorthi, 2024).

Kenar Bilişim (Edge Computing) ve Bulut Bilişimin Geleceği. Geleneksel BB sistemlerinde veriler genellikle bulut sağlayıcılarının merkezi altyapılarında işlenirken, kenar bilişim (Edge Computing), veri işleme süreçlerini uç cihazlara veya verinin üretildiği fiziksel konuma daha yakın noktalara taşıyarak gecikme sürelerini azaltmakta, ağ bant genişliği ihtiyacını düşürmekte ve veri güvenliğini artırmaktadır (Shi vd., 2016).

Bu yaklaşım, özellikle gerçek zamanlı veri işleme gerektiren uygulamalarda (Nesnelerin İnterneti (IoT), akıllı şehir sistemleri ve endüstriyel otomasyon gibi alanlarda) etkinliğiyle öne çıkmaktadır. Shi ve arkadaşları (2016), kenar bilişimin

yalnızca veri işleme süreçlerini optimize etmekle kalmayıp, aynı zamanda enerji verimliliği, veri gizliliği ve kullanıcı deneyiminin iyileştirilmesi gibi önemli katkılar sunduğunu belirtmektedir.

Kenar bilişim, BB ile birlikte çalışarak hibrit bir mimari oluşturabilir. Bu modelde, verilerin önemli bir kısmı uçta işlenirken; yalnızca özet veya kritik veriler merkezi bulut sistemlerine gönderilir. Bu yaklaşım, yüksek veri hacmi üreten uygulamalarda ağ üzerindeki yükü azaltarak hem zaman kazanımı hem de sistem kararlılığı sağlamaktadır (Shi vd., 2016).

Blockchain Tabanlı Bulut Çözümleri. Blockchain teknolojisi, merkezi olmayan yapısıyla BB sistemlerine yüksek düzeyde güvenlik, şeffaflık ve veri bütünlüğü sağlamaktadır. Özellikle dağıtık defter yapısı ve değiştirilemez (immutable) veri kaydı özellikleri sayesinde siber saldırıların ve veri manipülasyonunun önüne geçilmesi mümkün hale gelmiştir (Merlec ve In, 2024).

Bu teknoloji, merkezi sistemlerin zayıf yönlerini azaltarak, kullanıcıların veriler üzerinde daha fazla kontrol sahibi olmasını sağlar. Filecoin, Storj ve Sia gibi blockchain tabanlı platformlar, kullanıcıların verilerini şifrelenmiş biçimde dağıtık ağlarda depolamasına olanak tanımaktadır. Bu sistemlerde veriler, farklı düğümlerde bölünmüş ve çoğaltılmış halde tutulduğu için hem güvenlik hem erişilebilirlik açısından avantaj sunmaktadır (Xie vd., 2019).

Ayrıca, akıllı sözleşmeler (smart contracts) aracılığıyla depolama anlaşmaları otomatikleştirilmekte ve veri erişim süreçleri denetlenebilir hale gelmektedir. Bu sayede, veri erişimi yalnızca yetkili taraflara açılarak gizlilik korunmakta, aynı zamanda kullanıcı ve hizmet sağlayıcılar arasında güvenli ve izlenebilir bir işlem altyapısı oluşturulmaktadır (Badhe ve Arjunwadkar, 2024).

Blockchain'in sunduğu bir diğer katkı ise siber saldırılara karşı dayanıklılıktır. Merkeziyetsizlik ilkesi, tek bir arıza noktasının sistemin tamamını etkilemesini engelleyerek, hizmet sürekliliğini artırmaktadır (Han ve Kim 2020).

Sürdürülebilir Bulut Bilişim ve Yeşil Veri Merkezleri. BB sistemlerinin sürdürülebilirliği, yalnızca performans değil aynı zamanda çevresel etki açısından da değerlendirilmektedir. Buyya ve Beloglazov (2010), enerji verimliliği öncelikli veri merkezi mimarilerinin karbon emisyonunu azaltmada etkili olduğunu ve bu tür yapılarla sürdürülebilir dijital altyapılar geliştirilebileceğini vurgulamaktadır.

Bianzino ve arkadaşları (2012), ağ ekipmanları ve veri merkezlerinin enerji tüketimini analiz ettikleri çalışmada, özellikle otomatik kapatma, akıllı soğutma sistemleri ve kaynak kullanımına göre dinamik enerji yönetimi gibi tekniklerin önemli

kazanımlar sağladığını belirtmiştir. AI destekli tahmine dayalı yönetim sistemlerinin, enerji israfını önleyerek hem çevreye olan etkiyi azalttığı hem de işletme maliyetlerini düşürdüğü ifade edilmektedir.

Öte yandan, Beloglazov ve Buyya (2012) tarafından geliştirilen bir enerji farkındalıklı kaynak yönetimi modeli (Energy-Aware Resource Allocation), CPU kullanımı, iş yükü profili ve uygulama taleplerine göre sistem kaynaklarını esnek biçimde yönlendirerek enerji tasarrufu sağlamaktadır. Bu model, geleneksel sabit kaynak tahsisi yöntemlerine göre %30'a kadar enerji verimliliği sağlamış ve çeşitli simülasyon ortamlarında başarıyla test edilmiştir.

Bu bulgular, sürdürülebilir bulut mimarilerinin yalnızca çevre dostu olmakla kalmayıp, aynı zamanda sistem performansını korurken enerji maliyetlerini optimize etme kapasitesine de sahip olduğunu göstermektedir.

Kuantum Bilişim ve Bulut Güvenliği. Kuantum bilişim teknolojileri, BB sistemlerinde güvenlik paradigmalarını köklü biçimde dönüştürme potansiyeline sahiptir. Geleneksel simetrik ve asimetrik şifreleme yöntemlerinin, kuantum bilgisayarların süperpozisyon ve kuantum paralellik gibi özellikleri karşısında savunmasız kaldığı bilinmektedir. Bernstein ve arkadaşları (2017), kuantum bilgisayarların RSA ve ECC gibi yaygın kullanılan kriptografik algoritmaları kısa sürede çözebileceğini ortaya koymuştur.

Bu nedenle, kuantuma dayanıklı (post-quantum) kriptografi teknikleri üzerinde yoğun akademik çalışmalar yürütülmektedir. Arute ve arkadaşları (2019), kuantum üstünlüğü (quantum supremacy) deneylerini takiben, bu yeni çağda veri güvenliğini sağlayacak protokollerin geliştirilmesinin zorunlu hale geldiğini vurgulamıştır. Kuantum güvenli bulut sistemleri, hem kuantum rastgele sayı üreticileri (QRNG) hem de homomorfik şifreleme gibi ileri düzey yöntemleri içermektedir (Chen vd., 2016).

Bu gelişmeler doğrultusunda, kuantum destekli bulut altyapılarının, gelecekte dijital veri güvenliği standartlarını yeniden tanımlayarak BB'nin evriminde önemli bir kilometre taşı oluşturması beklenmektedir.

Bahsedilen tüm bu gelişmeler, BB'nin güvenlik ve performans açısından nasıl dönüşüm geçirdiğini ortaya koymaktadır. Bir sonraki bölümde, bulut ortamlarında büyük veri analizi kapsamında önemli bir yere sahip olan metin madenciliği ele alınacaktır.

2.11. Metin Madenciliği ve Konu Modellemesi

Günümüzde dijitalleşmenin hız kazanmasıyla birlikte, üretilen verilerin büyük bir kısmı metin tabanlı ve yapılandırılmamış formda olması, geleneksel veri madenciliği yöntemlerinin yetersiz kalmasına neden olmuş ve metin odaklı özel analiz tekniklerinin gelişmesini zorunlu kılmıştır. Bu doğrultuda ortaya çıkan metin madenciliği (text mining), büyük hacimli metin verilerinden anlamlı bilgi ve örüntüler çıkarmayı amaçlayan çok disiplinli bir analiz alanıdır (Feldman ve Sanger, 2007).

Metin madenciliği, doğal dil işleme (NLP) ve makine öğrenmesi gibi alanlardan yararlanarak metinlerin otomatik analizini sağlar. NLP, metinlerin dilbilgisel ve anlamsal açıdan işlenmesini sağlarken; makine öğrenmesi, sınıflandırma, kümeleme ve konu modelleme gibi görevlerde etkin şekilde kullanılmaktadır (Manning ve Schütze, 1999).

2.11.1. Metin Madenciliği

Günümüzde giderek artan veri miktarının önemli bir kısmını metin verileri oluşturmaktadır. Bu durum, bireylerin bilgi aktarımında doğal dilleri kullanmalarından kaynaklanmaktadır. Yapılandırılmış veri analizinde yaygın olarak kullanılan geleneksel veri madenciliği yöntemleri, metin verileri için yetersiz kalmaktadır. Bu nedenle, metinlerin analizine özgü yöntem ve araçlar geliştirilmiş; böylece metin madenciliği olarak adlandırılan araştırma alanı ortaya çıkmıştır (Feldman ve Sanger, 2007).

Metin madenciliği, büyük ölçekli ve yapılandırılmamış metin verilerini işleyerek bu verilerde gizli bilgileri açığa çıkarmayı, örüntüleri belirlemeyi ve anlamlı sonuçlar üretmeyi hedefleyen bir veri analizi yöntemidir. Bu süreç, doğal dil işleme (NLP), makine öğrenmesi (ML) ve derin öğrenme (DL) gibi çağdaş teknolojilerden yararlanarak ham metinleri analiz edilebilir ve yorumlanabilir bir forma dönüştürmektedir (Feldman ve Sanger, 2007).

Metin madenciliğinde kullanılan temel teknolojilerden biri olan NLP, metinleri dilbilgisel ve anlamsal açıdan analiz ederek makinelerin insan dilini anlamasını sağlamaktadır (Jurafsky ve Martin, 2020). ML, etiketli veya etiketlenmemiş verilerden öğrenerek metin sınıflandırma, duygu analizi ve konu modelleme gibi görevlerde etkin biçimde kullanılmaktadır (Manning ve Schütze, 1999). DL tabanlı yöntemler ise özellikle sinir ağları ile desteklenerek metin madenciliği alanında önemli ilerlemelere yol açmış, büyük veri kümelerinden öğrenme kapasitesi sayesinde analiz doğruluğunu önemli ölçüde artırmıştır (Lauriola vd., 2022).

Günümüzde yaygın olarak kullanılan BERT, GPT ve Transformer tabanlı modeller, gelişmiş dil anlama ve bağlamsal analiz yetenekleri ile metin madenciliğinde yüksek performans sağlamaktadır (Young vd., 2018). Bu gelişmeler sayesinde metin madenciliği, sadece teorik bir araştırma alanı olmaktan çıkmış, birçok sektörde somut uygulama alanı bulmuştur. Sonraki bölümde, metin madenciliği tekniklerinin farklı alanlardaki uygulamalarına odaklanılarak, bu teknolojilerin pratikte nasıl kullanıldığı incelenecektir.

2.11.2. Metin Madenciliğinin Uygulama Alanları

Günümüzde verilerin önemli bir bölümü metin formatında olup, metin madenciliği teknikleri; sosyal medya içerikleri, akademik makaleler, müşteri geri bildirimleri, finansal dokümanlar, elektronik postalar ve tıbbi kayıtlar gibi çok çeşitli alanlarda yaygın şekilde kullanılmaktadır (Aggarwal ve Zhai, 2012). Ayrıca metin madenciliği, bilgi çıkarımı, konu takibi, duygu analizi, metin sınıflandırma ve tahmine dayalı analizler gibi yöntemlerle birçok sektörde aktif olarak uygulanmaktadır (Weiss vd., 2010).

Kızılkaya (2018), metin madenciliği tekniklerinin sosyal medya içeriklerinde özellikle duygu analizi amacıyla kullanıldığını ve bu tekniklerin müşteri ilişkileri yönetimi ile Pazar araştırmaları alanlarında etkin sonuçlar sunduğunu belirtmektedir. Liu (2015), şirketlerin müşteri yorumlarını analiz ederek pazarlama stratejilerini geliştirebileceğini vurgularken; Okatan ve Çam (2024), dijital bankacılık uygulamalarında toplanan müşteri geri bildirimlerinin metin madenciliği ile analiz edilmesinin hizmet kalitesine önemli katkılar sağladığını ifade etmektedir.

Sağlık sektöründe ise metin madenciliği, hasta kayıtlarının analizi yoluyla hastalık teşhis süreçlerinin iyileştirilmesine destek sağlamaktadır. Bu bağlamda Kurban (2021), tıbbi tedavi alanlarının belirti bazlı analizinde metin madenciliği ve kümeleme algoritmalarını kullanarak branşlar arası benzerlikleri ortaya koymuştur.

Eğitim alanında Karakuş (2021), öğrenci geri bildirimlerinin sözlük tabanlı duygu analizi yoluyla değerlendirilmesinin öğretim süreçlerinin gelişimine katkı sunduğunu ortaya koyarken; Torun ve Şengül (2022), kripto para birimleriyle ilgili sosyal medya paylaşımlarını metin madenciliği teknikleriyle analiz ederek yatırımcı davranışlarını incelemiştir. Benzer şekilde Şirzad (2023), Black Friday indirim kampanyalarının tüketici karar süreçleri üzerindeki etkisini metin madenciliği yöntemleri aracılığıyla analiz etmiştir.

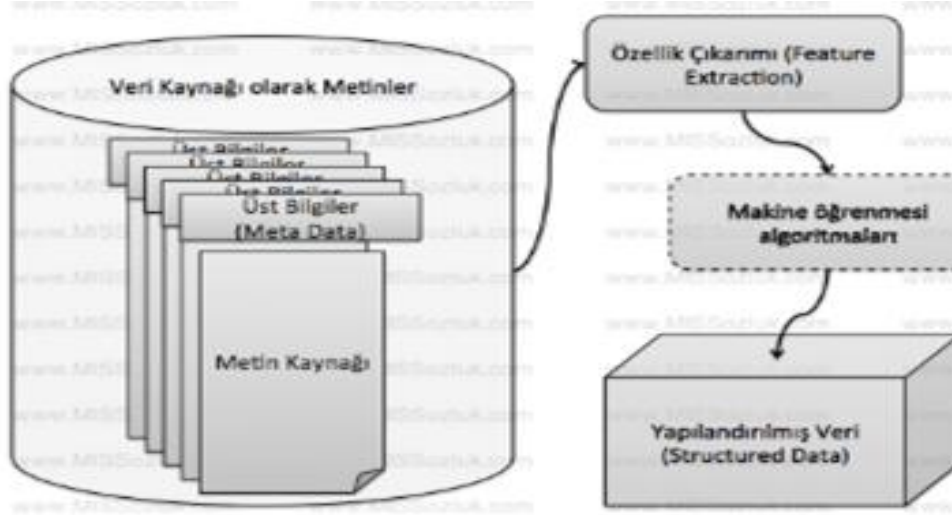
Bu geniş kullanım alanları ve sunduğu katkılar doğrultusunda, metin madenciliği sürecinin etkili biçimde yürütülebilmesi için belirli analiz aşamalarının sistemli şekilde izlenmesi gerekmektedir. Bir sonraki bölümde, bu temel aşamalar detaylı olarak ele alınacaktır.

2.11.3. Metin Madenciliğinin Aşamaları

Metin madenciliği, yapılandırılmamış metin verilerinden anlamlı bilgi çıkarma sürecinde belirli adımları içerir. Bu süreç; veri toplama, önışleme, özellik çıkarımı, modelleme (veri madenciliği) ve görselleştirme gibi aşamalardan oluşmaktadır. Her bir aşama, metin verilerinin analiz edilebilir ve yorumlanabilir hale getirilmesinde kritik bir rol oynar (Şeker, 2015; Sumathy ve Chidambaram, 2013).

Bu aşamaların doğru ve sistematik şekilde uygulanması, elde edilen analiz sonuçlarının doğruluğunu, güvenilirliğini ve tutarlılığını doğrudan etkilemektedir. Her aşamada yapılan algoritma ve parametre seçimleri, elde edilen sonuçların anlamlılığını belirlemekte; eksik, hatalı veya önyargılı işlemler ise analiz sürecinin güvenilirliğini zayıflatabilmekte ve yanlış çıkarımlara neden olabilmektedir. Bu nedenle, veri işleme sürecinin her adımı titizlikle planlanmalı ve kullanılan yöntemler, analiz hedeflerine uygun olarak belirlenmelidir.

Bu çalışmada, metin madenciliği süreçlerinin sistematik bir şekilde uygulanabilmesi için söz konusu aşamalar takip edilecek ve elde edilen veriler analiz edilerek yorumlanacaktır. Analiz süreçlerinin tutarlılığını ve güvenilirliğini sağlamak amacıyla, veri önışleme adımlarına özellikle dikkat edilecek ve modelleme sürecinde amaca en uygun yöntemler seçilecektir. Şekil 4'te metin madenciliği sürecine ait temel aşamalar görsel olarak sunulmuştur.



Şekil 4. Metin madenciliği sürecinde veri işleme ve yapılandırma adımları (Şeker, 2015).

Veri Toplama: Metin madenciliği sürecinin ilk adımını oluşturan veri toplama, analiz yapılacak metinlerin sosyal medya paylaşımları, müşteri geri bildirimleri, akademik makaleler, e-postalar ve kurumsal belgeler gibi çeşitli kaynaklardan elde edilmesini kapsamaktadır. Bu aşamada, veri setinin güvenilir olması, analiz edilecek konuyu yeterince temsil etmesi ve çalışmanın amacına uygunluk göstermesi oldukça büyük önem taşır. Verilerin doğru kaynaklardan, amacına uygun biçimde toplanması, elde edilecek sonuçların geçerliliği açısından kritik bir rol oynamaktadır (Aggarwal ve Zhai, 2012).

Veri toplama, metin madenciliği sürecinin sadece ilk adımı değil, aynı zamanda en temel ve yönlendirici aşamalarından biridir. Bu adımda yapılacak tercih ve yöntemler, analiz sonuçlarının doğruluğunu ve güvenilirliğini doğrudan etkiler. Özellikle büyük hacimli veri kümeleriyle çalışıldığında, verilerin eksiksiz, düzenli ve sistemli bir şekilde toplanması kritik hale gelir. Bu süreçte, veri çeşitliliği ve kalitesi; ortaya çıkacak sonuçların kapsamlı ve anlamlı olmasını sağlarken, eksik bilgiler, çelişkili içerikler ya da gürültülü veriler gibi sorunlar önileme aşamasında dikkatle ele alınmalı ve bu olumsuz etkiler en aza indirgenmelidir (Liu, 2015).

Önişleme: Genellikle tutarsız, gürültülü ve analize uygun olmayan bir yapıda bulunan ham metin verilerinin analize hazır hale getirilmesi, metin madenciliği sürecinde kritik öneme sahiptir. Bu aşamada; gereksiz karakterlerin temizlenmesi, büyük-küçük harf dönüşümleri, noktalama işaretlerinin kaldırılması, dilsel dolgu (stop-words) kelimelerin çıkarılması ve kök bulma (stemming) gibi işlemler gerçekleştirilerek metnin daha yapılandırılmış bir forma dönüştürülmesi sağlanır (Manning ve Schütze,

1999). Bu işlemler, doğal dil işleme ve veri analizi süreçlerinin daha sistematik ve tutarlı biçimde yürütülmesine olanak tanımaktadır.

Özellikle büyük ölçekli veri kümelerinde ön işleme adımının doğru ve titizlikle uygulanması, analiz süreçlerinin güvenilirliği ve doğruluğunu doğrudan etkilemektedir. Bu bağlamda, NLTK ve spaCy gibi yaygın doğal dil işleme araçları, kelime köklerine indirgeme (lemmatization), stop-word kaldırma ve sözcük türü etiketleme (POS tagging) gibi işlemler için etkin biçimde kullanılmaktadır (Bhagwat ve Pandit, 2024). Bu araçlar, metinlerdeki dilsel karmaşıklığı azaltarak, doğal dil işleme temelli analizlerin daha sağlıklı yürütülmesine katkı sağlamaktadır.

Özellik Seçme: Metin madenciliği sürecinde kritik bir aşama olarak kabul edilen özellik seçme metinlerin sayısal temsillerini oluşturmak amacıyla uygulanmaktadır. Kelime frekansı analizi, n-gram modellemeleri ve kelime gömme (word embedding) yöntemleri bu süreçte en çok kullanılan teknikler arasında yer almaktadır. Özellikle Word2Vec, GloVe ve FastText gibi kelime temsilleri, metinlerin vektörel forma dönüştürülmesini sağlayarak makine öğrenmesi algoritmalarının bu verileri daha anlamlı biçimde işleyebilmesine olanak tanır (Young vd., 2018).

Kelime gömme yöntemleri, kelimeler arasındaki anlamsal ve bağlamsal ilişkileri yakalayabilme yetenekleri sayesinde, klasik kelime frekansı temelli modellerin ötesine geçerek derin öğrenme tabanlı doğal dil işleme uygulamalarında yüksek başarı elde edilmesini sağlamaktadır. Young vd. (2018), bu yöntemlerin hem gömülü katmanlarda kullanılarak sistematik özellik çıkarımı sağladığını hem de düşük boyutlu, anlam yüklü temsillerle veri setinin boyutunu küçülttüğünü belirtmektedir. Böylece, yüksek doğrulukla çalışan ve daha az kaynak tüketen metin madenciliği modelleri geliştirilebilmektedir.

Veri Madenciliği: Veri madenciliği aşamasında, ön işleme ve özellik çıkarımı işlemleri tamamlanmış metin verileri; sınıflandırma, duygu analizi, konu modelleme ve ilişki keşfi gibi yöntemler kullanılarak anlamlı bilgiye dönüştürülür. Bu aşama, metin madenciliği süreçlerinin en kritik basamaklarından biridir ve ortaya çıkan bulgular, özellikle karar verme süreçlerinde değerli içgörüler sunar (Weiss vd., 2010).

Veri madenciliği tekniklerinin doğru uygulanması, metinlerdeki kalıpların ve eğilimlerin ortaya konmasını kolaylaştırır. Örneğin, duygu analizi; müşteri yorumlarında olumlu ya da olumsuz eğilimleri belirleyerek kurumların pazarlama ve müşteri ilişkileri stratejilerini şekillendirmelerine katkı sağlar. Benzer şekilde, konu modelleme yöntemler özellikle (LDA gibi algoritmalar) büyük hacimli metinlerdeki ana

temaları belirleyerek içerik sınıflandırmasını kolaylaştırır ve bilgi erişimini optimize eder (Blei vd., 2003).

Bu yöntemlerin kullanımı, sosyal medya analizinden akademik literatür taramasına; sağlık kayıtlarından finansal rapor analizine kadar birçok alanda yaygınlaşmıştır. Özellikle büyük veri ortamlarında, bu teknikler bilgiye dayalı karar alma süreçlerini destekleyen önemli araçlardır.

Görselleştirme: Görselleştirme, metin madenciliği sürecinin son aşamasını oluşturarak, analiz edilen verilerin anlaşılabilir ve etkili bir biçimde sunulmasına olanak tanır. Bu aşamada; kelime bulutları, ağ grafikleri ve etkileşimli panolar gibi araçlar kullanılarak, karmaşık ve büyük boyutlu metin verileri daha kolay yorumlanabilir hale getirilir (Liu, 2015). Görsel öğeler, verilerdeki desenleri ve ilişkileri ortaya koyarak, kullanıcıların bilgiye dayalı kararlar almasını kolaylaştırır.

Kelime bulutları ve grafiksel sunumlar, metinlerde öne çıkan anahtar kelimeleri ve kavramsal ilişkileri belirginleştirirken; etkileşimli panolar, kullanıcıların veriye farklı açılardan bakarak daha derinlemesine analiz yapmalarına olanak tanır. Bu görselleştirme teknikleri, yalnızca sonuçların etkili biçimde sunulmasını değil; aynı zamanda daha geniş bir kitle tarafından anlaşılmasını da sağlamaktadır (Heer vd., 2010).

Bu çalışmada, BB alanındaki akademik yayınlar metin madenciliği teknikleriyle analiz edilmiştir. Süreç boyunca veri toplama, ön işleme, özellik seçme, veri madenciliği ve görselleştirme aşamaları sistematik şekilde uygulanmıştır. Elde edilen sonuçların anlamlı bileşenlere ayrılması amacıyla bir sonraki bölümde LDA modeli üzerinden konu modelleme yapılacak; BB alanında öne çıkan başlıklar tespit edilecektir.

2.12. Konu Modelleme (Topic Modelling)

Bir önceki bölümde ele alınan metin madenciliği teknikleri, büyük ölçekli metin veri kümelerinden anlamlı bilgilerin elde edilmesini sağlayan temel analiz araçlarını kapsamaktadır. Bu teknikler arasında yer alan konu modelleme (topic modeling), yapılandırılmamış metinlerde yer alan gizli temaları ortaya çıkararak belgelerin tematik olarak sınıflandırılmasına imkân tanıyan istatistiksel bir yaklaşımdır. Bu çalışmada, BB alanındaki akademik yayınların analizi amacıyla, Blei ve arkadaşları (2003) tarafından geliştirilen LDA yöntemi kullanılacaktır.

2.12.1. Konu Modelleme: Tanımı, Önemi ve Uygulama Alanları

Konu modelleme, büyük ölçekli ve yapılandırılmamış metin belgeleri içerisindeki gizli temaları belirlemeye yönelik geliştirilen olasılıksal bir analiz yaklaşımıdır. Bu

yöntem, özellikle çok sayıda belgeden oluşan metin koleksiyonlarında, belgelerin içeriklerine göre tematik olarak sınıflandırılmasına olanak tanır. Bağlamsal farklılıklar ve kelimelerin çok anlamlılığı gibi dilsel karmaşıklıkların üstesinden gelinmesini sağlayan konu modelleme, bilgiye daha hızlı ve anlamlı şekilde erişim sağlamak amacıyla kullanılmaktadır (Blei vd., 2003).

Makine öğrenmesine dayalı bu yöntem, metin verilerinde tekrar eden örüntüleri keşfederek, belgelerde açıkça ifade edilmeyen ancak içerik bağlamında var olan temaları belirleyebilme yeteneğine sahiptir. LDA modeli, bu kapsamda en sık kullanılan yöntemlerden biridir. LDA, her belgenin çeşitli konuların karışımı olduğunu ve her konunun ise belirli bir kelime dağılımına sahip olduğunu varsayan bir olasılık modelidir (Blei vd., 2003). Bu model sayesinde, çok sayıda belgeden oluşan veri setlerinde hangi konuların öne çıktığı ve bu konuların belgelerle ilişkisi sistematik olarak analiz edilebilmektedir.

Konu modelleme yöntemleri; biyomedikal araştırmalardan pazarlamaya, sosyal medya analizinden bilgi yönetim sistemlerine kadar birçok alanda etkili şekilde kullanılmaktadır. Ayrıca büyük veri ortamlarında, karar destek sistemlerine entegre edilerek anlamlı bilgi çıkarımı ve içerik organizasyonu gibi işlemleri kolaylaştırmaktadır (Srivastava ve Sutton, 2017).

Bu çalışmada, BB alanındaki akademik yayınlardan oluşan belge kümesi üzerinde konu modelleme yöntemleri uygulanarak, bu alandaki temel araştırma eğilimlerinin ve içeriksel örüntülerin ortaya konması amaçlanmaktadır. Analizlerde LDA modeli kullanılacak olup, elde edilen temalar üzerinden literatürün yapısı yorumlanacaktır.

2.12.2. Konu Modelleme Yöntemleri

Konu modelleme, metin verilerindeki gizli temaları belirlemek ve belgeler arasındaki ilişkileri analiz etmek için kullanılan çeşitli yöntemleri içeren bir süreçtir. Bu yöntemler, modelin istatistiksel yaklaşımına ve kullanım amacına bağlı olarak farklı avantajlar ve kısıtlılıklar içermektedir (George ve Birla, 2018). Tablo 4'te, farklı konu modelleme yöntemleri, bunların temel özellikleri ve sınırlılıkları karşılaştırmalı olarak özetlenmiştir.

Tablo 4. Konu modelleme yöntemlerinin özellikleri ve sınırlamaları (George ve Birla, 2018, derlenmiştir).

Konu Modelleme Yöntemi	Özellikleri	Sınırlamaları
Latent Semantic Analysis (LSA)	1. Eş anlamlı kelimeleri belirleyebilir.	1. Konu sayısını belirlemek zordur.
	2. Kelimeler arasındaki ilişkileri analiz ederek metinlerin semantik temsillerini oluşturur.	2. Olasılıksal anlamda sağlam bir istatistiksel altyapıya sahip değildir.
	3. Metin madenciliği, metin sınıflandırma, metin kümeleme ve öneri sistemlerinde kullanılır.	3. Kullanımı karmaşıktır ve yükleme değerlerini olasıksal anlamda yorumlamak zordur.
Probabilistic Latent Semantic Analysis (PLSA)	1. Çok anlamlı kelimeleri ve bağlamı daha iyi ayırt edebilir.	1. Belge düzeyinde olasıksal modelleme yapamaz.
	2. Her kelime tek bir konudan türetilirken, bir belgedeki birçok kelime farklı konulardan gelebilir.	2. Aşırı öğrenmeye (overfitting) karşı hassastır.
	3. Dokümanlardaki konuların belirlenmesi için kelime dağılımlarını modelleyerek daha hassas sonuçlar sunar.	
Latent Dirichlet Allocation (LDA)	1. Her belgenin birden fazla konu içerdiğini varsayar.	1. Stop-word'lerin manuel olarak kaldırılması gerekir.
	2. Konu dağılımlarını olasıksal olarak belirler.	2. Konular arasındaki ilişkileri doğrudan gösteremez (CTM ile çözülebilir).
	3. Büyük ölçekli metin koleksiyonlarında yaygın olarak kullanılır.	3. Çanta kelimeler (bag-of-words) yaklaşımını kullanır, cümle yapısını modellemez.
Correlated Topic Model (CTM)	1. Konular arasındaki ilişkileri modelleyebilir.	1. Yüksek hesaplama maliyetine sahiptir.
	2. Kelimelerin birden fazla konuyla ilişkili olmasına izin verir.	2. Fazla genel/kapsamlı kelimeler içerebilir.
	3. Sosyal ağ analizi gibi konuların ilişkili olduğu veri kümelerinde tercih edilir.	
HDP (Hierarchical Dirichlet Process)	1. Konu sayısının otomatik olarak belirlenmesine olanak tanır.	1. Yüksek hesaplama maliyeti gerektirir.
	2. Konuların genişleyen veri kümeleriyle dinamik olarak uyumlanmasını sağlar.	2. LDA'ya kıyasla yorumlaması daha zordur.

Tablo 4'te özetlenen konu modelleme yöntemleri, farklı metin analizi ihtiyaçlarını karşılamak üzere geliştirilmiş çeşitli yaklaşımları içermektedir. Bu yöntemler,

yapılandırılmamış metin verilerinde gizli temaları ortaya çıkarma ve belgeler arasındaki semantik ilişkileri analiz etme süreçlerinde etkili araçlar sunmaktadır. Sonraki bölümde, bu yöntemlerden en sık kullanılanlar detaylı olarak ele alınmakta ve her birinin temel işleyiş prensipleri açıklanmaktadır.

Latent Semantic Analysis (LSA). Latent Semantic Analysis (LSA), metinlerin semantik yapısını istatistiksel yöntemlerle analiz ederek, kelimeler ve belgeler arasındaki gizli anlamsal ilişkileri ortaya çıkarmayı hedefleyen bir konu modelleme tekniğidir. LSA, geniş metin kümelerinde ortak kelimeleri ve bunların belgelerdeki dağılımlarını analiz ederek, içerik bazlı benzerlikleri ortaya çıkarmayı mümkün kılar (Deerwester vd., 1990). Bu amaçla, kelime-belge matrisinden elde edilen değerlerin tekil değer ayrıştırması (Singular Value Decomposition - SVD) ile boyutunun azaltılması sağlanarak, düşük boyutlu ve anlam yüklü vektörel temsiller oluşturulur (Alghamdi ve Alfalqi, 2015; Landauer vd., 1998).

NLP, bilgi erişimi, metin madenciliği ve öneri sistemleri gibi birçok alanda başarıyla uygulanan LSA yöntemi, yalnızca kelime sıklıklarına değil, kelimelerin bağlam içerisindeki kullanımına da odaklandığı için klasik “bag-of-words” modellerine kıyasla daha anlamlı sonuçlar üretebilmektedir (Aggarwal ve Zhai, 2012; Chen vd., 2011).

Uygulama alanları açısından LSA, sosyal medya analizlerinde duygu tespiti, haber gruplarının otomatik sınıflandırılması, belge kümelerinin kümeleme analizi ve içerik öneri sistemlerinin iyileştirilmesi gibi çok çeşitli alanlarda etkin şekilde kullanılmaktadır. Örneğin, Shah ve Sebastian (2020), COVID-19 pandemisi döneminde Hindistan hükümetinin twitter paylaşımlarını analiz ettikleri çalışmalarında LSA yöntemini kullanarak kriz iletişimi bağlamındaki toplumsal tepkileri incelemiştir. Benzer şekilde, Qomariyah vd. (2024), sosyal medya verileri üzerinde gerçekleştirdikleri çalışmada LSA ve LDA modellerini karşılaştırarak tematik içerik çözümlemesi yapmıştır.

LSA'nın uygulama potansiyeli yalnızca sosyal medya ile sınırlı değildir. Nesnelerin İnterneti (IoT) teknolojileriyle entegre edilen insan kaynakları yönetimi sistemlerinde, çalışan hareketliliğini ve görev verimliliğini ölçmek amacıyla da kullanıldığı görülmektedir (Jirasatjanukul vd., 2019). Arama motorları örneğinde ise, LSA'nın sunduğu semantik analiz yetenekleri, kullanıcı sorgularının daha doğru anlamlandırılmasını sağlayarak bilgi erişim kalitesini artırmaktadır (Chen vd., 2011).

Tüm bu avantajlarına rağmen, LSA'nın bazı sınırlamaları da bulunmaktadır. Öncelikle, kelime sırasını dikkate almayan bag-of-words yaklaşımına dayalı olması

nedeniyle bağlamsal ilişkileri tam olarak yansıtamayabilir. Ayrıca, konuların sayısının önceden belirlenmesi gerekliliği ve olasılıksal bir yapı sunmaması, modelin esnekliğini azaltmaktadır (Landauer vd., 1998). Dumais (2004) ve Wiemer-Hastings vd. (2004), özellikle yükleme değerlerinin yorumlanmasının uzmanlık gerektirdiğini ve çok anlamlı kelimelerin doğru biçimde temsil edilmesinde zorluklar yaşanabileceğini ifade etmektedir.

Bu nedenlerle, LSA'nın sınırlamalarını aşmak amacıyla Probabilistic Latent Semantic Analysis (PLSA) ve Latent Dirichlet Allocation (LDA) gibi gelişmiş modeller geliştirilmiştir. Bu modeller, kelimeler ile belgeler arasındaki ilişkileri olasılıksal çerçevede ele alarak daha tutarlı ve anlamlı konu modellemeleri yapılmasını mümkün kılmaktadır.

Probabilistic Latent Semantic Analysis (PLSA). LSA yönteminin daha gelişmiş bir uzantısı olan Probabilistic Latent Semantic Analysis (PLSA), çok anlamlılık (polysemy) ve bağlamsal farklılıklar gibi dilsel sorunları ele almak üzere geliştirilmiş bir olasılıksal modelleme tekniğidir. PLSA, kelimelerin belgeler içerisinde belirli konuların olasılıksal karışımı üzerinden üretildiğini varsayarak, metinlerin semantik yapısını daha hassas biçimde modellemeyi mümkün kılar (Alemayehu ve Fang, 2024; Hofmann, 2001).

Bu yöntemin temel avantajlarından biri, her kelimenin yalnızca tek bir bağlamda değerlendirilmemesi; aksine farklı konulardan türemiş olabileceğinin kabul edilmesidir. Böylece PLSA, belgelerdeki konu yapısını çok daha sistematik bir biçimde çözümleyebilir (Hofmann, 2001). Bu yapı, metinlerin semantik olarak sınıflandırılması ve içerik temelli analizlerde daha etkili sonuçlar elde edilmesine olanak sağlar.

Özellikle arama motorlarında PLSA, kullanıcının sorgularını yalnızca anahtar kelimeler üzerinden değil, semantik bağlamlarını da dikkate alarak yorumlamayı sağlar. Bu da daha anlamlı ve bağlamsal olarak ilgili sonuçların sunulmasına katkı verir (Alemayehu ve Fang, 2024). Google gibi büyük ölçekli sistemlerde PLSA, bu amaçla temel modelleme araçlarından biri olarak kullanılmaktadır.

PLSA'nın kullanım alanı sadece bilgi erişimiyle sınırlı değildir. Sosyal medya analizi ve duygu madenciliği gibi uygulamalarda da bu model etkin bir biçimde kullanılmaktadır. Ren ve Han (2014), kullanıcı yorumlarını analiz ederek marka algısını ve kamuoyu tepkilerini değerlendirmede PLSA'yı kullanmışlardır.

Ayrıca, multimedya içeriklerinin analizi ve özellikle video veritabanlarının düzenlenmesinde de PLSA yaygın biçimde kullanılmaktadır. Li vd. (2015), görüntü

sınıflandırma ve içerik benzerliği çıkarımı gibi görevlerde, PLSA'nın Fisher kernel teknikleriyle birleştirilerek başarılı sonuçlar sunduğunu ortaya koymuştur.

Bununla avantajların yanı sıra, PLSA'nın bazı önemli sınırlamaları da mevcuttur. Model, her belgeye özel ayrı konu dağılımları tahmin ettiğinden, büyük veri setleri üzerinde uygulandığında hesaplama maliyetleri artmakta ve eğitim süresi uzamaktadır (Steyvers ve Griffiths, 2007). Ayrıca, modelin sadece kelime düzeyinde olasılıksal yapıya sahip olması ve belge düzeyinde tümevarım yapamaması, PLSA'nın genelleme gücünü sınırlayan bir diğer faktördür (Blei vd., 2003; Hofmann, 1999).

Bu nedenlerle, PLSA'nın eksikliklerini gidermek ve daha sağlam belge düzeyinde çıkarım yapabilmek amacıyla LDA gibi daha gelişmiş modeller geliştirilmiştir.

Latent Dirichlet Allocation (LDA). LDA, konu modelleme alanında en yaygın ve etkili kullanılan yöntemlerden biridir. Blei, Ng ve Jordan (2003) tarafından geliştirilen bu olasılıksal model, büyük metin koleksiyonlarındaki gizli temaları keşfetmeye olanak sağlar. LDA, her bir belgenin birden fazla konuyu içerebileceğini ve her bir konunun belirli kelimelerden oluşan bir dağılım ile temsil edilebileceğini varsayarak, metinler arasındaki anlamsal ilişkileri derinlemesine analiz etmeye imkân tanır.

Bu çalışmada, LDA modeli kullanılarak BB alanındaki akademik literatür incelenecek ve belgelerde öne çıkan temalar ile bu temaların içeriksel bağlantıları ortaya konacaktır. LDA, literatürdeki temel araştırma eğilimlerini sınıflandırma, boşlukları belirleme ve alanın gelişimini daha sistematik şekilde yorumlama noktasında güçlü bir araç olarak kullanılmaktadır (Jelodar vd., 2019).

LDA, çok sayıda belgeden oluşan büyük veri kümelerinde konuların dağılımını modellemek için oldukça uygundur. Modelin en önemli avantajlarından biri, belgelerin yalnızca tek bir konuya değil, birden fazla konunun karışımına sahip olabileceğini kabul etmesidir. Bu durum, metinlerin daha esnek ve gerçekçi şekilde modellenmesini sağlar (Blei vd., 2003).

Tablo 4'te de belirtildiği gibi, LDA; metin sınıflandırma, belge kümeleme, öneri sistemleri, biyomedikal araştırmalar ve sosyal medya analizleri gibi çok çeşitli alanlarda yaygın olarak kullanılmaktadır (Jelodar vd., 2019). Model, yalnızca konu keşfi için değil, aynı zamanda bilgi erişimi ve içerik özetleme gibi görevlerde de önemli katkılar sunmaktadır.

Ancak, LDA'nın bazı sınırlamaları da mevcuttur. Model, belgeler arasındaki konular arası ilişkileri doğrudan modelleyemez ve bu durum, semantik yapının tam anlamıyla yansıtılmasını engelleyebilir. Ayrıca, stop-word'lerin (bağlamsal anlam

taşımayan kelimeler) manuel olarak temizlenmesi gerekmekte, bu da ek önışleme adımlarını zorunlu kılmaktadır (Blei vd., 2003).

Bu tür sınırlamaların üstesinden gelmek amacıyla Correlated Topic Model (CTM) gibi gelişmiş yapılar önerilmiş olup, CTM'nin konular arası ilişkileri modelleme konusundaki avantajları bir sonraki bölümde ele alınacaktır.

Correlated Topic Model (CTM). LDA modelinin konular arasındaki ilişkileri doğrudan modelleyememesi gibi sınırlılıklarını aşmak amacıyla Blei ve Lafferty (2007) tarafından önerilen bu yöntem, lojistik normal dağılım kullanarak konular arasındaki bağıntıları hesaba katar ve böylece özellikle bağlantılı temalara sahip metin koleksiyonlarında daha doğru analizler yapılmasını sağlar.

Özellikle sosyal medya, kamuoyu araştırmaları ve büyük veri analizlerinde öne çıkan model, konuların birlikte ortaya çıkma eğilimlerini dikkate alması, temalar arası bağlamları belirlemede güçlü bir avantaj sunar. Gerli vd. (2024), geniş çaplı sosyal medya içeriklerinde CTM'yi kullanarak kullanıcıların görüş ve eğilimlerini konu bazlı olarak analiz etmiş ve bu yaklaşımın klasik LDA modellerine göre daha yüksek bağlamsal tutarlılık sağladığını göstermiştir.

Ayrıca multimedya içeriklerinin sınıflandırılmasında da kullanılmakta olan CTM, görsel verilerin içerdiği hareket kalıplarının daha derinlemesine modellenmesini sağlayarak, video tabanlı içeriklerin tematik ayrıştırılmasında anlamlı katkılar sunmaktadır (Alghamdi ve Gotoh, 2018). Bunun yanında, COVID-19 döneminde kamuya açık sosyal medya paylaşımlarının analizinde de CTM modeli, halkın farklı temalar etrafındaki duygu dağılımlarını ayrıntılı biçimde ortaya koymak için kullanılmıştır (Sarıman ve Muta, 2020).

CTM'nin avantajları kadar bazı sınırlamaları da bulunmaktadır. Özellikle konu sayısının önceden belirlenmesi gerekliliği, esnekliği azaltmaktadır. Ayrıca, lojistik normal dağılıma dayanan yapısı nedeniyle yüksek hesaplama gücü gerektirmesi, büyük veri kümeleriyle çalışırken uygulama maliyetlerini artırmaktadır (Blei ve Lafferty, 2007). Ayrıca, modellenen konuların bazen aşırı genel veya aşırı spesifik olabilmesi, içeriklerin yorumlanmasını karmaşık hale getirebilir.

Tablo 4'te de özetlendiği üzere, CTM; konular arasındaki ilişkileri modelleyebilmesi açısından avantajlıdır. Ancak, yüksek hesaplama maliyeti ve konu sayısının analiz öncesinde belirlenme zorunluluğu gibi faktörler, bu yöntemin geniş ölçekli veri kümelerinde kullanımını sınırlamaktadır.

Bu kısıtlamaların üstesinden gelmek için geliştirilen Hierarchical Dirichlet Process (HDP) modeli, konu sayısını otomatik olarak belirleyebilmesi sayesinde daha dinamik bir yaklaşım sunmaktadır.

Hierarchical Dirichlet Process (HDP). CTM, konular arası ilişkileri modelleme yeteneğine sahip olmakla birlikte, konu sayısının önceden belirlenmesi zorunluluğu ve yüksek hesaplama maliyeti gibi sınırlamalar içermektedir. Bu eksiklikleri gidermek amacıyla geliştirilen Hierarchical Dirichlet Process (HDP) yöntemi, konu sayısını veriye bağlı olarak otomatik şekilde belirleyen ve büyük ölçekli veri kümeleri için daha esnek bir modelleme süreci sunan olasılıksal bir yaklaşımdır (Teh vd., 2006).

HDP'nin en büyük avantajlarından biri, konu sayısının önceden belirlenmesi zorunluluğunu ortadan kaldırarak veri kümesinin doğal yapısına uygun bir model oluşturmastır. Özellikle akademik literatür analizi, sosyal medya madenciliği ve biyomedikal metin sınıflandırmasında, değişken bilgi yapılarına daha uyumlu sonuçlar sunduğu belirtilmektedir (Rijcken vd., 2022).

Modelin hiyerarşik yapısı sayesinde, hem bireysel belgelerde hem de genel belge koleksiyonunda konu dağılımları daha gerçekçi biçimde modellenabilmektedir. Zamana bağlı olarak değişen içeriklerde konuların kendiliğinden genişleyebilmesi, HDP'yi uzun vadeli analizler için avantajlı hâle getirmektedir (Blei ve Jordan, 2006).

Ancak bu güçlü yönlerine rağmen, HDP'nin bazı sınırlamaları da bulunmaktadır. Özellikle yüksek hesaplama maliyeti, uzun eğitim süresi ve paralel işlem gereksinimi gibi teknik kısıtlamalar, modelin büyük ölçekli veri setlerinde uygulanabilirliğini sınırlamaktadır (Teh vd., 2006; Newman vd., 2009).

Bu nedenle, söz konusu sınırlamalar dikkate alınarak bu çalışmada LDA yöntemi tercih edilmiştir. LDA, belirli sayıda konu üzerinden analiz yapılmasına imkân tanırken; düşük hesaplama maliyeti, yüksek ölçeklenebilirlik ve yaygın kullanım desteği gibi avantajları sayesinde özellikle akademik yayınların analizinde etkin biçimde kullanılabilir. Tablo 4'te de özetlendiği üzere; HDP'nin otomatik konu belirleme esnekliği, CTM'nin konular arası korelasyon modelleme kabiliyeti ve LDA'nın düşük hesaplama maliyeti ile ölçeklenebilirliği karşılaştırıldığında, LDA; akademik literatür analizinde dengeli ve uygulanabilir bir model olarak öne çıkmaktadır.

Bir sonraki bölümde, LDA yöntemi detaylı olarak ele alınacak; BB alanında yayımlanan akademik makaleler bu yöntemle analiz edilerek alandaki temel temalar ve araştırma eğilimleri ortaya konacaktır.

3. LATENT DIRICHLET ALLOCATION (LDA) MODELİ VE BULUT BİLİŞİM YAYINLARININ ANALİZİNE YÖNELİK LİTERATÜR ÖZETİ

3.1. Akademik Yayınların Konu Modellemesinde LDA'nın Rolü

Bulut Bilişim (BB) alanındaki akademik yayınların hızla artması, literatürdeki tematik yapıların sistematik bir şekilde incelenmesini zorunlu hale getirmektedir. Bu doğrultuda, Latent Dirichlet Allocation (LDA) modeli; büyük ölçekli ve disiplinler arası metin verilerinden gizli konuları çıkararak, araştırma eğilimlerini haritalandıran etkili bir yöntem olarak öne çıkmaktadır (Blei vd., 2003).

LDA, konu dağılımlarını olasılıksal bir modelleme yöntemiyle analiz ederken, farklı araştırma alanlarını kümelere ayırarak, literatürdeki ilişkilerin daha açık ve anlamlı biçimde anlaşılmasına katkı sunmaktadır (Griffiths ve Steyvers, 2004). Bu yönüyle, SaaS, IaaS, güvenlik ve hibrit bulut gibi disiplinler arası konuları içeren BB araştırmalarının analizinde uygulanabilir bir yaklaşım sunmaktadır.

Ayrıca, LDA'nın zamansal eğilimleri izleyebilme yeteneği, belirli dönemlerde öne çıkan teknolojik gelişmeleri (örneğin sanallaştırma, kuantum bulut ve edge computing gibi) yayınlar üzerinden takip etme ve bu gelişmelerin literatürde nasıl yankı bulduğunu analiz etme açısından önemli bir potansiyel taşımaktadır (Newman vd., 2009).

Alternatif konu modelleme yöntemleri ile karşılaştırıldığında, LDA'nın akademik yayın analizindeki avantajları daha belirgin hâle gelmektedir. Bu karşılaştırmaya dair detaylar, bir önceki bölümde sunulan Tablo 4'te ayrıntılı şekilde ele alınmıştır. Özetle;

- PLSA modeli, aşırı öğrenme (overfitting) eğilimi nedeniyle akademik verilerde tutarsız sonuçlar üretmektedir,
- CTM modeli, konu korelasyonlarını modelleyebilse de lojistik normal dağılım yapısı nedeniyle yüksek hesaplama maliyeti doğurmaktadır,
- HDP modeli ise konu sayısını otomatik belirleme esnekliğine sahip olsa da, büyük veri kümeleri üzerinde çalışırken yüksek kaynak tüketimi nedeniyle pratik kullanımı sınırlı kalmaktadır (Blei ve Lafferty, 2006) (bkz. Tablo 4).

Tüm bu alternatiflerin karşısında, LDA ölçeklenebilir yapısı, düşük kaynak tüketimi ve paralel işlem uyumluluğu sayesinde, büyük hacimli veri analizlerinde öne çıkmaktadır. Scopus, IEEE Xplore ve SpringerLink gibi önde gelen akademik

platformlardan toplanan on binlerce makale, Dirichlet öncelikli yapısı sayesinde istikrarlı ve verimli biçimde analiz edilebilmektedir (Jelodar vd., 2019).

Bu çalışma kapsamında LDA modeli; BB literatürünün tematik yapısını haritalandırmak, zamansal eğilimleri ortaya koymak ve araştırma boşluklarını tespit etmek amacıyla kullanılmıştır. Modelin sunduğu bu stratejik analiz kabiliyeti, akademik üretimin dinamiklerini anlamada önemli bir rol üstlenmektedir. Bir sonraki bölümde, LDA'nın temel bileşenleri, genel yapısı ve nasıl çalıştığı ayrıntılı biçimde açıklanacaktır.

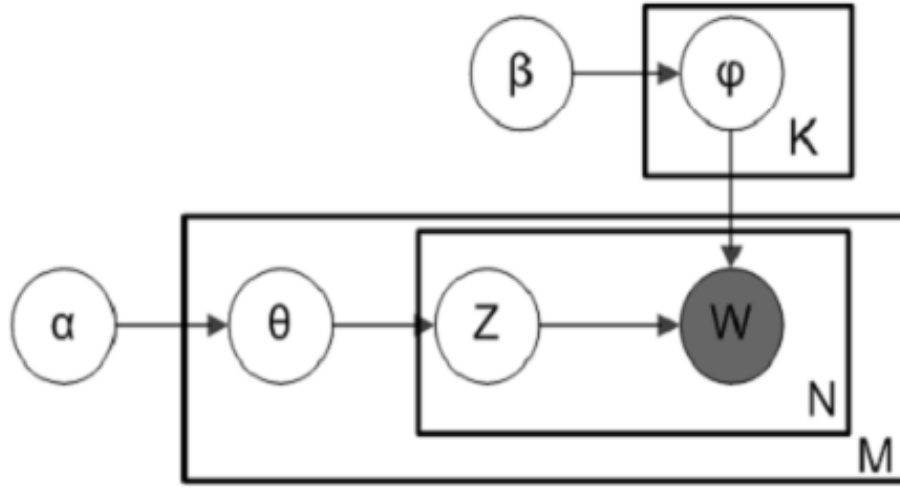
3.2. LDA Modeli ve Genel Yapısı

LDA, büyük ölçekli metin veri kümelerinde gizli tematik yapıları keşfetmek amacıyla geliştirilmiş olasılıksal bir modeldir (Blei, Ng ve Jordan, 2003). Bu model, belgeleri çeşitli konuların olasılıksal bir birleşimi olarak değerlendirirken, her konuyu da belirli kelimelerin dağılımıyla temsil eder (Griffiths ve Steyvers, 2004). LDA'nın sunduğu bu esnek yapı hem metin madenciliğinde hem de akademik literatür analizinde yaygın kullanım alanı bulmasını sağlamıştır (Jelodar vd., 2019).

LDA'nın temel prensibi, her belgenin belirli konulara farklı oranlarda sahip olduğu ve her konunun da belirli kelimelerden oluştuğu fikrine dayanır. Model, Dirichlet Dağılımı'nı kullanarak belgeler, konular ve kelimeler arasındaki ilişkileri olasılıksal olarak hesaplar (Blei, 2012). Örneğin, LDA modeliyle belirlenen 'bulut güvenliği' konusunun 'şifreleme' (encryption), 'kimlik doğrulama' (authentication) gibi anahtar kelimeler içermesi beklenirken, 'veri analitiği' (data analytics) temasının 'makine öğrenimi' (machine learning) ve 'büyük veri' (big data) gibi teknik terimlerle ilişkilendirilmesi mümkündür (Bkz. Tablo 7: LDA Modelinden Türetilen 20 Ana Konu).

LDA'nın çalışma prensibi, belge-konu-kelime hiyerarşisine dayanmaktadır. Her belge, Dirichlet dağılımı ile belirlenen bir konu karışımından oluşur ve her konu, belirli kelimelerin olasılık dağılımı ile ifade edilir (Steyvers ve Griffiths, 2007). Bu yapı sayesinde, akademik makalelerden haber içeriklerine kadar geniş bir metin yelpazesinde etkili bir analiz gerçekleştirilmesi mümkün olmaktadır.

LDA'nın grafiksel modeli, belgelerdeki kelimelerin birden fazla konuyla ilişkili olabileceğini ve her konunun belirli kelime dağılımlarına sahip olduğunu ifade eder. Şekil 5'te, LDA'nın temel bileşenleri ve modelin işleyişi genel hatlarıyla açıklanmaktadır.



Şekil 5. LDA için grafiksel model (Genç vd., 2024).

LDA modelinde kullanılan temel bileşenler şunlardır:

- **K - Konu Sayısı:** Modelin kaç farklı konu üreteceğini belirleyen hiperparametredir. Konu sayısı önceden belirlenir ve modelin performansı ile yorumlanabilirliği açısından kritik bir faktördür (Griffiths ve Steyvers, 2004).
- **θ (Theta) - Belgelerin Konu Dağılımı:** Her belge için konular üzerinde tanımlanan olasılık dağılımını ifade eder (Blei, Ng ve Jordan, 2003).
- **φ (Phi) - Konuların Kelime Dağılımı:** Her konu için kelimeler üzerinde tanımlanan olasılık dağılımıdır (Blei, Ng ve Jordan, 2003).
- **α (Alfa) - Belgelerin Konu Dağılımı Önceliği:** Bir belgedeki konu dağılımının yaygınlığını belirler. Düşük α değeri belgelerin belirli birkaç konuya odaklandığını, yüksek α değeri daha yaygın bir dağılım olduğunu gösterir (Wallach vd., 2009).
- **β (Beta) - Konuların Kelime Dağılımı Önceliği:** Konuların kelime dağılımını düzenleyen hiperparametredir. Düşük β değeri konuların daha özgül kelimelerle temsil edilmesini sağlar (Blei vd., 2003).

LDA, Dirichlet dağılımı üzerinden her belgenin birden fazla konuyla ilişkili olabileceği varsayımıyla çalışır. Bu çoklu konu aidiyeti sayesinde akademik metinlerin detaylı tematik analizi mümkün hale gelmektedir. Modelin performansı, modelin hiperparametrelerinin doğru belirlenmesine bağlıdır. Yüksek α değeri daha fazla konunun oluşmasına, düşük α değeri ise az sayıda baskın konunun ortaya çıkmasına neden olabilir (Genç vd., 2024; Blei vd., 2003).

LDA'nın esnek ve olasılıksal yapısı, akademik metin analizi, sosyal medya içerik keşfi ve büyük veri kümelerinin tematik analizi gibi birçok alanda güçlü bir araç

olmasını sağlamaktadır. Ancak, zaman içinde gelişen veri yapıları ve konu modelleme teknikleri, LDA'nın çeşitli evrimlerden geçmesine neden olmuştur.

3.3. LDA'nın Zaman İçindeki Evrimi

LDA'nın temel yapısı ve avantajlarının yanı sıra, modelin zaman içinde farklı alanlardaki gereksinimlere adapte edilerek nasıl evrildiğini incelemek, bu yöntemin neden bu kadar geniş kabul gördüğünü anlamak açısından önemlidir.

Model, ilk olarak Blei, Ng ve Jordan (2003) tarafından tanıtıldığında, metin madenciliği ve doğal dil işleme alanlarında büyük bir yenilik olarak kabul edilmiştir. Başlangıçta, büyük ölçekli metin koleksiyonlarını analiz etmek ve gizli konuları keşfetmek amacıyla geliştirilmiş olan model, zamanla çeşitli iyileştirmelerle daha esnek hale gelmiş ve farklı alanlara uyarlanmıştır.

Tablo 5'te görüldüğü üzere, LDA'nın temel modelinden hareketle, çok sayıda disiplinde özelleştirilmiş analiz uygulamaları dönüştürülmüş; bu sayede sağlık, hukuk, sosyal medya, enerji, pazarlama gibi pek çok alanda metin temelli büyük veri analizlerine etkin biçimde katkı sağlamıştır.

Örneğin, Xue ve arkadaşları (2020) çalışmasında, COVID-19 pandemisi süresince sosyal medya üzerinden kamuoyunun duygu ve söylemleri LDA yöntemiyle analiz edilerek halkın tepkileri tematik olarak sınıflandırılmıştır. Aynı veri Twitter özelinde incelendiğinde, pandemi döneminde toplumsal duygu değişimlerinin dijital ortamdaki yansımaları ayrıntılı olarak ortaya konmuştur.

Bilimsel yayınlar açısından ele alındığında, Gupta ve arkadaşları (2022), Uygulamalı Zekâ alanında yayımlanmış 30 yıllık akademik çalışmaları analiz etmiş ve araştırma eğilimlerini konu modellemesiyle tahmin etmiştir. Bu yöntem, bilimsel üretimin zaman içindeki dönüşümünü anlamada etkili bir araç olarak sunulmuştur.

Hukuk alanında ise Silveira ve arkadaşları (2021) tarafından gerçekleştirilen çalışmada, Amerikan Yüksek Mahkemesi kararları LEGAL-BERT destekli bir LDA yaklaşımıyla analiz edilmiş; yargı kararlarının tematik yapılarına ulaşılmıştır. Benzer biçimde, sosyal bilimler kapsamında Shi ve Min (2023), ESG (Çevresel, Sosyal ve Yönetişim) konularına dair haber içerikleri ile akademik yayınları karşılaştırmalı biçimde analiz etmiş; toplum gündemi ile akademik literatür arasındaki boşlukları ortaya koymuştur.

Kamu politikası ve enerji alanına odaklanan Pamula ve arkadaşları (2023), Polonya'daki üniversitelere ait kamu ihale belgelerini incelemiş ve üniversite binalarında enerji verimliliği konularını LDA yöntemiyle belirlemiştir. Çalışma, kamu

belgelerinden tematik çıkarım yapılmasının politika üretimi açısından önemini vurgulamaktadır. Pazarlama alanında yapılan çalışmalardan biri olan Çullu ve Arabelen (2023), pazarlama literatüründe yer alan 54 makaleyi analiz ederek, LDA'nın bu alandaki kullanımlarını bilimetrik yöntemlerle incelemiştir. Bu kapsamda, tüketici davranışları, marka konumlandırma ve sosyal medya pazarlaması gibi alt alanlarda LDA'nın etkili bir analiz aracı olduğu sonucuna ulaşılmıştır.

Yazılım ve eğitim bilişimi özelinde, Onan ve arkadaşları (2020) çalışması üniversite bilgi yönetim sistemine gelen 17.831 servis destek talebini LDA ile analiz etmiş; taleplerin metin temelli gizli yapıları ortaya konarak sistem yönetimi süreçlerine katkı sağlanmıştır.

Tablo 5. LDA yönteminin farklı alanlardaki akademik uygulamaları (Birden fazla kaynaktan derlenmiştir).

Uygulama Alanı	Örnek Çalışma / Model	Açıklama	Kaynak
Sağlık	COVID-LDA	COVID-19 süresince sosyal medya verileri üzerinden halkın duygu ve söylemlerinin LDA yöntemiyle tematik analizi yapılmıştır.	Xue vd. (2020)
Sosyal Medya / Twitter	COVID-LDA	Pandemi döneminde Twitter verileri kullanılarak kamuoyunun duygu durumları ve söylem eğilimleri LDA ile analiz edilmiştir.	Xue vd. (2020)
Bilimsel Yayın Analizi	LDA ile araştırma trendi tahmini	Uygulamalı zekâ alanında 30 yıl boyunca yayımlanmış makalelerden LDA ile temalar çıkarılmış, eğilim analizi yapılmıştır	Gupta vd. (2022)
Hukuk	LEGAL-BERT ile yargı metinleri analizi	ABD Yüksek Mahkemesi kararları, yasa atıfları ve paragraflar üzerinden LDA temelli analiz yapılmıştır	Silveira vd. (2021)
Sosyal Bilimler / Politika	ESG haber ve makale analizi	ESG haberleri ve akademik yayınlar LDA ile analiz edilerek, toplum gündemi ile araştırmalar arasındaki tematik farklar ortaya konmuştur.	Shi ve Min (2023)
Enerji / Kamu Politikası	Üniversite ihalelerinde enerji temaları analizi	Kamu üniversitelerindeki inşaat ihaleleri LDA ile analiz edilerek enerji verimliliği odaklı temalar belirlenmiştir	Pamula vd. (2023)

Tablo 5. (Devamı)

Pazarlama	LDA ile pazarlama literatürü analizi	LDA kullanılan 54 pazarlama araştırması bilimetrik analiz ve sistematik tarama ile incelenerek, konu modellemesinin kullanım alanları ve eğilimleri belirlenmiştir.	Çullu ve Arabelen (2023)
Yazılım / Eğitim Bilişimi	Üniversite bilgi yönetim sistemi (ÜBYS) servis taleplerinde konu modelleme	ÜBYS servis talepleri LDA ile analiz edilerek metin tabanlı temalar ve gizli örüntüler çıkarılmıştır.	Onan vd. (2020)

Sonuç olarak, Tablo 5'te özetlenen bu evrimsel süreç, LDA'nın farklı disiplinlerdeki ihtiyaçlara nasıl esnek bir şekilde cevap verebildiğini açıkça ortaya koymaktadır. Modelin bu esnek yapısı, LDA'nın temel bileşenlerinin ve işleyiş mekanizmasının detaylı şekilde anlaşılmasını daha da önemli hale getirmektedir. Bu nedenle bir sonraki bölümde, modelin içsel yapısı ve çalışma prensipleri daha ayrıntılı olarak ele alınacaktır.

3.4. LDA Modelinin Temel Bileşenleri

LDA, temelde Dirichlet dağılımına dayanan olasılıksal bir Bayes konu modelleme yöntemidir (Noble vd., 2021). Belge-konu ve kelime-konu dağılımlarını paylaşım esasına göre yapılandıran bu model, metin verilerinin analizinde ve yorumlanmasında önemli bir rol oynamaktadır. Belge-konu ve kelime-konu dağılımlarını paylaşım esasına göre yapılandıran bu model, metin verilerinin analizi ve yorumlanmasında önemli bir rol oynamaktadır.

Bu bölümde, LDA modelinin işleyişi dört temel başlık altında incelenmekte olup; Belge-Konu Dağılımı, Kelime-Konu Dağılımı, Olasılıksal Modelleme ve Dirichlet Dağılımı ile Gizli Değişkenler ve Bayesçi Yaklaşım başlıkları çerçevesinde modelin temel bileşenleri açıklanmaktadır. LDA'nın metin madenciliği süreçlerindeki işlevi ve sunduğu analitik avantajlar detaylı olarak ele alınmaktadır.

Belge-Konu Dağılımı. LDA, belgeleri çeşitli konuların olasılıksal karışımı olarak değerlendiren; her konuyu da belirli kelimelerin dağılımı üzerinden tanımlayan bir modeldir (Blei, 2012). Bu yapı, metinlerin tematik özelliklerini analiz edip anlamlandırmak amacıyla geliştirilmiştir. Modelin temel varsayımlarından biri, belgelerin bağımsız olduğu ve kelimelere sırasız kabul edildiği “Kelime Torbası (Bag of Words, BoW)” yaklaşımıdır.

BoW modeli, metinlerin içeriklerine göre analiz edilerek belirli konular altında gruplandırılmasına olanak tanır. Garg ve Kanjilal (2022) tarafından yapılan bir çalışmada, metin madenciliği teknikleriyle birlikte LDA uygulanarak kütüphanecilik alanında çevrim içi tartışma verileri analiz edilmiştir. Diğer bir çalışmada, Muppala ve Devi (2023), BoW algoritmasıyla metinleri işleyip, ön işleme adımları (tokenizasyon, stopword çıkarımı) sonrasında frekansa dayalı analiz gerçekleştirerek içerikleri konu temelli sınıflandırmıştır.

LDA'nın belge-konu dağılımlarını belirlemedeki başarısı yalnızca belgelerin hangi konulara odaklandığını değil, aynı zamanda belirli kelimelerin hangi temalarla ilişkili olduğunu da ortaya koymaktadır. Model, kelimeler ve konular arasındaki olasılıksal dağılımları analiz ederek, metinlerin semantik yapısının daha derinlemesine incelenmesine olanak tanır. Özellikle büyük hacimli veri setlerinde bu tür dağılımların ortaya konması, içerik analizini sistematik ve yorumlanabilir kılmaktadır (Blei, 2012).

Kelime-Konu Dağılımı. LDA, metinlerdeki tematik örüntüleri ortaya çıkarmak amacıyla kelimeler ile konular arasında olasılıksal bağlar kuran bir modelleme yaklaşımıdır (Blei, Ng ve Jordan, 2003). Bu yöntem, metinlerin içeriksel analizine olanak sağlayarak belge temelli konuların belirlenmesini ve metinlerin daha derinlemesine anlamlandırılmasını mümkün kılar (Yılmaz ve Polat, 2023).

Modelde her konu, belirli kelimelerin oluşturduğu bir olasılık dağılımıyla temsil edilir. Örneğin teknolojiyle ilişkili bir konuda, “yazılım”, “donanım” ve “internet” gibi kelimelerin yüksek olasılıkla yer alması beklenir. Bu durum, belgelerin konu temelli olarak sınıflandırılmasını kolaylaştırır.

Onan, Yalçın ve Atik (2020) tarafından yürütülen bir çalışmada, üniversite bilgi yönetim sistemi (ÜBYS) üzerine gelen binlerce servis destek talebi LDA yöntemiyle analiz edilmiş ve bu taleplerin kelime-konu dağılımları üzerinden tematik kümelere ayrıldığı görülmüştür. Model, öğrenci işleri, belge yönetimi, personel işlemleri ve akademik projeler gibi konulara ait yoğunlukların kelime düzeyinde ortaya konmasına yardımcı olmuştur.

Benzer şekilde, Łabuński ve arkadaşları (2023) tarafından gerçekleştirilen çalışmada, İngilizce yayımlanmış Psikolojik İlk Yardım (PFA) eğitim kılavuzları LDA yöntemiyle analiz edilmiş; travma sonrası müdahaleler, PTSD ve toplum temelli uygulamalar gibi temaların öne çıktığı görülmüştür. Bu analiz, akademik söylemin belirli kavram grupları etrafında nasıl şekillendiğini ve LDA'nın bu yapıyı çözümlemedeki etkinliğini göstermektedir.

Bu çalışmada da kelime-konu modeli, BB alanındaki literatürü analiz etmek için kullanılmıştır. Elde edilen sonuçlar; güvenlik, veri depolama, makine öğrenimi ve blok zincir gibi temaların sıklıkla tekrarlandığını göstermektedir. Bu temalar, BB altyapısından büyük veri çözümlerine kadar geniş bir kapsamı temsil etmektedir.

Sonuç olarak, kelime-konu dağılımı, literatürdeki temel kavramların konularla nasıl ilişkilendirildiğini ortaya koymakta; LDA'nın altında yatan olasılıksal modelleme yapısını ve Dirichlet dağılımının rolünü kavramak ise bu işleyişi daha iyi anlamayı sağlamaktadır. Bir sonraki bölümde bu kuramsal temeller detaylı olarak ele alınacaktır.

Olasılıksal Modelleme ve Dirichlet Dağılımı. Olasılıksal modelleme, büyük hacimli metin verilerinde gizli yapıları ortaya çıkarmak ve belgelerle konular arasındaki ilişkileri belirlemek amacıyla kullanılan etkili bir yöntemdir. Bu yaklaşımın temelinde yer alan LDA modeli, belgeleri çeşitli konuların olasılıksal karışımı olarak temsil ederken, her bir konuyu da belirli kelimelerin dağılımı ile tanımlar (Blei, Ng ve Jordan, 2003).

LDA'nın işleyişinde Dirichlet Dağılımı önemli bir rol oynar. Bu dağılım, modelin hem belge-konu hem de kelime-konu dağılımlarını tanımlamasına olanak tanır. Her belge, farklı konulara ait bir olasılık dağılımına sahipken; her konu da çeşitli kelimelerle temsil edilen bir dağılım içerir. Böylece model, metin içerisindeki tematik örüntüleri çıkararak çok boyutlu analiz yapılmasını sağlar (Syed ve Spruit, 2018).

Noble ve arkadaşları (2021) tarafından gerçekleştirilen bir çalışmada, LDA modeli veterinerlik kayıtlarından oluşan geniş ölçekli veriler üzerinde uygulanmış ve İngiltere'deki köpeklerde görülen hastalık kümeleri analiz edilmiştir. Bu çalışmada Dirichlet Dağılımı sayesinde semptom, tedavi ve coğrafi faktörler gibi öğeler belirli konulara başarılı şekilde gruplanmıştır. Elde edilen sonuçlar, LDA'nın sağlık, biyomedikal ve epidemiyolojik alanlarda metin temelli analizlerde ne denli etkili olduğunu göstermektedir.

Bu tez kapsamında da LDA temelli olasılıksal modelleme yöntemi kullanılarak BB literatürü analiz edilmiş, toplamda 20 tematik alan belirlenmiştir. Güvenlik, veri saklama, makine öğrenimi ve blok zincir gibi konuların öne çıktığı bu analizler sayesinde alanın mevcut eğilimleri değerlendirilmiş ve eksik araştırma alanları ortaya konmuştur.

Sonuç olarak, Dirichlet Dağılımı'nın sağladığı yapısal esneklik sayesinde LDA modeli, belgelerdeki tematik örüntüleri anlamlı bir biçimde sınıflandırmakta; araştırmacılara konular arası bağlamları görselleştirme ve yorumlama olanağı

sunmaktadır. Bir sonraki bölümde, bu modelin temelini oluşturan gizli değişkenler ve Bayesçi Yaklaşım detaylı olarak ele alınacaktır.

Gizli Değişkenler ve Bayesci Yaklaşım. Gizli değişkenler, doğrudan gözlemlenemeyen ancak verilerin altında yatan yapıları temsil eden kavramsal unsurlardır. Bu tür değişkenler, özellikle metin madenciliği gibi karmaşık veri kümelerinde tematik örüntüleri ortaya çıkarmak amacıyla kullanılır. Bayesci yaklaşım ise bu gizli değişkenlerin olasılık dağılımlarını tahmin ederek modelin performansını artıran istatistiksel bir çerçeve sunar (Syed ve Spruit, 2018).

LDA gibi konu modelleme yöntemleri, gizli değişkenleri temel alarak belgelerdeki kelimelerin belirli konular altında nasıl gruplandığını olasılıksal olarak modellemektedir. Bu yaklaşım, özellikle büyük hacimli metin veri kümelerindeki tematik yapıları analiz etmek için etkin bir yöntemdir (Blei, Ng ve Jordan, 2003).

Bayesci yöntemler, özellikle Markov Chain Monte Carlo (MCMC) ve Gibbs Sampling gibi tekniklerle karmaşık yapıların tahmin edilmesinde önemli avantajlar sağlar. De la Hoz-M., Fernández-Gómez ve Mendes (2021) tarafından geliştirilen LDAShiny uygulaması, bilimsel literatürün LDA temelli Bayesci yöntemlerle analizine olanak tanımakta ve bu süreçte gizli değişkenlerin hangi temaları temsil ettiğini anlamaya yardımcı olmaktadır. Uygulamada, Dirichlet Dağılımı kullanılarak belgelerdeki konuların ve konulardaki kelimelerin olasılık temelli yapısı modellenmektedir.

Bu tez kapsamında yapılan analizlerde de, LDA'nın gizli değişken yapısı kullanılarak BB alanındaki tematik yapı analiz edilmiştir. Belgelerin belirli konularla hangi oranda ilişkili olduğunu ortaya koyan bu yapılar, araştırma eğilimlerini belirlemede ve literatürdeki bilgi kümelerini sınıflandırmada etkili olmuştur.

Sonuç olarak, Bayesci yaklaşım ve gizli değişken temelli modelleme sayesinde, büyük metin veri kümelerinde içerik sınıflandırması sistematik ve yüksek doğrulukla gerçekleştirilebilmektedir. Bir sonraki bölümde, LDA'nın avantaj ve sınırlılıkları değerlendirilecek; bu yöntemin metin madenciliği uygulamalarındaki genel katkısı ele alınacaktır.

3.5. LDA Modelinin Avantajları ve Kısıtları

Model, metin madenciliği, biyoenformatik ve sosyal ağ analizi gibi çeşitli alanlarda yaygın olarak kullanılmaktadır. Modelin sunduğu avantajlar ve karşılaşılan zorluklar aşağıdaki tabloda özetlenmiştir.

Tablo 6. LDA modelinin avantaj ve dezavantajları (birden fazla kaynaktan derlenmiştir).

Kriter	Açıklama	Kategori	Kaynak
Büyük veri kümelerinde etkili analiz	Büyük hacimli metin veri kümelerinde gizli tematik örüntüleri yüksek doğrulukla tespit ederek içerik madenciliği ve veri analitiği süreçlerini optimize eder.	Avantaj	Albalawi vd. (2020); Du ve Liu (2021)
Olasılıksal yapı	Konuların kelimelerle ve belgelerle olasılık temelli ilişkilerini modelleyerek, veri içerisindeki belirsizliğe karşı esnek ve istatistiksel olarak güçlü analizler sunar.	Avantaj	Breuninger vd. (2021)
Gözetimsiz öğrenme	Etiketleme gerektirmeden büyük veri setlerinde analiz yapılmasını mümkün kılar.	Avantaj	Albalawi vd. (2020)
Çeşitli veri türlerine uygulanabilirlik	Farklı modalitelerdeki veriler (örneğin metin, görsel, sosyal medya içerikleri) üzerinde konu analizi yapılmasına olanak tanıyarak çeşitli disiplinlerde uygulanabilirlik sağlar.	Avantaj	Rasiwasia ve Vasconcelos (2013)
Konu belirlemede güçlü yapı	Belgelerdeki içerikleri temsil eden konuları otomatik olarak çıkartarak tematik haritalama ve bilgi yapısının görselleştirilmesi süreçlerinde etkili bir araç sunar.	Avantaj	Albalawi vd. (2020)
Parametre seçiminin zorlukları	Konu sayısı gibi parametrelerin dikkatlice belirlenmesi gereklidir; yanlış seçim aşırı uyum (overfitting) veya yetersiz uyum (underfitting) ile sonuçlanabilir.	Dezavantaj	Maier vd. (2018);
Kelime çantası modeli kısıtı	Kelimeleri bağımsız ve sırasız varsayan Bag-of-Words yaklaşımını temel aldığı için anlamsal bütünlük ve bağlamsal ilişkilerde bilgi kaybına neden olabilir.	Dezavantaj	(Syed ve Spruit, 2018).
Hesaplama maliyetleri	Büyük veri kümelerinde çalışırken yüksek işlem gücü gerektirir ve analiz süresi uzayabilir.	Dezavantaj	Yan vd. (2017)
Sonuçların yorumlanabilirliği	Üretilen konular her zaman insan sezgisine uymayabilir ve ek doğrulama gerektirebilir.	Dezavantaj	Maier vd. (2018)

LDA'nın avantajlarını ve kısıtlamaları Tablo 6'da özetlenerek modelin güçlü ve zayıf yönlerinin analiz edilmesine olanak tanınmıştır. Bu çerçevede, LDA'nın en büyük avantajlarından biri, büyük veri kümelerindeki gizli yapıları ortaya çıkarma yeteneğidir.

Model, metinsel verilerdeki ana temaları tespit edebilmek için belgeleri konuların karışımı olarak ele alır. Bu avantaj sayesinde, sosyal medya analizi ve belge sınıflandırması gibi uygulamalarda oldukça faydalı analizler sağlamaktadır (Albalawi vd., 2020; Du ve Liu, 2021).

Modelin olasılıksal yapısı, veri içerisindeki belirsizlikleri ve varyasyonları dikkate alarak sağlam ve esnek analizler gerçekleştirmesine olanak tanır. Özellikle sağlık verilerinin analizinde bu özellik, modelin güçlü yönlerinden biri olarak öne çıkmaktadır. Nitekim, mikrobiyom temelli çalışmalarda LDA, farklı tedavi yanıtları gösteren mikrobiyal alt grupların belirlenmesinde etkili bir biçimde uygulanmıştır. Bu uygulama, LDA'nın yalnızca metin madenciliği alanında değil, aynı zamanda biyolojik ve klinik araştırmalarda da geniş bir kullanım alanına sahip olduğunu ortaya koymaktadır (Breuninger vd., 2021).

Ayrıca, modelin gözetimsiz öğrenme özelliği, verileri etiketleme ihtiyacını ortadan kaldırarak araştırmacıların önceden bilgi sahibi olmadan geniş veri setlerinde analiz yapmasını kolaylaştırır (Maier vd., 2018). Etiketli veri setlerinin maliyetli ve zor temin edildiği durumlarda, bu özellik büyük bir avantaj sağlamaktadır. Ayrıca, model, metinler ve görseller gibi farklı veri türlerine uyarlanabilmekte olup, böylece modelin kullanım alanı genişlemektedir (Albalawi vd., 2020; Rasiwasia ve Vasconcelos, 2013).

LDA'nın bu avantajlarına rağmen, Tablo 6'da özetlenen bazı zorlukları da bulunmaktadır. Bu zorlukların başında, modelin performansını ve elde edilen sonuçların yorumlanabilirliğini doğrudan etkileyen konu sayısı gibi hiperparametrelerin dikkatli şekilde ayarlanması gerekliliği gelmektedir. Hatalı parametre seçimi, aşırı uyum (overfitting) veya yetersiz uyum (underfitting) gibi sorunlara yol açarak, zayıf konu tutarlılığı ve düşük içerik alaka düzeyi oluşturabilir. Özellikle konu sayısı, kelime sayısı ve Dirichlet öncelik parametreleri (α , β), modelin çıktısını ciddi biçimde şekillendirmektedir (Maier et al., 2018; Syed ve Spruit, 2017).

Ayrıca, modelin kelime çantası (bag-of-words) varsayımı kelime sırasını dikkate almadığı için bağlamın önemini göz ardı edebilir. Bu durum, nüanslı konuların anlaşılmasını zorlaştırabilmektedir (Syed ve Spruit, 2018). Büyük veri kümeleriyle çalışırken LDA'nın hesaplama açısından karmaşık olması bir diğer zorluktur. Algoritmanın yinelemeli yapısı nedeniyle, veri seti büyüdükçe hesaplama süresi ve kaynak ihtiyacı da artmaktadır. Bu ölçeklenebilirlik sorununu çözmek için paralel işleme yöntemleri önerilmiştir (Yan vd., 2017). Son olarak, LDA'nın ürettiği konuların insan sezgisine her zaman uymaması, yorumlamayı zorlaştırabilir. Bu durum, model sonuçlarının daha fazla doğrulama gerektirmesine neden olabilir (Maier vd., 2018).

Bu çalışmada, BB alanındaki akademik yayınların analizi sürecinde, modelin sunduğu bazı avantajlar etkin bir şekilde kullanılmıştır. Özellikle, LDA'nın büyük veri kümelerindeki gizli tematik yapıları belirleme yeteneği, BB literatüründeki temel konuların sistematik olarak incelenmesine olanak sağlamıştır. Modelin gözetimsiz öğrenme özelliği sayesinde, önceden belirlenmiş etiketlere ihtiyaç duyulmadan geniş veri setleri analiz edilmiştir; böylece veri toplama sürecinde esneklik sağlanmıştır.

Ancak, analiz sürecinde LDA modelinin bazı zorluklarıyla da karşılaşmıştır. Özellikle, konu sayısının belirlenmesi aşamasında, yanlış parametre seçiminden kaynaklanabilecek aşırı uyum (overfitting) veya yetersiz uyum (underfitting) riskleri göz önünde bulundurulmuştur. Ayrıca, LDA'nın kelime çantası (bag-of-words) yaklaşımı nedeniyle bağlamsal ilişkilerin tam olarak modellenememesi, belirli kavramların anlaşılmasını zorlaştırmıştır. Büyük veri kümeleriyle çalışırken modelin hesaplama maliyetleri ve uzun işlem süresi de önemli bir kısıt oluşturmıştır.

Sonuç olarak, LDA modeli, BB literatüründeki eğilimleri anlamada güçlü bir araç olarak kullanılmış olsa da, modelin parametrik kısıtlamaları ve hesaplama gereksinimleri analiz sürecinde dikkatle ele alınmıştır. Bu doğrultuda, literatürde BB çalışmalarının değerlendirilmesinde kullanılan temel bilimsel analiz yöntemleri kapsamlı biçimde incelenecektir.

3.6. Bulut Bilişim Literatüründe Yöntemsel İnceleme

BB teknolojilerinin giderek daha fazla alanda uygulanmaya başlaması, bu konudaki akademik ilginin de hızlı bir biçimde artmasına neden olmuştur. Özellikle sağlık, finansal sistemler, eğitim uygulamaları ve kamu bilişim altyapılarındaki dijital dönüşüm süreçleri, BB üzerine yapılan çalışmaların hızlanmasına ve içerik bakımından çeşitlenmesine zemin hazırlamıştır. Bu gelişmeler, söz konusu alandaki literatürün hem hacimsel hem de yapısal olarak incelenmesini gerekli kılmıştır.

Akademik çalışmalar genellikle BB'nin temel bileşenleri, güvenlik mekanizmaları, performans değerlendirmeleri ve yeni nesil mimariler üzerine yoğunlaşmaktadır. Özellikle veri güvenliği, büyük veri analitiği, makine öğrenimi ve blok zinciri entegrasyonu gibi konular, son dönemde öne çıkan araştırma alanları arasında yer almaktadır (Alouffi vd., 2021; Tian vd., 2022). BB literatüründe kullanılan analiz yöntemleri şu şekilde özetlenebilir:

Bibliyometrik analiz, bilimsel yayınların sayısal verilerini inceleyerek araştırma eğilimlerini belirleyen bir yöntemdir (Donthu vd., 2021). Özellikle belirli zaman

dilimlerinde arařtırmaların nasıl geliřtiđini ve hangi konulara y6neldiđini tespit etmek iin kullanılmıřtır (Mongeon ve Paul-Hus, 2016).

Atıf analizi, belirli alıřmaların bilimsel etkisini ve akademik literatürdeki konumunu deđerlendirmek iin kullanılan bir y6ntemdir (Zupic ve ater, 2015). BB arařtırmalarında en ok atıf alan makaleler ve yazarların akademik iliřkileri atıf analizleriyle belirlenmiřtir.

Ađ analizi, arařtırmacılar arasındaki iř birliklerini, atıf iliřkilerini ve akademik alanlar arasındaki bađlantıları inceleyerek belirli bir alandaki en etkili arařtırmacıları ve bilgi akıřını ortaya koymak iin kullanılmıřtır (Newman, 2018; Cobo vd., 2011).

Konu modelleme teknikleri ise b6y6k metin k6meleri iindeki temel temaları belirlemek ve akademik yayınlardaki konu eđilimlerini analiz etmek iin uygulanmıřtır (Blei vd., 2003). 6zellikle LDA gibi teknikler, BB arařtırmalarında hen6z yaygın olarak kullanılmamıř olmasına rađmen, bu alanın derinlemesine incelenmesi iin 6nemli bir potansiyel sunmaktadır (Blei vd., 2003; Haq vd., 2019).

Bu alıřmada, BB alanında yapılan akademik arařtırmalarda bařvurulan analiz y6ntemleri Tablo 7'de 6zetlenmiřtir. Tablo, bu y6ntemlerin odaklandığı konu bařlıklarıyla birlikte nasıl uygulandıđını karřılařtırmalı biimde sunmaktadır.

Tablo 7. Bulut biliřim literatüründe kullanılan analiz yaklařımları³

Analiz Y6ntemi	Avantajları	Dezavantajları
Bibliyometrik Analiz	Yayınlara sayısal eđilimlerini ve literatürdeki genel arařtırma yođunluklarını belirler. (Donthu vd., 2021; Mongeon ve Paul-Hus, 2016)	İeriđin detaylı analizini yapamaz, yalnızca nicel verilere dayanır. (Janmajaya vd., 2021)
Atıf Analizi	Bilimsel etkisi y6ksek alıřmalar ve yazarlar belirlenebilir. (Zupic ve ater, 2015; Wang vd., 2016)	Atıf iliřkilerini analiz ederken, alıřmaların ieriksel bađlamı g6z ardı edilir. (Bharati ve Singh, 2019)
Ađ Analizi	Arařtırmacılar arasındaki iř birlikleri ve bilgi akıřı incelenebilir. (Talley vd., 2011; Yaseen ve Albahri, 2023)	Yayın ierikleri hakkında dođrudan bilgi sađlamaz, sadece iliřkileri g6sterir. (Tian vd., 2022)
LDA (Konu Modelleme)	Metin k6melerindeki gizli temalar belirlenebilir, konuların zaman iindeki eđilimleri analiz edilebilir. (Blei vd., 2003; Haq vd., 2019; Sharma vd., 2022; Batool ve Byun, 2024)	Modelin dođru sonu vermesi iin geniř veri setlerine ihtiya duyulur, yanlış modelleme ierik kaybına yol aabilir. (Yu vd., 2018; Tian vd., 2022)

³ Tablo, Donthu vd. (2021), Blei vd. (2003), Sharma vd. (2022) ve Zupic ve ater (2015) alıřmalarından uyarlanarak yazar tarafından hazırlanmıřtır.

3.7. Literatürdeki Önemli Çalışmalar:

BB alanında yapılan akademik araştırmalarda çeşitli analiz teknikleri kullanılmıştır. Sriram ve Khajeh-Hosseini (2010), 2005-2009 yılları arasında yayımlanan 56 makaleyi inceleyerek BB alanındaki teknik temelleri ve araştırma gündemini analiz etmişlerdir. Çalışmada, BB'nin hizmet modelleri ve dağıtım modelleri (ayrıntılı olarak ele alınmıştır. Güvenlik, ölçeklenebilirlik, hizmet sağlayıcılar arası geçiş ve standart eksiklikleri gibi teknik sorunların ön planda olduğu belirlenmiştir.

Bu analiz sonucunda, BB'nin standartlaşma ihtiyacı ve güvenlik açıklarının kapatılması gerektiği vurgulanmıştır. Ayrıca, bulut hizmetlerinin etkin kullanılabilmesi için protokollerin geliştirilmesi ve açık standartların oluşturulması gerektiği belirtilmiştir. Çalışma, BB'nin gelecekteki araştırmalarında güvenlik ve standartlaşma konularının daha fazla öncelik taşıması gerektiğine dikkat çekmektedir.

Marston ve arkadaşları (2010), BB alanındaki mevcut literatürü kapsamlı bir şekilde inceleyerek, teknik temelleri ve iş dünyası perspektifini analiz etmişlerdir. Çalışmada, BB'nin hizmet modelleri ve dağıtım modelleri detaylı olarak ele alınmıştır

Analiz sonucunda, BB'nin benimsenmesinin önündeki temel engellerin güvenlik, ölçeklenebilirlik, sağlayıcılar arası geçiş zorlukları ve standart eksiklikleri olduğu vurgulanmıştır. Veri güvenliği ve evrensel standartların yokluğu kritik sorunlar olarak öne çıkmakt olup BB'nin etkin kullanımı için açık standartların geliştirilmesi ve yeni protokollerin oluşturulması önerilmektedir.

Sivakumaren ve arkadaşları (2012) tarafından yapılan çalışmada, 2001-2010 yılları arasında Web of Science veritabanında yayınlanan BB ile ilgili 510 yayının gelişimi ve dağılımı bibliyometrik analiz yöntemiyle incelenmiştir. Çalışmada, yayınların yıllara göre dağılımı, diller, belge türleri ve ülkeler gibi çeşitli kriterler ele alınmıştır.

Sonuçlara göre, en fazla yayının 2010 yılında yapıldığı ve yayınların büyük çoğunluğunun İngilizce olduğu belirlenmiştir. Ayrıca, BB konusundaki en fazla yayını ABD'nin ürettiği ve "IEEE Internet Computing" dergisinin en fazla yayına ev sahipliği yaptığı tespit edilmiştir. Bu sonuçlar, BB'nin hızla büyüyen bir araştırma alanı olduğunu ve ABD'nin bu alanda lider konumda bulunduğunu göstermiştir.

Heilig ve Voß (2014), 2008-2013 yılları arasında Scopus veritabanında indekslenen 15.376 BB yayını üzerinde yaptıkları çalışmada, BB araştırmalarındaki yayın eğilimleri ve etkilerini bibliyometrik analiz yöntemiyle incelemişlerdir. Çalışmada, yayınların yıllara göre dağılımı, anahtar kelime kümeleri ve atıf desenleri gibi meta-veriler analiz edilmiştir.

Analiz sonucunda, BB arařtırmalarında gvenlik, gizlilik, leklenebilirlik ve veri ynetimi gibi konuların n planda olduėu belirlenmiřtir. Ayrıca, arařtırmaların byk kısmının İngilizce olduėu ve en fazla yayının IEEE gibi dergilerde yayımlandıėı tespit edilmiřtir. Bu sonular, hızla byyen BB sistemlerinde standartlařma ve gvenlik konularına odaklanması gereken bir alan olduėunu gstermektedir.

Gupta ve arkadaşları (2015), 2004-2013 yılları arasında Scopus veritabanında bulunan 21. 397 BB yayını zerinde yaptıkları alıřmada, BB arařtırmalarındaki yayın eėilimlerini ve geliřimini bibliyometrik analiz yntemiyle incelemiřlerdir. alıřmada, yayınların yıllara gre daėılımı, lkeler, iř birliėi oranları ve atıf etkisi gibi eřitli kriterler ele alınmıřtır.

En fazla yayının in ve ABD kaynaklı olduėu tespit edilen bu arařtırmada en ok bilgisayar bilimi alanında alıřma yapıldıėı, en fazla yayının ise konferans bildirileri ve dergiler aracılıėıyla yayımlandıėı bildirilmiřtir. Bu sonular, BB'nin hızla byyen bir arařtırma alanı olduėunu ve uluslararası iř birliėi ve standartlařma konularına daha fazla nem verilmesi gerektiėini gstermektedir.

Jan, Wani ve Hafiz (2015), 2009-2013 yılları arasında Web of Science veritabanında yayınlanan 1. 879 BB yayını zerinde yaptıkları alıřmada, BB arařtırmalarındaki yayın eėilimlerini ve geliřimini scientometrik analiz yntemiyle incelemiřlerdir. alıřmada, yayınların yıllara gre daėılımı, lkelere gre sıralaması, atıf etkisi ve yazar iř birliėi gibi eřitli kriterler ele alınmıřtır. Ayrıca, yayınların byme oranı ve ikiye katlanma sresi gibi metrikler kullanılarak BB alanındaki arařtırma faaliyetlerinin artıř hızı deėerlendirilmiřtir.

Analiz sonucuna gre, BB arařtırmalarının byk lde bilgisayar bilimi alanında yoėunlařtıėı, en fazla yayının ABD ve in kaynaklı olduėu ve oėunun konferans bildirileri aracılıėıyla yayımlandıėı belirlenmiřtir. Anahtar kelime analizleri, alıřmaların aėırlıklı olarak teknolojik konulara odaklandıėını, sosyoekonomik boyutların ise yeterince ele alınmadıėını gstermektedir. Bu bulgular, BB'nin potansiyelinden tam anlamıyla yararlanılabilmesi iin sosyoekonomik konulara da ynelmenin ve mevcut bilgi birikiminin daha iyi anlařılmasının nemini ortaya koymaktadır.

Wang ve diėerleri (2016), 2004-2014 yılları arasında 20 prestijli akademik dergi ve 2 uluslararası konferansta yayımlanan 214 BB yayını, atıf ve eř-atıf analizi yntemleriyle incelemiřlerdir. alıřmada, yayınların atıf iliřkileri, eř-atıf matrisi ve sosyal aė analizi gibi yntemler kullanılarak BB arařtırmalarının temel temaları ve geliřim ařamaları deėerlendirilmiřtir.

Analiz sonucuna göre, BB arařtırmalarının temeller, SaaS modeli, gvenlik ve risk, literatr incelemeleri, benimseme ve etkiler ile modelleme olmak zere altı ana tema etrafında toplandıđı belirlenmiřtir. Ayrıca, BB arařtırmalarının kuluřka, keřif ve geliřme olmak zere ç ařamadan geçtiđi ve bu sreçte teknik konuların yanı sıra sosyoekonomik sorunların da ele alınması gerektiđi vurgulanmıřtır.

Senyo, Addae ve Boateng (2016), 2009-2015 yılları arasında 67 hakemli dergide yayımlanan 285 BB yayını zerinde yaptıkları çalıřmada, BB arařtırmalarındaki eđilimleri ve bilgi yapısını meta-analiz yntemiyle incelemiřlerdir. Çalıřmada, yayınların teorik çerçeveleri, arařtırma yntemleri, cođrafi dađılımı ve analiz dzeyleri gibi çeřitli kriterler ele alınmıřtır.

Analiz sonucuna göre, BB arařtırmalarının byk lçde teknolojik boyut zerinde yođunlařtıđı, buna karřın iř dnyası, kavramsallařtırma ve uygulama alanları gibi konuların yeterince arařtırılmadıđı belirlenmiřtir. Ayrıca, çalıřmaların çođunun deney ve simlasyon gibi teknik yntemler kullandıđı, ancak teorik çerçeveler ve modellerin sınırlı olduđu tespit edilmiřtir. Bu bulgular, BB arařtırmalarının daha dengeli bir yaklařımla ele alınması gerektiđini ve sosyo-teknik konular zerinde daha fazla çalıřma yapılmasının gerekliliđini gstermektedir.

Junfeng Yu ve arkadařları (2018) tarafından yapılan çalıřmada, 2005-2018 yılları arasında Web of Science veritabanında yayınlanan 8.178 BB ile ilgili yayının geliřimi ve dađılımı bibliyometrik analiz yntemiyle incelenmiřtir. Çalıřmada, yayınların yıllara gre dađılımı, diller, belge trleri ve lkeler gibi çeřitli kriterler analiz edilmiřtir.

Analiz sonularına gre, en fazla yayının 2017 yılında (1. 710 yayın) yapıldıđı ve yayınların tamamının İngilizce olduđu belirlenmiřtir. En fazla yayını Çin retmiř (2.487 yayın) ve Chinese Academy of Sciences, bu alanda en fazla yayına sahip kurum olmuřtur. Ayrıca, BB konusundaki iř birliđi ađlarında Çin ve ABD arasında gçl bir uluslararası iř birliđi olduđu gzlemlenmiřtir. Çalıřma, BB'nin hızla byyen bir arařtırma alanı olduđunu ve Çin'in bu alanda lider konumda olduđunu vurgulamaktadır.

Bharati ve Singh (2019), 2010-2019 yılları arasında Web of Science veritabanında indekslenen 11. 776 BB yayını zerinde bibliyometrik analiz gerekleřtirmiřtir. Çalıřmada, yayınların yıllara gre dađılımı, en retken yazarlar, lkeler ve en ok yayın yapan dergiler gibi kriterler ele alınmıřtır. R. Buyya, 138 makale ile en retken yazar olurken; en fazla yayın sırasıyla Çin, ABD ve Hindistan kaynaklıdır. Future Generation Computer Systems, en ok makale yayımlayan dergi olmuřtur. 2018 yılı, 2667 yayınlı en yksek retkenlik yılıdır.

Analiz sonucunda, BB arařtırmalarının gvenlik, byk veri, sanallařtırma ve SaaS gibi temel temalarda yoęunlařtıęı grlmřtır. Yayınların byk çoęunluęu bilgisayar bilimi ve bilgi ynetimi alanında olup, çoęu İngilizcedir. Ayrıca, konferans bildirilerinin BB arařtırmalarında nemli bir yer tuttuęu belirlenmiřtir. alıřma, zellikle gvenlik, veri ynetimi ve enerji verimlilięi konularında daha fazla arařtırmaya ihtiya duyulduęunu vurgulamaktadır.

Garg, Sidhu ve Rani (2019), 2009–2018 yılları arasında Scopus veritabanında yayımlanan 15.591 BB gvenlięi (BBG) yayını zerinde bibliyometrik analiz gerekleřtirmiřtir. alıřmada; yayınların atıf desenleri, anahtar kelime kmeleri, yayın trleri ve arařtırma retkenlięi gibi eřitli ltler incelenmiřtir. Bulgulara gre, BBG alanında en ok alıřılan konular gvenlik, gizlilik, kimlik doęrulama ve eriřim kontroldr. Yayınların oęu bilgisayar bilimi alanında yoęunlařmakta olup, %64.46'sı konferans bildirisi trndedir.

En fazla yayın yapan lkeler sırasıyla in (%29.84) ve Hindistan (%23.86) olurken, en ok atıf alan arařtırmalar Kanada, ABD ve Avustralya kaynaklıdır. En etkili yazar Kui Ren, en aktif kurumlar ise Chinese Academy of Sciences ve Xidian University olmuřtur. alıřma, BBG alanında standartlařma, veri gizlilięi ve geliřmiř gvenlik teknolojileri konularında daha fazla arařtırma yapılması gerektięini vurgulamaktadır.

Sudolska, Lis ve Blař (2019) tarafından yapılan alıřmada, 2008-2018 yılları arasında Scopus veritabanında indekslenen 14.158 BB yayınının geliřimi ve daęılımı bibliyometrik analiz yntemiyle incelenmiřtir. alıřmada, yayınların yıllara gre daęılımı, en retken yazarlar, lkeler ve atıf aęları gibi eřitli kriterler ele alınmıřtır. Arařtırmada, BB ile ilgili ana temalar ve arařtırma sınırları da belirlenmiřtir.

Sonuçlara gre, en fazla yayının 2017 yılında yapıldıęı ve en ok yayını in, Hindistan ve ABD gibi lkelerin rettięi tespit edilmiřtir. Ayrıca, R. Buyya en retken yazar olarak belirlenmiř ve Lecture Notes in Computer Science en fazla yayına sahip dergi olmuřtur. Bu bulgular, BB'nin hızla byyen bir arařtırma alanı olduęunu ve zellikle gvenlik, veri ynetimi ve hizmet kalitesi gibi konularda daha fazla arařtırmaya ihtiya olduęunu gstermektedir.

Bharati ve Gautam (2019), 2010–2019 yılları arasında Web of Science veritabanında yayımlanan 11. 776 BB yayını zerinde yaptıkları alıřmada, BB arařtırmalarındaki yayın eęilimlerini ve geliřimini bibliyometrik analiz yntemiyle incelemiřlerdir. alıřmada; yayınların yıllara gre artıřı, lkelere gre daęılımı, en

retken yazarlar, yayın trleri, diller ve en ok atıf alan alıřmalar gibi eřitli kriterler ele alınmıřtır.

Analiz sonucuna gre, en fazla yayının in (3556) ve ABD (2459) kaynaklı olduđu, Hindistan'ın (997) ise nc sırada yer aldıđı belirlenmiřtir. R. Buyya, 138 yayın ile en retken yazar olurken, yayınların %92.89'unun makale formatında olduđu tespit edilmiřtir. Ayrıca, BB arařtırmalarının %98.74'nn İngilizce olduđu ve BB'nin hızla byyen bir arařtırma alanı olarak ne ıktıđı grlmřtir. En ok yayın yapan dergi ise Future Generation Computer Systems olmuřtur.

Haq, Li ve Hassan (2019), 2010–2018 yılları arasında IEEE, Springer, Wiley, Elsevier ve ACM gibi yksek etki faktrl yayınevlerinde yayımlanan 200 arařtırma makalesini incelemiřlerdir. alıřmada BB ve byk veri konularındaki eđilimleri belirlemek amacıyla metin madenciliđi ve zellikle LDA yntemi kullanılmıřtır. Bu alıřma, BB arařtırmalarında LDA temelli bilgi ıkarımı yapan nadir rneklerden biridir.

Analiz sonucunda, arařtırmaların zellikle uygulama alanları, benimseme faktrleri ve karřılařılan engeller gibi konulara odaklandıđı grlmřtir. IEEE, bu alandaki en fazla yayına sahip kaynak olarak ne ıkmıřtır. Ayrıca, LDA ile yapılan konu modelleme srecinde, BB'nin byk veriyle entegrasyonu, sađlık, eđitim, IoT ve gvenlik gibi eřitli alanlardaki kullanımları ortaya konmuřtur. alıřma, BB arařtırmalarında bilgi yapısını ve eđilimleri analiz etmede LDA kullanımının etkili bir yntem olduđunu gstermektedir.

Janmajaya ve arkadařları (2021), 2010–2020 yılları arasında Web of Science veritabanında indekslenen Endstri 4.0 konulu yayınları bibliyometrik analiz yntemiyle incelemiřlerdir. alıřmada, yayınların yıllara gre dađılımı, en etkili yazarlar, lkeler, kurumlar, atıf yapısı ve iř birliđi ađları gibi kriterler ele alınmıřtır. Ayrıca, k-means kmeleme ve LDA yntemleri kullanılarak, Endstri 4.0 arařtırmalarında ne ıkan temalar belirlenmiřtir.

Analiz sonularına gre, BB Endstri 4.0'ın temel bileřenlerinden biri olarak tanımlanmıř ve BB'nin zellikle veri ynetimi, gvenlik, altyapı hizmetleri ve uygulama alanları gibi konularda yođunlařtıđı tespit edilmiřtir. alıřma, BB'nin Endstri 4.0 ekosisteminde kritik bir rol oynadıđını gstermekte ve bu alandaki arařtırma eđilimlerinin artarak devam ettiđini ortaya koymaktadır.

Li ve arkadařları (2021), BB sistemlerinde blockchain tabanlı gven ynetimi zerine kapsamlı bir inceleme gerekleřtirmiřtir. alıřmada, 2017–2020 yılları arasında yayımlanmıř 35 akademik makale, tematik ierik analizi ve uygulama alanlarına gre

taksonomik sınıflandırma yöntemiyle değerlendirilmiştir. Yazarlar, mevcut çözümleri sistematik şekilde incelemiş ve bu çözümleri güven mimarilerine göre kategorize etmiştir.

Analiz sonuçlarına göre, BB'deki merkezi güven modelleri ağ tıkanıklığı ve tek nokta arızası gibi sorunlara neden olmakta; blockchain ise dağıtık, şeffaf ve izlenebilir yapısıyla bu sorunlara etkili çözümler sunmaktadır. Blockchain tabanlı modellerin özellikle kimlik doğrulama, erişim kontrolü ve veri güvenliği gibi alanlarda öne çıktığı belirlenmiştir. Çalışma, bu alanda hâlâ gelişime açık birçok yön olduğunu ve daha fazla araştırmaya ihtiyaç duyulduğunu vurgulamaktadır.

Büyük Veri ve Bulut Bilişim Teknoloji Yol Haritası (2021), Türkiye'nin dijital dönüşüm hedefleri doğrultusunda hazırlanmış; bibliyometrik analiz, Delphi yöntemi ve Teknoloji Hazırlık Seviyeleri (THS) kullanılarak BB ve büyük veri alanlarındaki araştırmalar değerlendirilmiştir. Çalışmada, veri güvenliği, SaaS, büyük veri işleme ve bulut altyapısı gibi temel temaların öne çıktığı belirtilmiştir. Türkiye'de bu alanlarda yerli teknolojilerin geliştirilmesi ve kamu-özel sektör iş birliğinin artırılması gerektiği vurgulanmıştır.

Analize göre BB araştırmaları özellikle veri güvenliği, standartlaşma ve entegrasyon konularına odaklanmaktadır. Dağıtık kayıt defteri (DLT) ve blockchain tabanlı güvenlik çözümleri, stratejik uygulama alanları arasında öne çıkmaktadır. Yol haritası, Ar-Ge yatırımları, insan kaynağı geliştirme ve üniversite-sanayi iş birliği gibi unsurların güçlendirilmesi gerektiğini ortaya koymaktadır.

Ayaz, Çelik ve Özyurt (2021), Web of Science veritabanında yer alan 48.692 BB yayını üzerinde bibliyometrik analiz uygulayarak BB alanındaki araştırma eğilimlerini ve gelişimini incelemişlerdir. Çalışmada, yayınların yıllara göre dağılımı, ülkeler, en etkili yazarlar, atıf ağları, anahtar kelime kümeleri, kurumlar ve dergiler gibi çeşitli kriterler analiz edilmiştir. Araştırma kapsamında VOSviewer ve BibExcel yazılımları kullanılarak kapsamlı bir bibliyometrik haritalama yapılmıştır.

Analizlere göre, BB yayınlarında 2016 yılına kadar hızlı bir artış yaşanmış, bu tarihten sonra ise araştırmalar daha spesifik alt temalara yönelmiştir. Çin, ABD ve Hindistan en fazla katkı sağlayan ülkeler olarak öne çıkarken; fog computing, mobile cloud computing, data security ve trust gibi konular sıkça araştırılmış, ayrıca edge computing ve blockchain gibi yeni temaların da yükselişte olduğu belirlenmiştir. Bu bulgular, BB alanında güvenlik, performans ve veri yönetimi gibi konuların öncelikli araştırma alanları haline geldiğini göstermektedir.

Saadat M. Alhashmi ve arkadaşları (2023) tarafından yapılan çalışmada, 1991–2022 yılları arasında Scopus veritabanında yayımlanan 2.335 yapay zekâ (YZ) temalı yayın, bibliyometrik analiz ve LDA yöntemiyle incelenmiştir. Çalışma, yayınların yıllara göre dağılımı, dil, belge türü, ülke katkısı, en çok atıf alan dergiler ve anahtar kelime eğilimleri gibi çeşitli kriterlere odaklanmıştır. 2021 yılı, 616 yayınla en yüksek yayım yılı olurken, yayınların büyük çoğunluğunun İngilizce olduğu saptanmıştır.

Çalışmada, bulut ve fog bilişim mimarilerinin sağlık verilerinin işlenmesi ve güvenliği açısından oynadığı kritik rol özellikle vurgulanmıştır. 2021 yılında yayımlanan sağlık temalı 938 makalenin yaklaşık %20'si bu teknolojilere odaklanmış; yalnızca bu konularda 186 yayın yapılmıştır. Cloud computing, IoT ve veri gizliliği gibi kavramlar, çalışmada en çok işlenen temalar arasında yer almıştır. Bu bulgular, yapay zekâ ile bütünleşik bulut mimarilerinin sağlık hizmetlerinde yükselen bir araştırma alanı olduğunu ortaya koymaktadır.

Teh ve Subramaniam (2023), 2008–2022 yılları arasında Scopus veritabanında indekslenen 20.514 BB yayını üzerinde bibliyometrik analiz yöntemiyle gerçekleştirdikleri çalışmalarında, BB alanındaki literatürün gelişimini ve eğilimlerini incelemiştir. Yayınların yıllara göre dağılımı, anahtar kelime kümeleri, belge türleri, ülkeler, yazarlar ve iş birlikleri gibi çeşitli kriterler detaylı olarak değerlendirilmiştir.

Analiz sonuçlarına göre, 2008'den itibaren BB yayınlarında sürekli bir artış görülmüş; ancak 2020 yılında COVID-19 pandemisi nedeniyle belirgin bir düşüş yaşanmıştır. Ayrıca, blockchain, IoT, yapay zekâ gibi yeni teknolojilerin BB ile birlikte ele alındığı çalışmaların sayısında artış gözlemlenmiştir. Çalışma, BB araştırmalarının kapsamlı biçimde sürdürülmesi ve gelişen teknolojilerle entegrasyonunun teşvik edilmesi gerektiğini vurgulamaktadır.

Alhashmi ve arkadaşları (2024), 1991–2022 yılları arasında Scopus veritabanında indekslenen 2.335 makale üzerinden yürüttükleri çalışmada, sağlık alanındaki AI uygulamalarını bibliyometrik analiz ve LDA yöntemi ile incelemiştir. Çalışma, makine öğrenimi, derin öğrenme ve BB gibi alanlarda yürütülen yayınları değerlendirerek bu alanlardaki temel eğilimleri ortaya koymuştur.

Analizlerde, cloud/fog computing, blockchain, IoT gibi teknolojilerin AI destekli sağlık hizmetleri araştırmalarında hızla önem kazandığı görülmüştür. 2021 yılında sadece bu konulara odaklanan 186 yayının yapılması, BB teknolojilerinin sağlık verisi yönetimi ve güvenliğinde kritik rol oynadığını göstermektedir. LDA sonuçları, hasta verisi güvenliği, karar destek sistemleri ve görüntü analizi gibi temaların öne çıktığını ortaya koymuştur.

Toklu ve arkadaşları (2024), 15 Mayıs 2024 tarihine kadar YÖK Ulusal Tez Merkezi (YÖKTEZ) arşivinde yer alan 360 BB konulu lisansüstü tezi bibliyometrik analiz yöntemiyle incelemişlerdir. Çalışmada tezlerin yıllara göre dağılımı, üniversite türleri, ana bilim dalları, enstitüler, anahtar kelimeler ve kullanılan yöntemler gibi çeşitli ölçütler dikkate alınmıştır.

Analiz sonuçlarına göre, BB konusunda en fazla tezin 2022 yılında yayımlandığı, en çok çalışmanın ise İstanbul Teknik Üniversitesi'nde yapıldığı belirlenmiştir. Ayrıca, bilgisayar mühendisliği hem konu hem de ana bilim dalı düzeyinde ilk sırada yer alırken, tezlerin çoğunun fen bilimleri enstitüsü kapsamında hazırlandığı tespit edilmiştir. “Bulut bilişim”, 186 kez ile en sık kullanılan anahtar kelime olmuştur. Çalışma, Türkiye’de BB alanındaki akademik ilginin giderek arttığını ve bu alanda daha fazla çalışmaya ihtiyaç olduğunu vurgulamaktadır.

Tablo 8. Literatür özeti (birden fazla kaynaktan derlenmiştir).

YAZAR ve YIL	KONU
Ilango Sriram ve Ali Khajeh-Hosseini (2010)	Research Agenda in Cloud Technologies
Sean Marston, Zhi Li, Subhajyoti Bandyopadhyay, Juheng Zhang ve Anand Ghalsasi (2011)	Cloud computing — The business perspective
K.S. Sivakumaren, S. Swaminathan ve G. Karthikeyan (2019)	Growth and Development of Publications on Cloud Computing: A Scientometric Study
Leonard Heilig ve Stefan Voß (2014)	A Scientometric Analysis of Cloud Computing Literature
Gupta B.M., Dhawan S.M. ve Gupta Ritu (2015)	Cloud computing research: A bibliometric analysis
Rosy Jan, Wasim Rashid Wani ve Obaid Hafiz (2015)	Scientometric Analysis of Cloud Computing
Nianxin Wang, Huigang Liang, Yu Jia, Shilun Ge, Yajiong Xue ve Zhining Wang (2016)	Cloud computing research in the IS discipline: A citation/co-citation analysis
Senyo Prince Kwame, Addae Erasmus ve Boateng Richard (2018)	Cloud computing research: A review of research themes, frameworks, methods and future research directions
Junfeng Yu, Zijiang Yang, Shiwei Zhu, Beibei Xu, Sisi Li, Mingjun Zhang (2018)	A Bibliometric Analysis of Cloud Computing Technology Research
Vijay Kumar Bharati ve Mahender Pratap Singh (2019)	Rejuvenating Libraries from the Cloud: A Bibliometric Analysis of Cloud Computing
Deepak Garg, Jagpreet Sidhu, Shalli Rani (2019)	Emerging trends in cloud computing security: a bibliometric analysis
Agata Sudolska, Andrzej Lis, Róża Błaś (2019)	Cloud Computing Research Profiling: Mapping Scholarly Community and Identifying Thematic Boundaries of the Field

Tablo 8. (Devamı)

Muhammad Inaam Ul Haq, Qianmu Li, Shoaib Hassan (2019)	Gerçek başlık: <i>A systematic review and meta-analysis of cloud computing and big data in Industry 4.0</i>
Vijay Kumar Bharati, Sunil Kumar Gautam (2019)	Mapping of Scholarly Publication on Cloud Computing during 2010 - 2019: A Bibliometric Study Based on Web of Science
Janmajaya, Manvendra; Shukla, Amit K.; Muhuri, Pranab K.; Abraham, Ajith (2021)	Industry 4.0: Latent Dirichlet Allocation and clustering based theme identification of bibliography
Wenjuan Li, Jiyi Wu, Jian Cao, Nan Chen, Qifei Zhang, Rajkumar Buyya (2021)	Blockchain-based trust management in cloud computing systems: a taxonomy, review and future directions
T.C. Cumhurbaşkanlığı Bilim, Teknoloji ve Yenilik Politikaları Kurulu (2021)	Büyük Veri ve Bulut Bilişim Teknoloji Yol Haritası Sonuç Raporu
Ahmet Ayaz, Kamil Çelik, Özcan Özyurt (2021)	Pattern detection in cloud computing: Bibliometric mapping of publications in the field from past to present
Saadat M. Alhashmi, Ibrahim Abaker Targio Hashem, Islam Al-Qudah (2023)	Artificial Intelligence applications in healthcare: A bibliometric and topic model-based analysis
R. Teh, A. Subramaniam (2023)	A bibliometric analysis on cloud computing literature development
M. Alhashmi, Ibrahim Abaker Targio Hashem, Islam Al-Qudah (2024)	Artificial Intelligence applications in healthcare: A bibliometric and topic model-based analysis
Rümeysa Toklu, Merve Gök, Emel Güven, Tamer Eren (2024)	Türkiye’de Bulut Bilişim Üzerine Yapılmış Tezlerin Bibliyometrik Analizi

Mevcut literatür taramaları, BB alanındaki akademik çalışmalarda özellikle veri madenciliği, bilgi çıkarımı ve konu modelleme gibi ileri düzey analiz yöntemlerinin nadiren kullanıldığını ortaya koymaktadır. Bu nedenle literatürdeki araştırmalar genellikle yüzeysel düzeyde kalmakta olup büyük veri kümelerindeki gizli temaların ve alt alanların belirlenmesinde yetersizdir. BB literatüründe karşılaşılan bu sınırlılıklar, anlamlı bilgilerin ortaya çıkarılması açısından metodolojik bir eksikliği işaret etmektedir.

Söz konusu eksiklikleri gidermek amacıyla bu çalışmada, Scopus veri tabanında yer alan toplam 18.689 akademik yayın analiz edilerek BB alanındaki araştırma trendlerini, teknolojileri, araçları, kavramları ve uygulamaları belirlemek hedeflenmiştir. Literatürdeki bu metodolojik boşluğu doldurmak için, büyük metin tabanlı veri kümelerindeki gizli temaların belirlenmesinde etkili olan LDA yöntemi tercih edilmiştir. LDA’nın seçilmesinin temel nedeni, mevcut araştırmaların ötesine geçerek,

yayınlar arasındaki ilişkileri, tematik kümeleri ve gizli kalmış eğilimleri objektif ve kapsamlı bir biçimde ortaya koyabilme yeteneğidir.

Sonuç olarak, bu çalışma BB alanında mevcut literatür boşluklarını doldurmaya yönelik anlamlı bir adım teşkil etmektedir. İzleyen bölümde, BB yayınlarına LDA yöntemi uygulanarak elde edilen tematik dağılımlar analiz edilecektir.



4. LDA MODELLEME SÜRECİ VE TEMATİK YAPININ BELİRLENMESİ

4.1. LDA Tabanlı Yöntemsel Yaklaşım ve Analiz Süreci

Bu araştırmada benimsenen yöntem, BB literatüründe sıklıkla kullanılan geleneksel bibliyometrik yaklaşımlardan ayrılmakta ve daha derinlemesine analiz imkânı sunmaktadır. Örneğin, Alfalah ve Veri (2024) gibi çalışmalarda yalnızca sistematik literatür taraması ve temel istatistiksel yöntemler kullanılmış; metinlerin bağlamsal içeriği ve tematik örüntüleri ise büyük ölçüde göz ardı edilmiştir. Bu durum, BB alanında mevcut eğilimleri yüzeysel bir şekilde yansıtmakta ve kavramsal yapıların derinliğini sınırlamaktadır.

Bu doğrultuda, bu çalışmada Scopus veritabanından elde edilen geniş çaplı akademik yayın kümesi üzerine odaklanılmış ve yalnızca yayınların sayısal dağılımları değil, içerik düzeyindeki tematik yapıları da analiz edilmiştir. Yayın metinleri, metin madenciliği, bigram-trigram modelleri ve LDA gibi ileri düzey doğal dil işleme yöntemleriyle işlenmiş; böylece BB literatürünün altında yatan anahtar temalar ve kavramsal bağlamlar nicel verilere dayanarak ortaya konmuştur.

Bu yöntemsel yaklaşım, yalnızca araştırma konularının sıklığını belirlemekle kalmayıp, temalar arası ilişkileri, bağlamsal yoğunlukları ve yapısal örüntüleri de görselleştirilebilir biçimde analiz etmeye imkân tanımaktadır. Böylece BB alanındaki araştırma boşlukları ve gelişim eğilimleri daha objektif ve kapsamlı biçimde değerlendirilebilmiştir.

Araştırma yöntemi dört aşamada yürütülmüştür:

1. Veri Toplama
2. Ön İşleme
3. LDA Modelleme Süreci
4. Veri Görselleştirme ve Sonuçların Değerlendirilmesi

Bu metodoloji, önceki çalışmaların metodolojik sınırlamalarını aşarak, daha derinlemesine ve kapsamlı bir analiz sunmayı hedeflemektedir. LDA yöntemi ile elde edilen tematik haritalama, BB araştırmalarındaki mevcut durumun ve gelecekteki eğilimlerin anlaşılmasına katkı sağlamaktadır.

4.2. Veri Toplama

Bu çalışmada, 2014–2024 yılları arasında yayımlanmış akademik yayınları içeren veri kümesi, Scopus veri tabanı kullanılarak oluşturulmuştur BB alanına ilişkin literatür, temel anahtar kelime olan “*cloud computing*” üzerinden, belirli sınırlayıcı filtreler uygulanarak taranmıştır. Seçilen yayınlar, niteliklerine göre sınıflandırılmış ve analiz süreci için yapılandırılmıştır.

4.2.1. Veri Kaynağı ve Formatı

Araştırmada kullanılan yayınlar, güvenilirliği ve disiplinler arası kapsamı ile bilinen Scopus veri tabanından elde edilmiştir. Scopus, bibliyometrik analizlerde yaygın biçimde kullanılan, yüksek kaliteli meta-veri sağlayan bir kaynak olarak akademik çalışmalarda sıklıkla tercih edilmektedir (Baas vd., 2020).

Veri toplama sürecinde, Tablo 9’ da özetlenen ve yayın yılı (2014–2024), belge türü (makale ve konferans bildirisi), disiplin alanı (Bilgisayar Bilimleri ve Mühendislik), dil (İngilizce), erişim türü (Açık Erişim) gibi ölçütler uygulanmıştır. Elde edilen veriler, analiz sürecinin düzenli yürütülebilmesi amacıyla CSV formatında dışa aktarılmış ve metin madenciliği, konu modelleme gibi analizlerde kullanılmak üzere yapılandırılmıştır.

Tablo 9. Veri toplama sürecinde uygulanan kriterler.

Kriter	Açıklama
Anahtar Kelime	"cloud computing"
Arama Alanları	Başlık, Özet, Anahtar Kelimeler
Tarih Aralığı	2014 - 2024
Yayın Türleri	Makale (Article), Konferans Bildirisi (Conference Paper)
Disiplin Alanları	Bilgisayar Bilimleri (COMP), Mühendislik (ENGI)
Dil	İngilizce
Erişim Türü	Tam Altın Açık Erişim veya Açık Erişim

4.2.2. Veri Setinin Oluşturulması

Belirlenen filtreleme kriterleri doğrultusunda, Scopus veri tabanından toplam 18.689 yayından oluşan kapsamlı bir veri seti elde edilmiştir. Bu yayınlar, yalnızca makaleler (articles) ve konferans bildirileri (conference papers) ile sınırlı tutulmuş; dil, erişim türü ve konu sınırlamaları da dikkate alınarak seçilmiştir.

Veri seçimi süreci, yalnızca araştırma sorularını doğrudan yanıtlayabilecek ve yüksek akademik standartlara sahip çalışmaların dahil edilmesini sağlamıştır. Böylece oluşturulan veri seti, BB literatüründeki tematik yapıları analiz edebilmek için sağlam

ve güvenilir bir temel sunmuştur. Bu aşamanın ardından, elde edilen verilerin analiz sürecine uygun hale getirilmesi amacıyla ön işleme sürecine geçilmiştir.

4.3. Ön İşleme

Veri toplama sürecini takiben, modelleme işlemlerinin doğru ve güvenilir biçimde yürütülebilmesi için veriler bir dizi ön işleme adımına tabi tutulmuştur. Bu aşama, ham verilerin temizlenmesi, yapılandırılması ve analiz modellerine uygun biçime getirilmesini kapsamaktadır.

4.3.1. Kütüphaneleri Kullanımı

Çalışmada kullanılan Python kütüphaneleri, veri temizleme, modelleme ve görselleştirme gibi süreçlerde önemli bir rol oynamıştır. Tablo 10'da analiz sürecinde kullanılan kütüphaneler ve işlevleri özetlenmiştir.

Tablo 10. Analiz sürecinde kullanılan python kütüphaneleri ve işlevleri.⁴

Kütüphane	Kullanım Amacı
pandas	CSV dosyalarının okunması, veri filtreleme ve düzenleme işlemleri.
numpy	Matematiksel işlemler (örneğin ortalamalar, standart sapma vb.).
gensim	LDA (Latent Dirichlet Allocation) modelleme ve metin işleme (örneğin sent_to_words fonksiyonu).
spaCy	Metinlerden kelimelerin lemmatize edilmesi (köklerine indirgenmesi).
nltk	İngilizce durak kelimelerin (stopwords) temizlenmesi ve çıkarılması.
matplotlib	Basit veri görselleştirmeler (örneğin, çubuk grafikleri).
seaborn	Daha karmaşık ve estetik veri görselleştirme (örneğin, korelasyon ısı haritaları).
plotly	Dinamik ve etkileşimli görselleştirme grafikleri.
pyLDAvis	LDA modelindeki konu dağılımlarının görselleştirilmesi (interaktif).
cartopy	Coğrafi veri görselleştirme (örneğin, haritalar üzerinde veri işleme).
networkx	Ağ analizleri ve veri bağlantılarının görselleştirilmesi.

Araştırma sürecinde, veri işleme, analiz ve görselleştirme aşamalarında kullanılan Python kütüphaneleri, çalışmanın metodolojik altyapısını güçlü bir şekilde desteklemiştir. Analiz sırasında, yalnızca anlamlı çıktılara odaklanılabilmesi amacıyla `warnings` modülü kullanılarak gereksiz sistem uyarıları devre dışı bırakılmıştır. Bu uygulama, çıktıların daha sade ve anlaşılır hale gelmesini sağlamış; böylece analiz sürecinin hem tutarlılığı hem de sonuçların güvenilirliği artırılmıştır.

⁴ Bu tablo, çalışmada kullanılan açık kaynak Python kütüphaneleri ve bunların analiz sürecindeki işlevlerine göre yazar tarafından derlenmiştir.

4.3.2. Veri Temizleme

BB alanındaki akademik yayınların analiz edilebilmesi için, verilerin anlamlı ve düzenli bir biçimde hazırlanması gerekmektedir. Bu kapsamda, metinlerin gereksiz unsurlardan arındırılarak analiz ve modelleme sürecine uygun hale getirilmesi, kritik bir adımdır. Temizleme işlemleri, hem verideki gürültünün azaltılmasını hem de modelleme doğruluğunun artırılmasını amaçlamaktadır.

Önceki aşamada belirlenen kriterler doğrultusunda oluşturulan veri kümesi, *pandas* kütüphanesi kullanılarak yüklenmiş ve 2014–2024 yılları arasındaki yayınlarla sınırlandırılmıştır. *Pandas*, veri filtreleme sürecinde hız ve güvenilirlik sağlayarak yalnızca amaca uygun akademik çalışmaları analiz kapsamına almıştır. Tablo 9, bu süreçte uygulanan filtreleme kriterlerini detaylı olarak sunmaktadır.

Veri setinin yüklenmesinin ardından metin işleme aşamasına geçilmiştir. Bu adımda ham metinler düzenlenerek analiz için anlamlı bir yapıya dönüştürülmüştür. Literatürde de belirtildiği üzere, metin ön işleme işlemleri analiz doğruluğunu artırmakta ve bağlamsal anlamın korunmasına katkı sağlamaktadır (Manning vd., 2008).

Bu süreçte gerçekleştirilen başlıca veri temizleme adımları şunlardır:

- **E-posta adreslerinin çıkarılması.** Anlamsal analizde katkı sağlamayan adresler temizlenmiştir.
- **Fazladan boşlukların silinmesi.** Metin yapısının düzenlenmesi sağlanmıştır.
- **Tek tırnak karakterlerinin temizlenmesi.** Modellemeyi zorlaştıran veya anlam taşımayan işaretler kaldırılmıştır.

Bu adımların ardından, veri kümesi hem içeriksel hem biçimsel olarak daha tutarlı bir yapıya kavuşmuştur. Aşağıda, temizleme işlemi sonrası elde edilen örnek bir metin sunulmaktadır:

Örnek Çıktı: *"Edge cloud computing has developed rapidly. Technologies bring convenience to online education. We propose a deep learning model for music recommendation. Experimental results demonstrate its effectiveness."*

Bu örnek, ön işleme sürecinin veriyi ne kadar odaklı ve analiz edilebilir hale getirdiğini göstermektedir. Ayrıca bu süreçte, analiz sırasında yalnızca anlamlı çıktıları odaklanılabilmesi amacıyla *warnings* modülü kullanılarak gereksiz sistem uyarıları devre dışı bırakılmıştır. Bu uygulama, çıktıların daha sade ve anlaşılır hale gelmesini sağlamış; böylece analiz sürecinin hem tutarlılığı hem de sonuçların güvenilirliği artırılmıştır.

Temizleme işlemleri tamamlandıktan sonra, veri kümesi modelleme için uygunluk açısından kontrol edilmiştir. Bu kontroller; anlamlı içerik, biçimsel tutarlılık ve analiz odaklılık ölçütlerine göre gerçekleştirilmiştir:

- **Anlamlı içerik:** Yalnızca bilgi değeri taşıyan ifadeler korunmuştur.
- **Tutarlılık:** Tüm metinler ortak bir yapıya dönüştürülmüştür.
- **Odaklanmış veri seti:** Gereksiz unsurların çıkarılması, modelleme sürecini kolaylaştırmıştır.

4.3.3. Stopwords Listesi

Stopwords kaldırma işlemi, metinlerde anlamsız kelimelerin temizlenmesi yoluyla analiz sürecinin daha odaklanmış ve anlamlı hale gelmesine olanak tanımaktadır. Literatürde stopwords temizlemenin, metin analizi süreçlerinde veri gürültüsünü azalttığı ve analizin doğruluğunu artırdığı belirtilmektedir (Manning vd., 2008).

Bu çalışmada, genel İngilizce stopwords listesinin yanı sıra BB bağlamına özgü teknik terimler de belirlenmiş ve listeye eklenmiştir. Böylece, yalnızca anlam taşıyan kelimelerin modele dahil edilmesi sağlanmış, analizin bağlamsal anlamı korunarak sonuçların doğruluğu artırılmıştır.

Tablo 11. Stopwords kategorileri ve örnekler.

Kategori	Örnek Stopwords
Genel İngilizce	"the", "is", "and", "to", "more"
Çalışmaya Özgü	"paper", "study", "model", "method", "result"
Genişletilmiş	"some", "these", "over", "me", "just"

Tablo 11’ de görüldüğü üzere, "paper", "model", "method" gibi çalışmaya özgü kavramlar ile "the", "and", "is" gibi sık kullanılan genel İngilizce kelimeler listeye dahil edilmiştir. Bu kapsamlı stopwords listesi, veri setindeki anlamsız tekrarları ortadan kaldırmış, metinlerin bağlamsal bütünlüğünü artırarak analiz sürecine anlamsal derinlik kazandırmıştır (Rajaraman ve Ullman, 2011).

4.3.4. Tokenizasyon ve N-gram Modelleri

Tokenizasyon, metinlerin daha küçük ve anlamlı birimlere (*token*) ayrılması işlemidir ve doğal dil işleme alanında temel bir adımdır. Bu işlem, metinlerin dilsel yapısını daha işlenebilir bir forma dönüştürerek modelleme süreçlerinde doğruluğun artırılmasına katkı sağlamaktadır. Özellikle *n*-gram modelleme gibi daha gelişmiş

analizlerde, *token*'ların oluşturduğu yerel bağlam, bağlamsal ilişkilerin belirlenmesinde kritik rol oynamaktadır (Joulin vd., 2016).

Joulin ve arkadaşları (2016), geliştirdikleri *fastText* modelinde, geleneksel *bag-of-words* yaklaşımına ek olarak *n-gram* özelliklerini kullanarak metinlerdeki sözcük sıralarının taşıdığı bağlamı daha iyi yakalamayı başarmışlardır. Bu sayede, sınıflandırma performanslarında anlamlı iyileşmeler gözlemlenmiştir. Özellikle *bigram* ve *trigram* kullanımı, sınıflandırma doğruluğunu artırmaktadır.

Ayrıca, *stopwords* temizleme işlemiyle birlikte kullanıldığında, yalnızca anlam taşıyan kelimelerin modele dahil edilmesi sağlanmakta ve modelleme süreci daha hedef odaklı hâle gelmektedir. Bu da modelin anlamsal doğruluğunu artırmakta ve bilgi yoğunluğunu yükseltmektedir (Joulin vd., 2016).

Bu çalışma kapsamında gerçekleştirilen tokenizasyon süreci aşağıdaki adımlardan oluşmuştur:

- **Küçük Harfe Dönüştürme:** Metinlerdeki tüm harfler küçük harfe çevrilerek dilsel tutarlılık sağlanmıştır.
- **Noktalama İşaretlerinin Temizlenmesi:** Noktalama işaretleri çıkarılmış ve yalnızca anlam taşıyan kelimeler bırakılmıştır.
- **Kelime Bazlı Ayırma:** Metinler, kelime düzeyinde bölünerek her kelime ayrı bir token olarak ele alınmıştır.

Bu adımların tamamlanmasının ardından, metinler modellemeye uygun ve tutarlı bir yapıya kavuşturulmuştur. Uygulanan bu işlemler, metinlerin hem daha anlamlı bir biçimde analiz edilmesine hem de bağlamsal ilişkilerin daha doğru bir biçimde modellenmesine olanak tanımıştır.

Örnek:

Orijinal metin:

“With the advent of the big data era, edge cloud computing has developed rapidly.”

Tokenize edilmiş çıktı:

['with', 'the', 'advent', 'of', 'the', 'big', 'data', 'era', 'edge', 'cloud', 'computing', 'has', 'developed', 'rapidly']

Bu örnek, tokenizasyon işleminin metinlerin dilsel yapılarını düzenleyerek analiz süreçlerini nasıl kolaylaştırdığını ve bağlamsal ilişkileri nasıl güçlendirdiğini açıkça göstermektedir. Tokenizasyon işlemiyle elde edilen bu yapıların, kelimeler arasındaki daha karmaşık bağlamsal ilişkileri yansıtabilmesi için *bigram* ve *trigram* gibi çoklu sözcük dizileri oluşturulmaktadır.

4.3.5. Bigram ve Trigram Modellerinin Oluşturulması

Bigram ve trigram modelleri, kelimeler arasındaki bağlamsal ilişkileri belirlemede önemli bir rol oynamaktadır. Literatürde bu modeller, sıkça birlikte kullanılan teknik ifadelerin tespitinde etkili araçlar olarak değerlendirilmektedir (Joulin vd., 2016; Beneki vd., 2020).

Özellikle *cloud computing* ve *data storage* gibi çok sözcüklü ifadeler, BB literatüründeki ana temaların anlaşılmasına katkı sunmuştur. Bu tür ifadelerin analize dahil edilmesi, yalnızca tekil kelimelerle değil, anlamlı kelime gruplarıyla çalışılmasını mümkün kılarak, metinlerin bağlamsal içeriğini daha net biçimde ortaya koymaktadır (Joulin vd., 2016).

Bu modeller, metin madenciliği süreçlerinde yalnızca sözcük frekanslarını değil, kelimeler arasındaki ilişkileri de dikkate alarak, daha yapısal ve anlamlı analizler yapılmasına olanak sağlamaktadır. Böylece, literatürdeki eğilimler daha net belirlenmiş, teknik kavram kümeleri daha doğru şekilde ortaya çıkarılmıştır (Beneki vd., 2020).

Bigram ve trigram modellerinin bu çalışmaya sağladığı temel katkılar üç başlık altında toplanabilir:

- **Tematik Odaklanma:** Sık kullanılan çoklu kelime ifadelerinin belirlenmesiyle, analiz edilen temaların daha açık biçimde sınıflandırılması sağlanmıştır (Beneki vd., 2020).
- **Bağlamsal Derinlik:** Kelimeler arasındaki anlam ilişkileri korunarak metinlerin anlamsal bütünlüğü desteklenmiştir (Joulin vd., 2016).
- **LDA Modelleme İçin Hazırlık:** Elde edilen kelime grupları, LDA modellemesi öncesi güçlü bir bağlamsal temel sunmuştur (Joulin vd., 2016).

Bu çalışmada oluşturulan bazı bigram örnekleri şunlardır: *cloud_computing*, *data_storage*, *digital_music*, *matrix_factorization* ve *neural_network*; trigram örnekleri ise *edge_cloud_computing*, *digital_music_recommendation*, *deep_learning_model*, *online_music_education* ve *weighted_matrix_factorization* şeklindedir. Bu ifadeler, literatürdeki teknik kavramların doğru biçimde sınıflandırılmasına ve tematik analizlere sağlıklı bir şekilde dahil edilmesine olanak sağlamıştır.

Bu doğrultuda, bigram ve trigram modelleriyle elde edilen bağlamsal yapı, bir sonraki aşamada uygulanacak olan LDA modelleme sürecine sağlam bir temel oluşturmuştur.

Bu temele dayanarak oluşturulan metinsel korpus, LDA için gerekli sayısal temsile dönüştürülmeden önce son hâlini almıştır.

4.3.6. Stopwords ve Boş Cümlelerin Kaldırılması

Veri ön işleme sürecinin önemli adımlarından biri olan *stopwords*'lerin kaldırılması, metinlerde anlam taşımayan veya düşük bilgi değeri taşıyan kelimelerin ayıklanarak analiz sürecinin daha odaklanmış hale gelmesini sağlamıştır. Bu işlem, yalnızca bilgi içeriği taşıyan kelimelerin modele dahil edilmesini sağlayarak, modelleme sürecine daha sağlam ve hassas bir zemin oluşturmuştur.

Bu çalışmada kullanılan *stopwords* listesi, 3.3.3. bölümünde tanımlanmış ve üç katmanlı bir yapı ile oluşturulmuştur. Liste; genel İngilizce stopwords ("the", "is", "and" gibi), çalışmaya özgü teknik terimler ("paper", "model", "method" gibi) ve genişletilmiş bağlamsal ifadeler ("some", "these", "just" gibi) içermektedir. Bu kapsamlı yaklaşım sayesinde, yalnızca bağlamsal anlam taşıyan kelimeler analiz sürecine dahil edilmiş, metinlerdeki anlamsız tekrarlar azaltılmış ve bağlamsal bütünlük korunmuştur. Literatürde de bu tür işlemlerin analiz doğruluğunu artırdığı ve veri gürültüsünü azalttığı sıklıkla vurgulanmaktadır (Manning vd., 2008; Rajaraman ve Ullman, 2011).

Stopwords ve *lemmatizasyon* işlemlerinin ardından, bazı satırlarda yalnızca boş string ("") içeren boş cümleler kalmıştır. Bu cümleler, bilgi içeriği taşımadıkları için analiz sürecine herhangi bir katkı sağlamamaktadır ve veri setinde gereksiz gürültüye neden olmaktadır. Aşağıda bu duruma ilişkin örnekler sunulmuştur:

Boş cümleler kaldırılmadan önce veri şu şekildedir:
["", "Edge cloud computing has developed rapidly.", "", "Various technologies have brought great convenience.", "", "We propose a deep learning model.", ""]

Boş cümleler kaldırıldıktan sonra ise sadeleşmiş hali şu şekilde olmuştur:
["Edge cloud computing has developed rapidly.", "Various technologies have brought great convenience.", "We propose a deep learning model."]

Bu işlem sayesinde yalnızca anlamlı içeriğe sahip cümleler analiz sürecine dahil edilmiş ve daha temiz, odaklanmış bir korpus oluşturulmuştur. Boş cümlelerin kaldırılması, veri setinin yapısal tutarlılığını artırmış; sözcük vektörlerinin oluşturulması, n-gram analizi ve model eğitimi gibi sonraki adımlar için daha güçlü bir temel sağlamıştır.

Literatürde de belirtildiği üzere, veri setinden anlamsız içeriklerin temizlenmesi, metin madenciliği süreçlerinin doğruluğunu ve etkinliğini önemli ölçüde artırmaktadır (Liu, Hu, ve Cheng, 2005). Bu doğrultuda oluşturulacak olan temiz ve anlamlı korpus, bir sonraki aşama olan LDA temelli konu modelleme süreci için sağlam bir zemin sunmaktadır.

4.3.7. Korpus ve Sözlük Oluşturma

Önceki adımlarda gerçekleştirilen stopwords temizliği, lemmatizasyon ve boş cümlelerin kaldırılması işlemleri sonucunda, içerik açısından zengin ve bağlamsal bütünlüğü korunmuş bir metin yapısı elde edilmiştir. Bu aşamada, temizlenmiş metinlerin analiz sürecine uygun hale getirilmesi amacıyla korpus ve sözlük oluşturma adımları uygulanmıştır. Bu süreç, metinlerin matematiksel olarak temsil edilmesini sağlayarak, tematik analizlerin doğruluğunu artırmış ve literatürün sistematik bir biçimde incelenmesine olanak tanımıştır.

Sözlük, metinlerdeki benzersiz kelimelere kimlik numarası atayan ve bu kelimeleri sayısal olarak temsil eden temel bir yapı olarak oluşturulmuştur. Her bir kelimeye özgü bir kimlik atanmış, böylece analiz süreçlerinde yalnızca bilgi değeri taşıyan içeriklerin dikkate alınması sağlanmıştır. Bu yapı, modelleme aşamaları için daha düzenli ve anlamlı bir veri temeli sunmuştur. Literatürde de Blei ve arkadaşları (2003), kelimelerin sıklık ve bağlamsal ilişkilerine göre düzenlenmesinin analiz süreçlerini optimize ettiğini vurgulamaktadır.

Sözlük yapısı oluşturulurken belirli eşik değerleri çerçevesinde kelime frekanslarına göre filtreleme yapılmıştır:

- **Nadir Kelimeler:** Veri setinde çok az yer alan ve bağlamsal anlam taşımayan ifadeler çıkarılmıştır.
- **Aşırı Sık Kullanılan Kelimeler:** Anlam dışı etki yaratabileceği düşünülen ve metinlerde tekrar eden kelimeler elenmiştir.

Bu yöntemle, analizde yalnızca tematik açıdan anlamlı kelimelere odaklanılmış ve modelleme sürecinin doğruluğu önemli ölçüde artırılmıştır.

Korpus ise metinlerin matematiksel bir forma dönüştürülmesini sağlayan analiz altyapısı olarak yapılandırılmıştır. Temizlenmiş belgeler, Bag-of-Words (BOW) temsili ile her kelimenin kimlik numarası ve frekans değeriyle ifade edilmiştir. Bu yapı, metinlerin sistematik analiz edilmesine olanak tanımakta ve özellikle LDA gibi modelleme yöntemlerinin uygulanmasını kolaylaştırmaktadır (Blei vd., 2003; HaCohen-Kerner vd., 2020).

Kelime frekansları, belgelerdeki tematik vurguları belirlemede kritik bir rol oynamaktadır:

- **Sıkça Geçen Kelimeler:** Literatürde öne çıkan ana temaları temsil etmektedir.
- **Nadiren Kullanılan Kelimeler:** Anlam derinliği oluşturmadığı için analizden çıkarılmaktadır.

Örnek Korpus Çıktısı:

Kelime Dağılımı → "advent": 1, "author": 1, "big": 1, "compute": 1, "computing": 1

Diğer Örnekler → "education": 1, "dataset": 1, "development": 1, "deep": 1

Bu örnek, metinlerdeki kelime frekanslarının nasıl sayısal olarak temsil edildiğini ve analiz sürecinde nasıl kullanıldığını göstermektedir. Bu yapı, tematik analizlerin doğruluğunu artırmada ve modelleme sürecinin etkinliğini yükseltmede önemli bir katkı sağlamıştır.

Elde edilen sözlük ve korpus yapısı, metinlerin konusal örüntülerini ortaya çıkarmak amacıyla yürütülecek LDA modellemesi için gerekli analitik zemini hazırlamıştır.

4.4. LDA Modelleme Süreci

Oluşturulan bu korpus temelinde, literatürdeki tematik yapıların tespiti ve akademik eğilimlerin analiz edilmesi amacıyla LDA modeli sıklıkla kullanılan güçlü bir tekniktir. Bu aşamada, veri seti üzerinde korpusun oluşturulması, modelin eğitilmesi ve sonuçların görselleştirilmesi gibi adımlar izlenmiş, böylece metinlerden anlamlı temalar çıkarılmıştır (Blei, Ng ve Jordan, 2003).

4.4.1. LDA Modelinin Eğitilmesi

LDA modeli, önceden yapılandırılmış sayısal veriler üzerinde eğitilmiştir. Modelin performansını artırmak ve elde edilen temaların güvenilirliğini sağlamak amacıyla çeşitli hiperparametreler dikkatle seçilmiştir. Özellikle:

- **num_topics=20:** Veri setinde yer alan içeriklerin tematik çeşitliliğini yansıtacak şekilde 20 konu belirlenmiştir.
- **random_state=100:** Her çalıştırmada tutarlı sonuçlar alınmasını sağlamıştır.
- **passes=30:** Modelin veriler üzerinden çoklu geçişlerle daha derin öğrenim sağlamasına imkân tanımıştır.

Tablo 12, modelin eğitiminde kullanılan parametreleri ve işlevlerini özetlemektedir.

Tablo 12. LDA model parametreleri ve işlevleri.

Parametre	Açıklama
num_topics=20	Analiz edilen konu sayısını belirler. Veri setine göre 20 anlamlı konu seçilmiştir.
random_state=100	Modelin her çalıştırıldığında aynı sonucu üretmesini sağlar. Tutarlılığı artırır.
update_every=1	Modelin her iterasyonda güncellenmesini sağlar. Daha hızlı ve verimli öğrenim sağlar.
chunksize=500	Her iterasyonda işlenecek belge sayısını belirler. Verimlilik sağlar.
passes=30	Veri seti üzerinden 30 kez geçiş yaparak derinlemesine öğrenim sağlar. Overfitting'i önler.
alpha='auto'	Konuların dağılımını otomatik optimize eder.
eta='auto'	Kelimelerin konularla ilişkisini otomatik optimize eder.
per_word_topics=True	Her kelimenin konu dağılımını görmeyi sağlar. Daha ayrıntılı analiz sunar.

Bu parametrelerle eğitilen model, literatürdeki tematik yapıları doğru bir şekilde belirlemiş ve metinlerin anlamlı gruplar halinde organize edilmesini sağlamıştır.

4.4.2. LDA Modeli Sonuçları ve Görselleştirilmesi

Bu çalışmada kullanılan LDA modeli, elde edilen tematik yapıların analizini daha görünür ve anlaşılır kılmak amacıyla pyLDavis aracı ile görselleştirilmiştir. Bu araç, LDA modelinin çıktılarının hem yapısal ilişkilerini (konular arası uzaklık ve benzerlikler) hem de tematik yoğunluklarını (anahtar kelimelerin baskınlığı) interaktif ve sayısal olarak incelemeye olanak sağlamaktadır.

Görselleştirme iki temel bileşenden oluşmaktadır:

- **Intertopic Distance Map (Sol Panel):** Bu grafik, model tarafından tespit edilen 20 temanın, çok boyutlu uzayda birbirine olan konumsal uzaklıklarını gösterir. Her daire bir temayı temsil ederken, büyüklüğü o temanın korpus üzerindeki marjinal katkısını ifade etmektedir.
- **Top-30 Most Salient Terms (Sağ Panel):** Bu grafik, analiz edilen tüm metinlerde en sık rastlanan ve ayırt edici özellik taşıyan 30 anahtar kelimeyi sıralar. Kelimelerin temalara göre frekans dağılımı ve temsil gücü bu panel üzerinden incelenebilmektedir.

İlgili görselleştirmelerin istatistiksel değerlendirmeleri bir sonraki bölümde sunulmaktadır. Analiz sonuçları Şekil 9’da detaylı olarak gösterilmiştir.

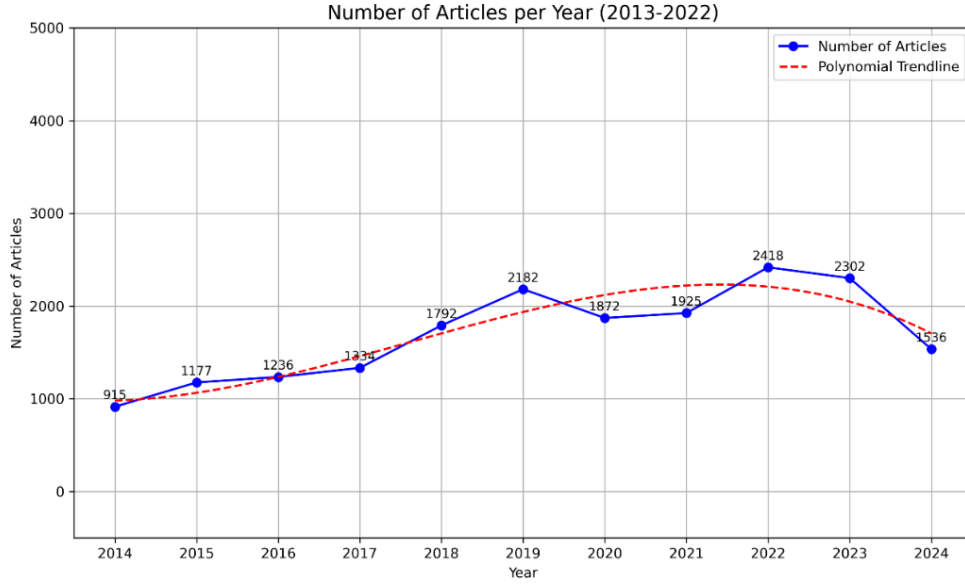
4.5. Bulut Bilişim Literatüründe Tematik ve İstatistiksel Eğilimler

Bu bölümde, BB alanında son on yılda yayımlanmış literatür LDA modeli desteğiyle kapsamlı biçimde analiz edilmiştir. Amaç; yalnızca içerikte öne çıkan temaları belirlemek değil, aynı zamanda bu temaların yayın frekansları, yıllara göre dağılımları ve literatürdeki önem derecelerini ortaya koymaktır. Tematik kümeler, frekans analizi ve trend görselleştirmeleri ile desteklenmiş, böylece BB araştırmalarının yapısal eğilimleri somut olarak sunulmuştur.

4.5.1. Yıllara Göre Makale Sayısının Analizi

Bu bölümde, BB alanında yayımlanan akademik çalışmaların 2014–2024 yılları arasındaki yıllara göre dağılımı analiz edilmiştir. Çalışma kapsamında toplam 18.689 akademik yayın incelenmiştir. Yayın eğilimlerini daha anlamlı sunmak amacıyla ikinci dereceden polinom eğri modellemesi uygulanmış, bu yöntemle verilerdeki dalgalanmalar yumuşatılarak uzun vadeli eğilimler ortaya konmuştur. Şekil 6, analiz edilen dönem boyunca BB literatüründeki yıllık üretim değişimlerini görsel olarak sunmaktadır.

- 2014 yılında yayın sayısı 915 olarak belirlenmiştir.
- 2015–2017 döneminde yayın sayıları sırasıyla 1177, 1236 ve 1337 olup, bu yıllarda istikrarlı bir artış kaydedilmiştir.
- 2018 yılında 1792 yayına ulaşılmış ve bu yıl itibarıyla dikkat çekici bir artış gözlemlenmiştir.
- 2019 yılında 2182 yayın ile bu yükseliş trendi devam etmiştir.
- 2020 yılında yayın sayısı 1872’ye gerilemiş; ancak 2021’de tekrar 1925 olarak artış göstermiştir.
- 2022 yılı, 2418 yayın ile literatürdeki en yüksek yayın sayısının gözlemlendiği zirve yılı olmuştur.
- 2023 yılında yayın sayısı 2302 olarak gerçekleşmiş; 2024 yılı itibarıyla ise 1536 yayın kaydedilmiştir.



Şekil 6. 2014–2024 yılları arasında BB konulu yayın sayılarının yıllık dağılımı

4.5.2. En Çok Yayın Yapan Dergilerin Analizi

BB literatüründe en çok yayın yapan dergilerin analizi, bu alandaki akademik katkıların yoğunlaştığı mecraları belirlemek açısından önemli bir kriterdir. Şekil 7’de, 2014–2024 yılları arasında BB konulu yayınlara en fazla katkıda bulunan ilk 10 akademik dergi ve bu dergilerin toplam yayınlar içindeki oranları görselleştirmiştir.

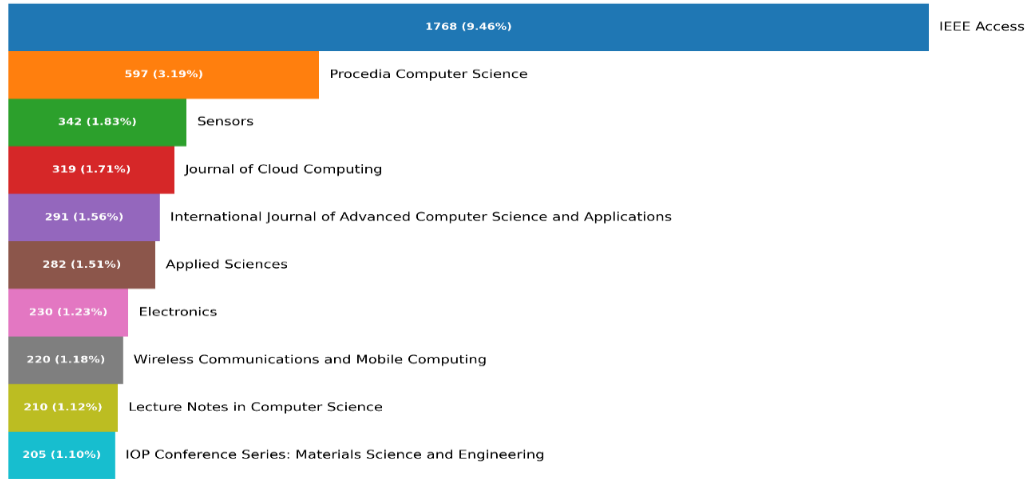
Bu analizde:

- IEEE Access dergisi, 1768 yayın ve %9,46’lık oranla alanda en fazla katkı sağlayan dergi olarak öne çıkmaktadır. Açık erişim yapısı ve disiplinler arası kapsayıcılığı, bu derginin tercih edilmesinde etkili olduğu değerlendirilmektedir.
- Procedia Computer Science, 597 yayın (%3,19) ile ikinci sırada yer almakta olup, hızlı bildiri yayımı özelliğiyle dikkat çekmektedir.
- Sensors dergisi ise 342 yayın (%1,83) ile üçüncü sırada yer almakta olup, BB literatürüne kayda değer katkı sunduğu değerlendirilmektedir.

Bu üç dergi, toplam yayınların yaklaşık %14,48’ini temsil ederek literatürdeki tematik odakların yönünü belirlemede etkin rol oynamaktadır. Bunun yanı sıra, *Journal of Cloud Computing*, *Applied Sciences*, *International Journal of Advanced Computer Science and Applications* gibi dergiler de önemli yayın hacimleriyle çeşitlilik sunmaktadır.

Sunulan grafik sayesinde, araştırmacılar kendi çalışmalarını hangi platformlarda yayımlayacaklarına yönelik daha stratejik kararlar alabilirken, literatürdeki akademik yoğunluk ve eğilimler hakkında da genel bir farkındalık kazanabilir.

Top 10 Journals by Number of Publications



Total Publications: 18689

Şekil 7. 2014–2024 yılları arasında BB alanında en çok yayın yapan dergiler

4.5.3. Bulut Bilişim Araştırmalarının Coğrafi Dağılımı

BB alanında gerçekleştirilen akademik çalışmaların coğrafi dağılımı, küresel araştırma eğilimlerinin anlaşılması açısından önemli bir gösterge sunmaktadır. Bu kapsamda, 2014–2024 döneminde en fazla yayın gerçekleştiren ilk 10 ülkenin verileri analiz edilmiş ve Şekil 8’de görselleştirilmiştir.

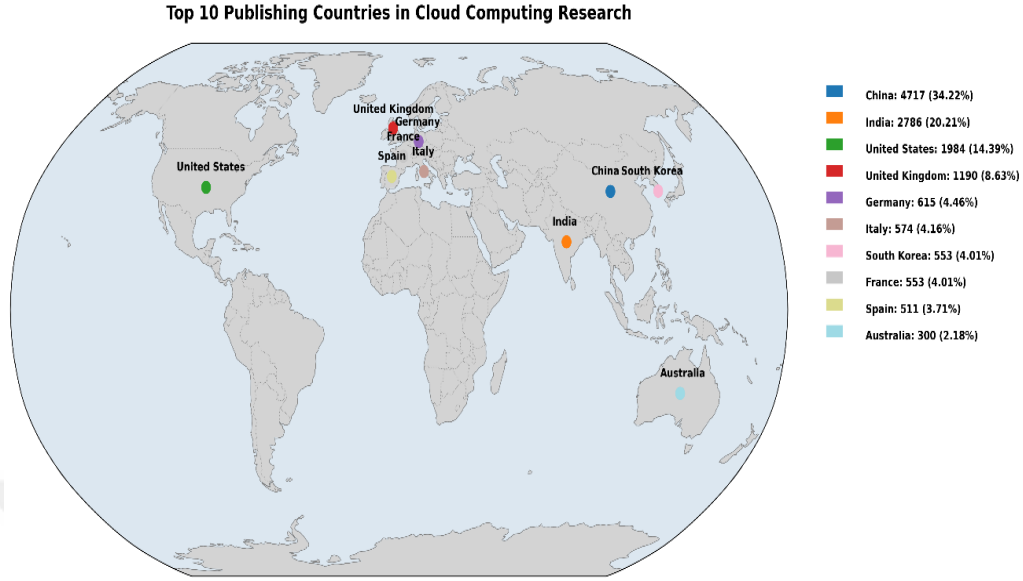
Analiz bulgularına göre, Çin 4177 yayın (%34,22) ile BB alanında en fazla katkı sağlayan ülke konumundadır. Hindistan, 2766 yayınla toplam yayınların %20,21’ini oluşturarak ikinci sırada yer almaktadır. Amerika Birleşik Devletleri ise 1984 yayın (%14,39) ile üçüncü sırada bulunmaktadır. Bu üç ülke toplamda literatürün yaklaşık %68,82’sini üretmiş ve alandaki araştırma yoğunluğunu büyük ölçüde belirlemiştir.

Avrupa ülkeleri ise daha dengeli bir dağılım sergilemektedir. Birleşik Krallık 1109 yayın (%8,63), Almanya 615 yayın (%4,46), İtalya 574 yayın (%4,14), Fransa 553 yayın (%4,03) ve İspanya 511 yayın (%3,71) ile öne çıkan Avrupa ülkeleri arasında yer almaktadır.

Asya-Pasifik bölgesinden Güney Kore 553 yayınla toplam yayınların %4,03’ünü üretmiş; Avustralya ise 300 yayınla %2,18 oranında katkı sağlamıştır.

Bu veriler ışığında değerlendirildiğinde, BB literatürünün büyük ölçüde Asya ve Kuzey Amerika ekseninde şekillendiği, Avrupa’nın ise istikrarlı ve çok yönlü bir katkı sunduğu görülmektedir. Çin ve Hindistan gibi ülkelerin yüksek yayın hacmi, özellikle

bu coğrafyalarda BB teknolojilerine yönelik artan yatırım ve akademik ilgiyi yansıtmaktadır.



Şekil 8. 2014–2024 yılları arasında Bulut bilişim alanında en çok yayın yapan 10 ülkenin coğrafi dağılımı.

4.5.4. BB Yayınlarının LDA ile Analizi

Bu bölümde, LDA modeli ile elde edilen BB literatürüne ait tematik kümelenme ve anahtar kelime frekansları, istatistiksel olarak değerlendirilmiştir. Modelin çıktıları, iki bileşenli görselleştirme üzerinden analiz edilmiştir:

1. Tematik Yoğunluk – Intertopic Distance Map:

Şeklin sol tarafında yer alan Intertopic Distance Map, elde edilen 20 temanın iki boyutlu düzlemdeki konumlarını sunmaktadır. Her bir daire, bir temayı temsil ederken, büyüklükleri o temaların toplam korpus içindeki marjinal dağılımlarını ifade etmektedir.

- **Yakın konumlu temalar (1, 2, 3):** Sıklıkla işlenen, içeriksel olarak birbiriyle ilişkili konulara işaret etmektedir.
- **Uzak konumlu temalar (15, 18, 20):** Niş araştırma alanlarını ve daha az işlenmiş özgül tematik kümeleri göstermektedir.

2. Anahtar Kelime Frekans Dağılımı – Top-30 Most Salient Terms:

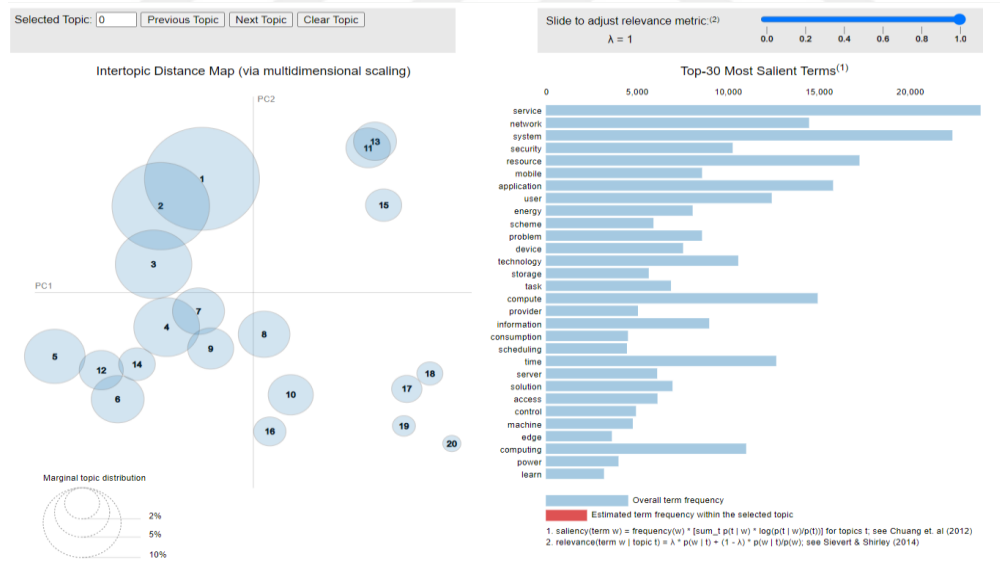
Şeklin sağ paneli, analiz edilen metinlerde en sık karşılaşılan 30 terimi göstermektedir. $\lambda = 1$ ayarında oluşturulan grafik, yalnızca kelime frekansına göre sıralama yapmaktadır.

Şekil 9’da sunulan Top-30 Most Salient Terms grafiği incelendiğinde, BB literatüründe öne çıkan anahtar kelimelerin frekans ve tematik belirleyicilik açısından

nasıl dağıldığı istatistiksel olarak analiz edilmiştir. Görsel çıktılar üzerinden yapılan değerlendirmeye göre:

- En yüksek frekansa sahip terim: “*service*” olup, yaklaşık 21. 000 kez geçmektedir.
- Bunu sırasıyla “*network*” (~19.000), “*system*” (~18.500), “*security*” (~17.000) ve “*resource*” (~15.500) gibi literatürde baskın kavramlar takip etmektedir.
- İlk 10 anahtar kelimenin ortalama frekansı yaklaşık 16.800 olarak hesaplanmıştır.
- En düşük frekansa sahip terim ise “*learn*” olup, yaklaşık 4.000 kez geçmiştir.
- Kelime frekanslarındaki dağılımın çeşitliliği, yaklaşık 4.600 birimlik standart sapma değeri ile dikkat çekmektedir. Bu durum, analiz edilen belgelerde bazı kavramların belirgin şekilde baskın olduğunu, diğerlerinin ise daha sınırlı kullanımda kaldığını ortaya koymaktadır.

Bu istatistiksel veriler, LDA modeli tarafından elde edilen temaların metinsel yoğunluklarını ve temalar arası kavramsal farklılıkları anlamada önemli bir zemin sunmaktadır.



Şekil 9. LDA modeline dayalı tematik dağılım haritası (Intertopic Distance Map) ve en baskın anahtar kelimeler grafiği (Top-30 Most Salient Terms)

4.5.5. LDA Modelinden Türetilen Genel Kelime Bulutu

Önceki bölümde sunulan tematik dağılımların ardından, bu alt bölümde LDA modelinden türetilen genel kelime bulutu analiz edilmektedir. Şekil 10’da yer alan bu görsel, tüm konulardan elde edilen anahtar kelimelerin yoğunluğunu ve sıklığını görsel olarak özetlemektedir. Analiz, literatürde en çok geçen, semantik olarak baskın ve LDA

4.5.6. LDA Tabanlı Tematik Yapılar ve Frekans Analizi

Bu bölümde, LDA modeli kullanılarak BB literatüründe öne çıkan tematik yapılar, içeriksel örüntülere dayalı olarak istatistiksel biçimde sınıflandırılmıştır. LDA modeliyle elde edilen 20 ana tema, ilgili anahtar kelimeler üzerinden adlandırılmış olup, bu yapılar Tablo 13'te özetlenmiştir. Bu anahtar kelimeler, ilgili başlıkların tematik eksenini oluşturmakta ve konuların anlamlandırılmasında temel rol üstlenmektedir.

Tablo 13. LDA modelinden türetilen 20 ana konu ve anahtar kelimeleri

No	Konu Başlığı	Anahtar Kelimeler
1	Software Tools and Web Development	software, development, tools
2	Risk Factors and Organizational Adoption	risk, adoption, organization
3	Virtual Machine Migration and Resource Management	virtual, machine, resource
4	Cloud Computing Technologies and Innovations	cloud, technology, innovation
5	Encryption and Security Mechanisms	encryption, security, mechanism
6	Functionality and Workflow Optimization	workflow, optimization, functionality
7	Cloud Services and User Engagement	services, user, engagement
8	Patient Health Monitoring and Medical Systems	health, monitoring, medical
9	Network Traffic Control and Data Communication	network, traffic, communication
10	Resource Allocation and Scheduling Optimization	resource, allocation, scheduling
11	Mobile Devices and Edge Computing Technologies	mobile, devices, edge
12	Information Technology and Industrial Applications	technology, information, industrial
13	Machine Learning and Training Processes	machine, learning, training
14	Energy Efficiency and Power Management	energy, efficiency, power
15	Image Analysis and Graphical Representation	image, analysis, graphical
16	Prediction Accuracy and Detection Algorithms	prediction, accuracy, detection
17	Application Frameworks and System Architectures	application, frameworks, architecture
18	Security, Privacy, and User Access Management	security, privacy, access
19	Data Storage and Big Data Analytics	data, storage, analytics
20	Proposal-Based Frameworks and Applications	proposal, framework, applications

Her temada, metin kümesi içerisinde ortak geçen anahtar kavramlara göre başlıklandırma yapılmıştır. Bu başlıklar; yazılım geliştirme, güvenlik mekanizmaları,

kaynak yönetimi, veri analitiği, kenar bilişim, kullanıcı etkileşimi gibi çok sayıda tematik kümeyi kapsamaktadır.

Temalar aynı zamanda frekanslarına göre üç kategori altında sınıflandırılmıştır. Temaların metin kümesindeki marjinal dağılımlarına göre elde edilen frekans oranları aşağıda listelenmiştir:

Frekans Değeri En Yüksek Temalar:

- Emerging Technologies in Cloud Computing – **0.2212**
- Distributed Applications and Frameworks – **0.1602**

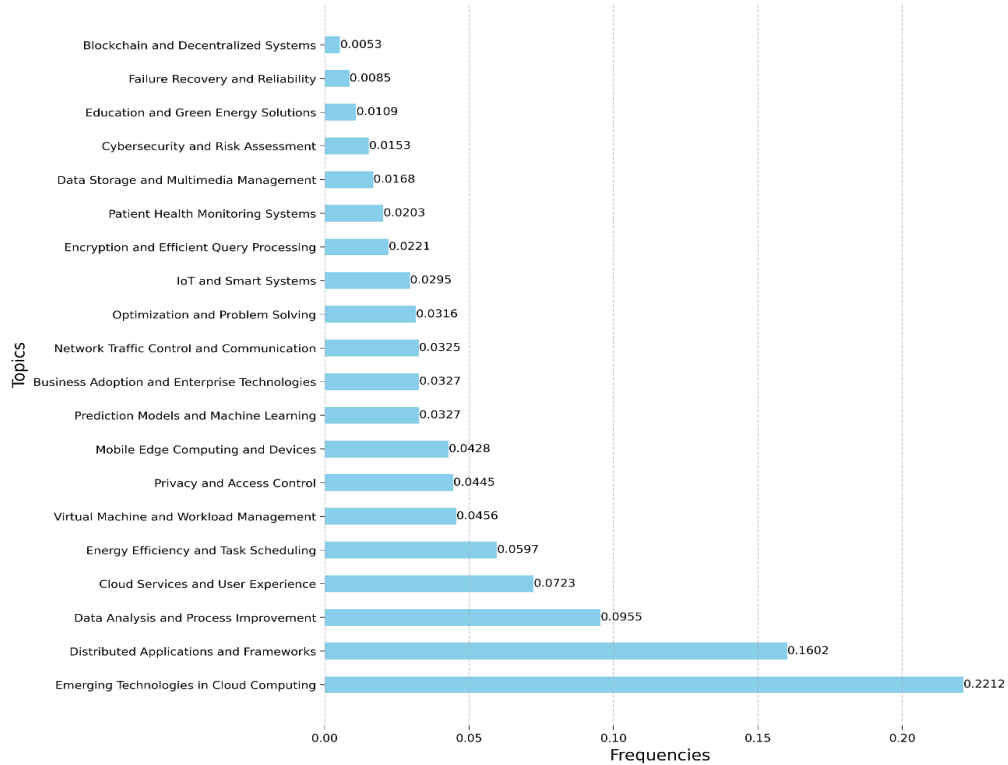
Orta Düzey Frekansa Sahip Temalar:

- Cloud Services and User Experience – **0.0723**
- Energy Efficiency and Task Scheduling – **0.0597**

Frekansı Düşük Temalar:

- Patient Health Monitoring Systems – **0.0203**
- Blockchain and Decentralized Systems – **0.0053**

Bu frekans değerleri, temaların tüm metin korpusu içerisindeki ağırlıklarını göstermekte olup, LDA modeli tarafından hesaplanan oranlara dayanmaktadır. Elde edilen istatistiksel dağılımın görsel temsili Şekil 11'de sunulmuştur.



Şekil 11. 2010-2014 yılları arasında yayınlanan bulut bilişim konularının frekans dağılımı.

4.5.7. BB Konuları için Kelime Kümesi Analizi

Bu bölümde, LDA modeli ile belirlenen anahtar kelime kümeleri üzerinden BB literatüründeki tematik yapılar ve bu yapılar arasındaki bağlantılar analiz edilmiştir. Şekil 12, her bir temanın öne çıkan anahtar kelimelerini ve bu temaların birbirleriyle olan ilişkilerini çok boyutlu bir ağ yapısı içerisinde göstermektedir. Analizde toplamda 20 farklı tematik küme yer almakta olup, bu kümeler belirgin biçimde renk kodlamasıyla gruplandırılmıştır.

Görsel Bileşenlerin Yapısı.

- Her bir renkli düğüm grubu (node cluster), belirli bir tematik başlığı temsil etmektedir.
- Düğümler, temada öne çıkan anahtar kelimeleri göstermektedir.
- Düğümler arası bağlantılar (kenarlar), anahtar kelimeler arasında metin korpusunda ortak kullanım ilişkisine işaret etmektedir.
- Düğümlerin yoğunlaştığı bölgeler, metinlerde daha sık geçen ve birbirleriyle daha fazla ilişkilendirilen kavramları temsil etmektedir.

Ağ Grafiğinde Temsil Edilen Temalar. Şekil 12’de listelenen 20 başlık aşağıda sunulmuştur:

1. Virtual Machine and Workload Management
2. Cybersecurity and Risk Assessment
3. Patient Health Monitoring Systems
4. Blockchain and Decentralized Systems
5. Data Storage and Multimedia Management
6. Energy Efficiency and Task Scheduling
7. Encryption and Efficient Query Processing
8. Cloud Services and User Experience
9. Education and Green Energy Solutions
10. Emerging Technologies in Cloud Computing
11. Distributed Applications and Frameworks
12. Failure Recovery and Reliability
13. Prediction Models and Machine Learning
14. Network Traffic and Communication
15. IoT and Smart Systems
16. Mobile-Edge Computing and Devices
17. Privacy and Access Control
18. Business Adoption and Enterprise Technologies

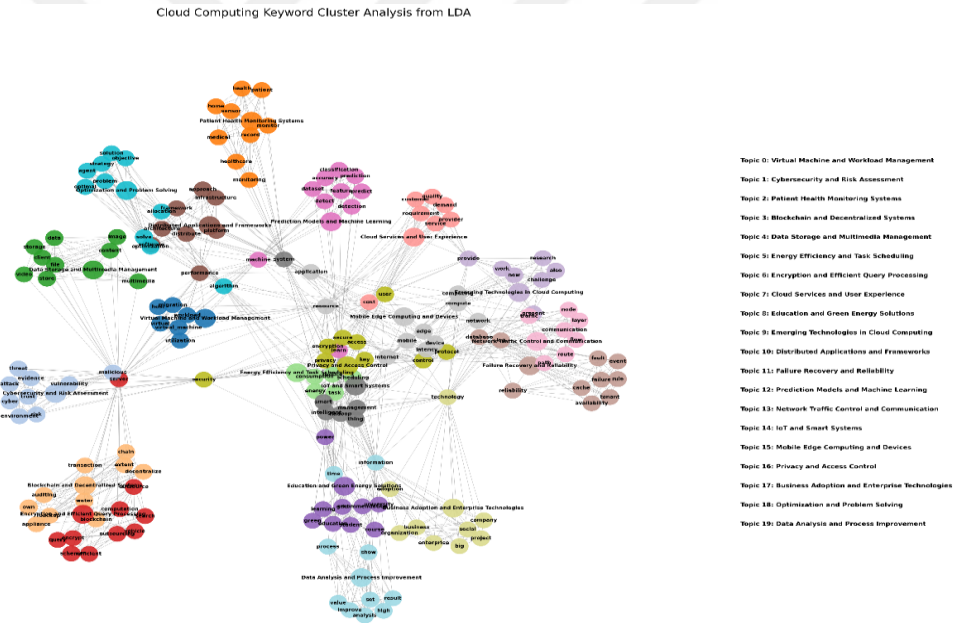
19. Optimization and Problem Solving

20. Data Analysis and Process Improvement

Gözlenen Ağ Özellikleri.

- Tematik kümelerin bir kısmı merkezi konumda yer almakta ve diğer birçok temayla ilişkili görünmektedir (örneğin: Topic 10, Topic 11, Topic 8).
- Daha uzak konumda bulunan kümeler (örneğin: Topic 4 – Blockchain, Topic 3 – Health Monitoring), daha bağımsız tematik alanları temsil etmektedir.
- Ağ bağlantı yoğunluğu açısından incelendiğinde, belirli anahtar kelimeler (örneğin: “cloud,” “management,” “system,” “data”) birden fazla kümede yer almakta olup, genel literatürde merkezi konulara işaret etmektedir.

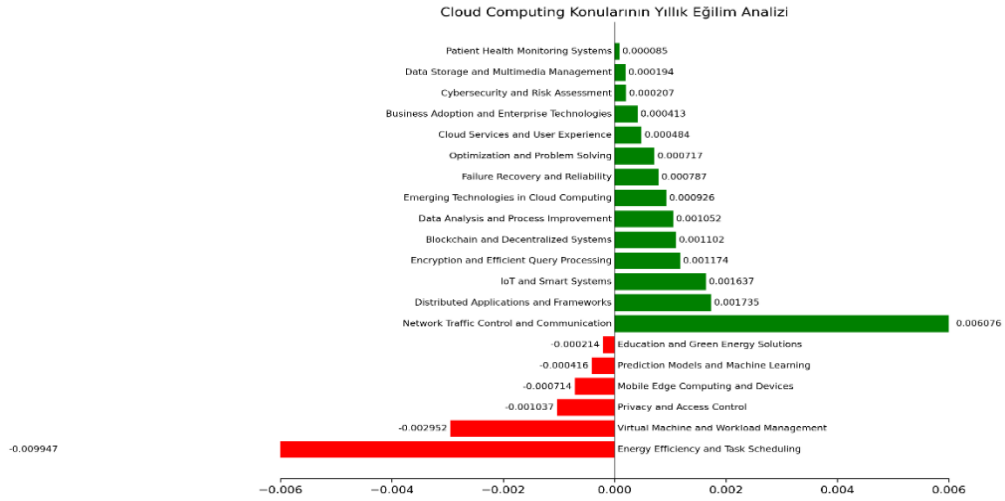
İlgili yapısal tematik analiz, literatürdeki anahtar kavramların ilişkisel düzeyde kümelendiğini ve bu kümelerin birbirine olan mesafeleriyle bütünsel bir bağlam sunduğunu göstermektedir. Görsel analiz çıktısı aşağıda sunulmaktadır:



Şekil 12. Bulut bilişim literatüründe LDA tabanlı tematik kelime kümeleme ağı

4.5.8. Konu Dağılımları ve Trend Analizi

Önceki bölümlerde belirlenen temaların istatistiksel karşılaştırmaları bu bölümde trend analizleri ve yıllık dağılımlar üzerinden detaylandırılarak sunulmaktadır. Elde edilen bulgular, literatürdeki temaların artış veya azalış yönündeki hareketlerini sayısal değerlerle ortaya koymaktadır. Şekil 13 ve Tablo 14, bu dağılımları hem grafiksel hem de sayısal olarak sunmaktadır.



Şekil 13. Bulut bilişim araştırma konularının yıllık eğilim analizi (2014–2024)

Yükselen Temalar. Trend analizine göre bazı konular son on yılda anlamlı bir yükseliş göstermiştir. Özellikle “Network Traffic Control and Communication” başlığı +0.006076 ile en yüksek artış eğilimine sahiptir. Bunu sırasıyla +0.001735 ile “Distributed Applications and Frameworks” ve +0.001637 ile “IoT and Smart Systems” konuları takip etmektedir. Bu konular, yıl bazlı frekans değerlerinde istikrarlı bir artış sergilemiş ve literatürdeki ağırlıklarını her yıl artırmıştır.

Düşüş Gösteren Temalar. Negatif trend gösteren konular arasında “Energy Efficiency and Task Scheduling” (-0.009947), “Virtual Machine and Workload Management” (-0.002952) ve “Privacy and Access Control” (-0.001037) dikkat çekmektedir. Bu temalar 2014–2024 dönemi boyunca frekans açısından azalan bir seyir izlemiş, yıllar içerisinde görece önemlerini yitirmiştir.

Dengeli veya Kararsız Eğilim Gösteren Temalar. Bazı konuların eğilimleri sınırlı artış veya azalış göstermiştir. Örneğin “Cloud Services and User Experience” (+0.000484) ve “Optimization and Problem Solving” (+0.000717) gibi temalar, literatürde istikrarlı bir biçimde varlığını sürdürmekle birlikte belirgin bir değişim göstermemiştir. Aynı şekilde “Prediction Models and Machine Learning” (-0.000416) gibi bazı konular da çok hafif azalma trendiyle dikkat çekmektedir.

Tablo 14. Bulut bilişim konularının yıllık frekans değerleri ve trend skorları

	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	Trend
Network Traffic Control and Communication	0.030403	0.034799	0.037705	0.038955	0.046485	0.048969	0.062656	0.067448	0.077707	0.083078	0.086823	0.006076
Distributed Applications and Frameworks	0.031053	0.037644	0.040479	0.050858	0.055139	0.053268	0.054795	0.057064	0.059663	0.049248	0.049612	0.001735
IoT and Smart Systems	0.035184	0.042736	0.038864	0.043229	0.045326	0.046761	0.038368	0.051018	0.049230	0.056068	0.052596	0.001637
Encryption and Efficient Query Processing	0.019668	0.017629	0.016887	0.017003	0.021038	0.020944	0.025989	0.027719	0.027747	0.026674	0.026475	0.001174
Blockchain and Decentralized Systems	0.094062	0.094412	0.091111	0.090016	0.089881	0.095752	0.096561	0.097020	0.102289	0.099069	0.103732	0.001102
Data Analysis and Process Improvement	0.013584	0.015688	0.016131	0.020942	0.023591	0.021913	0.024365	0.026512	0.026508	0.024373	0.021162	0.001052
Emerging Technologies in Cloud Computing	0.010314	0.012562	0.011517	0.012114	0.013441	0.014610	0.017635	0.016347	0.018510	0.017217	0.020225	0.000926
Failure Recovery and Reliability	0.010841	0.010983	0.012763	0.014176	0.012996	0.015005	0.016984	0.017246	0.018619	0.016311	0.018353	0.000787
Optimization and Problem Solving	0.021215	0.023629	0.023572	0.021556	0.022600	0.027769	0.025002	0.023582	0.025503	0.028373	0.030736	0.000717
Cloud Services and User Experience	0.038665	0.030736	0.037073	0.039607	0.051941	0.048135	0.050019	0.043830	0.041076	0.040997	0.037395	0.000484
Business Adoption and Enterprise Technologies	0.028111	0.033727	0.040920	0.043258	0.041277	0.038647	0.037697	0.032941	0.038245	0.040152	0.038506	0.000413
Cybersecurity and Risk Assessment	0.035993	0.036962	0.039166	0.040066	0.041973	0.037913	0.040541	0.038470	0.042097	0.038055	0.038836	0.000207
Data Storage and Multimedia Management	0.022695	0.024253	0.019989	0.020808	0.019725	0.020498	0.022401	0.024794	0.028871	0.020734	0.022320	0.000194
Patient Health Monitoring Systems	0.007832	0.008581	0.005853	0.007709	0.009222	0.008062	0.009474	0.008420	0.008125	0.008250	0.008274	0.000085
Education and Green Energy Solutions	0.024390	0.023278	0.021970	0.021333	0.017584	0.019288	0.017415	0.020695	0.020703	0.021892	0.021842	-0.000214
Prediction Models and Machine Learning	0.031586	0.029391	0.033987	0.031921	0.032476	0.031296	0.030289	0.029178	0.028716	0.027980	0.028257	-0.000416
Mobile Edge Computing and Devices	0.041961	0.033599	0.037588	0.035018	0.037158	0.039919	0.036069	0.036843	0.034379	0.033949	0.027376	-0.000714
Privacy and Access Control	0.038222	0.032047	0.032507	0.032174	0.028714	0.028486	0.028662	0.027240	0.027797	0.024794	0.026031	-0.001037
Virtual Machine and Workload Management	0.080634	0.082292	0.078816	0.077808	0.073214	0.069143	0.063725	0.061006	0.058677	0.058596	0.055338	-0.002952
Energy Efficiency and Task Scheduling	0.342506	0.333885	0.323865	0.304321	0.280750	0.278193	0.268512	0.251311	0.241867	0.254329	0.256163	-0.009947

Tablo 14'te yer alan veriler, 20 ana konunun her yılıki göreceli frekanslarını (oran bazlı) ve bunların trend skorlarını içermektedir. Örnek olarak “Energy Efficiency and Task Scheduling” konusu 2014 yılında %34.25 gibi yüksek bir paya sahipken, 2024 yılına gelindiğinde bu oran %25.61’e düşmüş; böylece trend değeri -0.009947 olarak hesaplanmıştır. Diğer yandan “Network Traffic Control and Communication” konusu 2014’te %3.4 oranındayken, 2024’te bu oran %8.68’e yükselmiş ve +0.006076’lık pozitif bir trend göstermiştir.

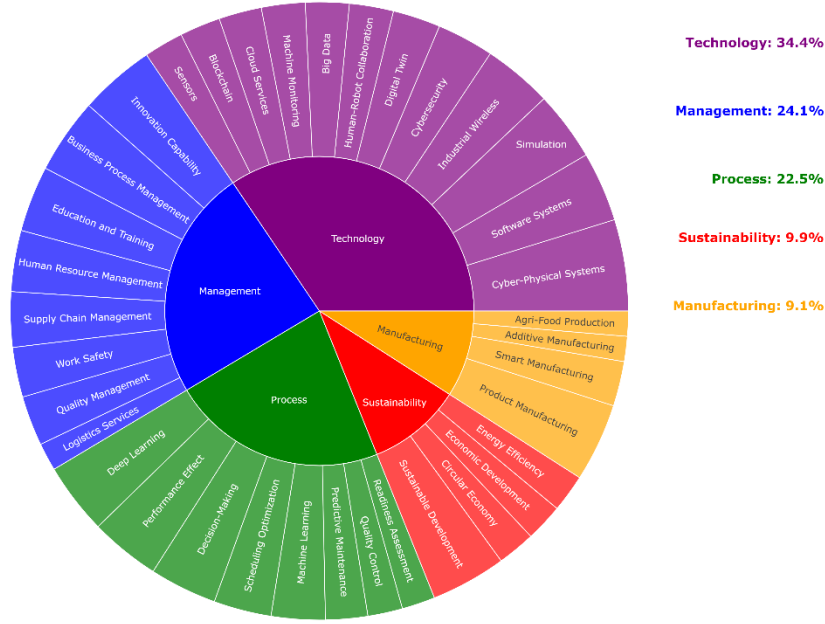
Bu sayısal bulgular, her temanın literatürdeki gelişim sürecini yıl bazında analiz etmeye olanak sağlamakta ve konular arası dinamiklerin karşılaştırılmasına imkân sunmaktadır. Böylece, hangi konuların zamanla öne çıktığı ya da önemini yitirdiği istatistiksel olarak tespit edilmiştir.

4.5.9. BB Taksonomi Görselleştirmesi

Bu bölümde, Bulut Bilişim (BB) literatüründeki tematik yapıların hiyerarşik sınıflandırması gerçekleştirilmiştir. Latent Dirichlet Allocation (LDA) analizinden elde edilen temalar, içerik benzerliği esas alınarak beş ana kategori altında gruplanmıştır: **Technology (Teknoloji)**, **Process (Süreç)**, **Sustainability (Sürdürülebilirlik)**, **Manufacturing (Üretim)** ve **Management (Yönetim)**. Bu tematik yapı, Şekil 14’te sunulan **sunburst (güneş ışını) tipi diyagram** kullanılarak görselleştirilmiştir.

Sunburst görselleştirmesi, üst düzey temaları ve bu temalar altında yer alan alt konuları hiyerarşik biçimde temsil etmektedir. Her bir segmentin büyüklüğü, ilgili

konunun literatürdeki temsil oranına karşılık gelecek şekilde orantısal olarak belirlenmiştir.



Şekil 14. Bulut bilişim temalarının taksonomik dağılımı

Şekil 14, BB araştırmalarında ele alınan temaların yapısal dağılımını bütüncül biçimde sunmakta; üst kategoriler ve bunlara bağlı alt temalar arasındaki ilişkileri sezgisel olarak görselleştirmektedir. Görselin detaylı değerlendirmesi ve tematik dağılım sonuçları bir sonraki bölümde açıklanmıştır. Ayrıca, her ana temanın öne çıkan alt başlıkları da ilgili alt bölümlerde ayrıntılı olarak ele alınmıştır.

Bu bölümde gerçekleştirilen tematik sınıflandırma, frekans analizi ve eğilim incelemeleri; BB literatüründe öne çıkan konuları, yapısal yoğunlukları ve zamansal değişimleri istatistiksel olarak ortaya koymuştur. Elde edilen bu bulgular doğrultusunda, şimdi bulguların anlamı, literatüre katkısı ve gelecekteki araştırmalara yönelik çıkarımların ele alındığı “Sonuçlar ve Tartışma” bölümüne geçilmektedir.

5. SONUÇLAR VE DEĞERLENDİRME

Bu çalışma kapsamında, 2014–2024 yılları arasında BB alanında yayımlanan 18.689 akademik yayının tematik dağılımı ve zaman içindeki eğilimleri detaylı biçimde incelenmiştir. Analiz süreci, iki temel eksen doğrultusunda yapılandırılmıştır.

İlk olarak, bibliyometrik analiz çerçevesinde yayın sayılarının yıllara göre dağılımı, en çok yayın yapan dergiler ve ülkeler gibi sayısal veriler değerlendirilmiştir. Bu veriler, BB alanındaki üretim yoğunluğunun zamanla nasıl şekillendiğini ve coğrafi olarak hangi bölgelerde yoğunlaştığını ortaya koymaktadır.

İkinci olarak ise tematik içerik analizi gerçekleştirilmiş ve bu doğrultuda Latent Dirichlet Allocation (LDA) yöntemi kullanılarak konusal kümelenmeler oluşturulmuştur. Bu model, literatürde öne çıkan temaları, bu temaların birbirleriyle olan ilişkilerini ve zamansal değişimlerini detaylı biçimde analiz etme imkânı sunmuştur. Böylece çalışma, BB alanındaki akademik üretimin niceliksel yönleri ile niteliksel eğilimlerini bir araya getirerek kapsamlı bir değerlendirme sunmaktadır.

Bibliyometrik Bulguların Analizi:

Bu çalışma kapsamında, 2014–2024 yılları arasında yayımlanan BB odaklı 18.689 akademik yayın incelenmiş ve yıllık yayın eğilimleri, dergilerdeki dağılım ve coğrafi üretkenlik gibi temel bibliyometrik göstergeler değerlendirilmiştir.

Yıllık yayın sayıları üzerine gerçekleştirilen zaman serisi analizinde, 2014–2019 yılları arasında belirgin bir artış eğilimi gözlemlenmiştir. Bu artışın, BB teknolojilerinin hem akademik çevrelerde hem de sektörel uygulamalarda hızla yaygınlaşmasıyla ilişkili olduğu düşünülebilir. 2020 yılı itibarıyla yayın sayılarında gözlenen düşüş, büyük olasılıkla küresel COVID-19 pandemisinin araştırma üretkenliği üzerindeki olumsuz etkilerini yansıtmaktadır. İzleyen yıllarda bu düşüş eğilimi tersine dönmüş; özellikle 2022 yılında 2418 yayına ulaşılması, BB alanında yeniden güçlü bir araştırma ivmesi yakalandığını göstermektedir. Öte yandan, 2024 yılına ilişkin veri setinde gözlemlenen görece azalma, veri toplama işlemlerinin yıl tamamlanmadan gerçekleştirilmiş olmasıyla açıklanabilir (bkz. Şekil 6).

Dergi dağılımı analizinde, IEEE Access dergisinin 1768 yayınlı toplam literatürün yaklaşık %9,46'sını oluşturduğu belirlenmiştir. Bu durum, açık erişim politikalarının BB alanındaki akademik görünürlüğü artırmadaki etkisine işaret etmektedir. IEEE Access'in disiplinler arası yayın yapısı, BB çalışmalarının farklı alanlardan gelen

araştırmacılar tarafından ele alınmasına olanak tanımaktadır. Ayrıca, Procedia Computer Science, Sensors ve Journal of Cloud Computing gibi dergilerin yayın hacmi, uygulama odaklı ve teknik çalışmaların da BB literatüründe önemli bir yer tuttuğunu göstermektedir. Bu çeşitlilik, alanın hem kuramsal hem de uygulamalı yönlerinin birlikte geliştiğini ortaya koymaktadır.

Ülke bazlı üretkenlik analizinde, Çin ve Hindistan sırasıyla %34,2 ve %20,2 oranında katkı sağlayarak, küresel BB literatürünün yarısından fazlasını oluşturmuştur. Bu bulgular, Asya ülkelerinin son yıllarda BB alanına yaptığı yatırımların ve geliştirdiği araştırma politikalarının etkili sonuçlar verdiğini düşündürmektedir. Amerika Birleşik Devletleri ise %10,6'lık katkı oranıyla, güçlü teknolojik altyapısı ve Google Cloud, AWS gibi lider bulut hizmet sağlayıcılarına ev sahipliği yapması sayesinde, küresel rekabette güçlü bir aktör konumunu sürdürmektedir. Avrupa ülkeleri ise daha dengeli ve çok yönlü bir katkı sunmakta olup; Birleşik Krallık, Almanya, Fransa ve İtalya gibi ülkeler hem teorik hem de uygulamalı çalışmaları destekleyen çeşitli araştırma çıktıları üretmektedir.

Bu genel bibliyometrik bulgular, BB alanındaki akademik üretimin hem küresel hem de sektörel gelişmelerle nasıl paralel bir biçimde şekillendiğini göstermekte ve sonraki tematik analizlere sağlam bir temel oluşturmaktadır.

Tematik İçerik Analizi:

Bibliyometrik bulguların ardından, BB alanında yayımlanan çalışmaların içeriksel yapısına daha derinlemesine bakmak amacıyla tematik analiz süreci gerçekleştirilmiştir. Bu kapsamda, LDA yöntemi kullanılarak oluşturulan konu kümeleri üzerinden literatürde öne çıkan temalar, anahtar kavramlar ve bunlar arasındaki ilişkiler ayrıntılı biçimde değerlendirilmiştir. Aşağıdaki analizler, LDA modelinin sunduğu çok boyutlu veri yapılarının istatistiksel ve görsel çıktıları temel alınarak yapılandırılmıştır.

LDA modelinin çıktısı olarak elde edilen Top-30 Most Salient Terms grafiği (bkz. Şekil 9), BB literatüründe öne çıkan kavramları hem sıklık hem de tematik belirleyicilik açısından özetlemektedir. Grafik, yalnızca metinlerde en sık karşılaşılan terimleri göstermekle kalmayıp, aynı zamanda bu terimlerin LDA temaları içindeki bağlamsal ağırlıklarını da ortaya koymaktadır.

Grafikte ön plana çıkan “service”, “network”, “system” ve “security” gibi yüksek frekanslı terimler, BB sistemlerinin temel bileşenlerini temsil etmektedir. Örneğin, “service” ve “system” kavramları genellikle Cloud Services and User Experience temasıyla ilişkilendirilirken, “security” terimi Encryption and Access Management ya da Cybersecurity gibi tematik kümelerle örtüşmektedir.

Buna karşılık, “learn”, “model” ve “prediction” gibi daha düşük frekanslı terimler, LDA tarafından tanımlanan Machine Learning and Training Processes veya Prediction Accuracy and Detection Algorithms gibi yükselen temalarla ilişkilendirilebilmektedir. Bu kavramların sınırlı temsil oranları, literatürde henüz geniş kapsamlı biçimde ele alınmadıklarını göstermekte; ancak AI ve veri analitiğine yönelik artan akademik ilginin etkisiyle, bu temaların önümüzdeki yıllarda daha fazla ön plana çıkması beklenmektedir.

Grafikte dikkat çeken bir diğer unsur, bazı terimlerin birden fazla tematik gruba örtüşebilmesidir. Örneğin, “data” ve “information” gibi kavramlar hem Data Storage and Analytics hem de IoT and Smart Systems temaları içinde yer alabilmektedir. Bu durum, LDA modelinin yalnızca konuları ayırtırmakla kalmayıp, aynı zamanda temalar arası kavramsal geçişkenlikleri de başarıyla temsil ettiğini göstermektedir.

Ayrıca, ilk 10 terimin ortalama frekansı yaklaşık 16.800 civarındayken, en düşük frekanslı terimlerden biri olan “learn” yaklaşık 4.000 tekrar ile listelenmiştir. Frekans dağılımı incelendiğinde yaklaşık 4.600 birimlik standart sapma değeri, literatürde belirli kavramların yoğun biçimde yer aldığına, bazılarının ise marjinal kaldığına işaret etmektedir. Bu varyans, kavramsal yoğunluk ve içeriksel çeşitlilik açısından dikkate değer bir bulgudur.

Bu doğrultuda, LDA yöntemiyle elde edilen çıktılar temel alınarak yapılan kelime bulutu analiziyle, BB literatüründe öne çıkan kavramların genel yoğunluk düzeyi ve içeriksel çeşitliliği ilk olarak incelenmiştir.

LDA modeli ile oluşturulan genel kelime bulutu görselleştirmesi (bkz. Şekil 10), BB literatüründe en sık karşılaşılan anahtar kavramların yoğunluk ve görünürlük düzeylerini özetlemektedir. Analiz edilen kelime bulutunda “resource”, “information”, “system”, “server” ve “technology” gibi terimlerin öne çıktığı görülmektedir. Bu kavramlar, BB literatürünün temel yapı taşlarını temsil etmekte ve altyapı, veri yönetimi ve bilgi sistemleri gibi çekirdek temalarla doğrudan ilişkilendirilmektedir.

Kelime bulutunda ayrıca “algorithm”, “performance” ve “security” gibi teknik odaklı terimlerin de belirgin bir yer tuttuğu gözlemlenmektedir. Bu durum, performans optimizasyonu, algoritmik geliştirmeler ve güvenlik önlemlerinin BB araştırmalarında ne denli kritik öneme sahip olduğuna işaret etmektedir. Öte yandan, daha küçük yazı tipiyle temsil edilen “blockchain”, “query”, “latency” ve “monitoring” gibi kavramlar ise, literatürde henüz geniş çapta işlenmemiş ancak niş ve potansiyel taşıyan araştırma alanlarını göstermektedir.

Kelime bulutunda gözlenen kavramsal çeşitlilik, BB alanının yalnızca birkaç ana başlığa sıkışmadığını, aksine çok katmanlı ve disiplinler arası bir yapıya sahip olduğunu açıkça ortaya koymaktadır. Bazı kavramların birden fazla tematik alanda yer alması (örneğin, “data” ve “information” gibi) temalar arası örtüşmenin yüksek olduğunu ve BB araştırmalarının dinamik bir yapıya sahip olduğunu göstermektedir. Bu bağlamda, kelime bulutu, LDA sonuçlarının daha geniş kavramsal bir perspektifle okunmasına olanak tanımakta ve sonraki tematik analizlere sağlam bir temel oluşturmaktadır.

Kelime bulutunda öne çıkan kavramlar genel eğilimlere dair ön bilgi sunarken, tematik frekans analizi ise bu kavramların hangi konu başlıkları altında yoğunlaştığını ve literatürdeki dağılım oranlarını daha sistematik biçimde ortaya koymaktadır. Bu bağlamda, LDA modelinden türetilen 20 ana temanın marjinal frekans değerleri incelenmiş ve BB alanındaki kavramsal yoğunluk haritası oluşturulmuştur.

LDA modeliyle elde edilen 20 ana tema, BB literatüründe kavramsal yoğunluğun hangi alanlarda toplandığını niceliksel olarak ortaya koymaktadır (bkz. Şekil 11). Her tema, içerdiği anahtar kelimelerin metinlerdeki ortak örüntülerine göre adlandırılmış ve frekans değerlerine göre sıralanmıştır.

Grafik incelendiğinde, “Emerging Technologies in Cloud Computing” (%22,12) ve “Distributed Applications and Frameworks” (%16,02) temalarının, tüm literatür korpusunun neredeyse üçte birini kapsadığı görülmektedir. Bu iki tema hem teknolojik yeniliklerin hem de uygulama mimarilerinin BB araştırmalarındaki baskın eksenlerini temsil etmektedir. Özellikle, BB’nin giderek artan yeni nesil teknoloji adaptasyonları ve dağıtık sistem çözümleri ile entegre edilmesi, bu sonuçları desteklemektedir.

Orta frekans grubunda yer alan “Cloud Services and User Experience” (%7,23) ve “Energy Efficiency and Task Scheduling” (%5,97) gibi temalar, BB’nin operasyonel performansı ve kullanıcı deneyimi boyutlarında yoğunlaştığını göstermektedir. Bu alanlar, literatürde sürdürülebilir hizmet sunumu, kaynak tahsisi ve sistem verimliliği gibi önemli uygulamalı problemlere odaklanıldığını işaret etmektedir.

Öte yandan, daha düşük frekanslı temalar arasında “Patient Health Monitoring Systems” (%2,03) ve “Blockchain and Decentralized Systems” (%0,53) gibi başlıklar yer almaktadır. Bu durum, söz konusu konuların henüz geniş bir akademik kapsama ulaşmadığını, ancak özellikle dijital sağlık ve merkeziyetsiz altyapılar gibi yükselen alanlarda stratejik öneme sahip olduklarını düşündürmektedir.

Bu frekans analizi genel olarak değerlendirildiğinde, BB literatürünün büyük ölçüde yeni teknolojik gelişmelere, uygulama mimarilerine ve operasyonel optimizasyona odaklandığı; ancak bunun yanında niş araştırma alanlarının da literatürde

kendine yer bulmaya başladığı görülmektedir. Bu yapı, alandaki araştırma eğilimlerinin hem çeşitlilik hem de derinlik açısından zenginleştiğini göstermektedir.

LDA modeliyle belirlenen tematik yapılar üzerine yapılan frekans analizi sonrasında, temalar arası kavramsal ilişkileri ve yoğunlukları daha derinlemesine değerlendirmek amacıyla Şekil 12'de sunulan anahtar kelime kümelenme haritası incelenmiştir. Şekil 12'te, BB literatüründeki 20 ana tema, anahtar kavramlar arasındaki içeriksel ilişkilere göre çok boyutlu bir ağ yapısında görselleştirilmiştir.

Grafikte her bir renkli küme, belirli bir tematik başlığı temsil etmekte ve düğümler arasındaki bağlantılar ortak anahtar kelime kullanımlarına işaret etmektedir. Merkezi konumda bulunan temalar (örneğin, Emerging Technologies in Cloud Computing ve Distributed Applications and Frameworks) daha yüksek kavramsal yoğunluğa sahipken; çevrede yer alan Blockchain and Decentralized Systems veya Patient Health Monitoring Systems gibi temalar, literatürde daha özgül ve niş alanları temsil etmektedir. Ağ grafiğinde “cloud”, “system”, “resource”, “service” gibi yüksek bağlantılı terimlerin birçok tema arasında köprü işlevi gördüğü gözlemlenmektedir. Bu durum, BB araştırmalarında farklı temalar arasında güçlü kavramsal geçişkenlikler bulunduğunu göstermektedir. Özellikle merkezde yer alan temalar etrafında kümelenen diğer başlıklar, literatürdeki bilgi üretiminin merkezî ve çevresel yapısını açıkça ortaya koymaktadır.

Bu bulgular, BB literatürünün yalnızca belirli başlıklarda yoğunlaşmadığını, aynı zamanda farklı uygulama alanlarına yayılan zengin ve çok boyutlu bir kavramsal ağ yapısına sahip olduğunu göstermektedir.

LDA modeliyle belirlenen temalar üzerine yapılan analizler, BB literatürünün zaman içerisindeki kavramsal evrimini ve tematik değişim dinamiklerini ortaya koymaktadır. Bu analizde, temaların yıllara göre dağılımları ve genel trend eğilimleri hem grafiksel hem de sayısal olarak sunulmuş ve değerlendirilmiştir.

Şekil 13'te, 2014–2024 yılları arasındaki yıllık eğilim değerlerini gösteren çubuk grafik yer almaktadır. Tablo 14 ise her bir temanın yıllık göreceli frekanslarını ve trend skorlarını detaylı bir şekilde sunmaktadır.

Trend analizine göre bazı temalar son on yılda belirgin bir artış göstermiştir. Özellikle “Network Traffic Control and Communication” başlığı +0.006076 değer ile en güçlü yükseliş eğilimi sergilemiştir. Bu başlığı sırasıyla “Distributed Applications and Frameworks” (+0.001735) ve “IoT and Smart Systems” (+0.001637) konuları takip etmektedir. Bu temaların zaman içinde artan ağırlıkları, BB altyapısının giderek daha

çok ağ yönetimi, IoT entegrasyonu ve dağıtık sistemler üzerine yoğunlaştığını göstermektedir.

Buna karşın, bazı temalar literatürde azalan bir önem göstermektedir. “Energy Efficiency and Task Scheduling” teması, -0.009947 değer ile en büyük düşüşü sergilemiş olup, 2014 yılında %34,25 seviyesinde olan payı, 2024 yılında %25,61’e gerilemiştir. Benzer şekilde, “Virtual Machine and Workload Management” (-0.002952) ve “Privacy and Access Control” (-0.001037) gibi temalar da düşüş eğilimindedir. Bu bulgu, özellikle kaynak yönetimi ve sanallaştırma gibi konularda doaygunluğa ulaşıldığını düşündürmektedir.

Bazı temalar ise zaman içinde daha durağan bir seyir izlemiştir. “Cloud Services and User Experience” (+0.000484) ve “Optimization and Problem Solving” (+0.000717) gibi konular, literatürde istikrarlı biçimde varlığını sürdürmüş, ancak ciddi bir yükseliş ya da düşüş göstermemiştir. Bu durum, kullanıcı deneyimi ve optimizasyon gibi konuların sürekli önem taşıdığını ve alanın temel yapı taşları arasında yer aldığını göstermektedir.

Tablo 14’te sunulan sayısal veriler, temaların yıllık görelı frekans değışimlerini sistematik bir biçimde takip etmeye olanak sağlamış, Şekil 13 ise bu değışimleri daha görsel bir şekilde kavramamıza yardımcı olmuştur. Örneğin, enerji verimliliği alanındaki düşüşe karşın, ağ trafiği kontrolü ve IoT sistemlerine olan ilginin yükselmesi, bulut teknolojilerinin uygulama alanlarında yaşanan dönüşümü açıkça yansıtmaktadır.

Bu analizler, BB alanında hangi araştırma eksenlerinin ön plana çıktığını, hangilerinin ise zamanla görelı önemini kaybettiğini anlamak açısından kritik öneme sahiptir.

Diğer bir analiz ise, LDA modelinden elde edilen tematik yapıları görsel ve hiyerarşik olarak gruplamak amacıyla gerçekleştirilmiştir. Bu doğrultuda BB alanındaki araştırmaların kavramsal yoğunluklarını ve tematik ilişkilerini ortaya koymak amacıyla bir taksonomik yapı geliştirilmiştir. Söz konusu yapı, alt temaların üst temalarla olan ilişkilerini yansıtacak şekilde sınıflandırılmış ve sunburst (güneş ışını) diyagramı ile görselleştirilmiştir. Şekil 14’te yer alan bu diyagram, temaların dağılımını oranlara dayalı olarak segmentlere ayırmakta ve BB alanındaki araştırma eğilimlerini hiyerarşik düzlemde anlamaya yönelik bütüncül bir bakış açısı sunmaktadır.

Tematik Dağılım Bulguları:

- **Technology (Teknoloji)** kategorisi, literatürde en geniş tematik kapsama sahip olup toplamın yaklaşık %34,4’ünü temsil etmektedir. Bu başlık altında özellikle

Cyber-Physical Systems, Software Systems, Simulation, Cybersecurity ve *Cloud Services* gibi teknoloji tabanlı temalar yoğunluk göstermektedir.

- **Management (Yönetim)** başlığı, toplam tematik yapının yaklaşık **%24,1**'ini oluşturmaktadır. Bu kategori, *Innovation Capability, Business Process Management, Education and Training* ve *Supply Chain Management* gibi organizasyonel ve yönetsel alt başlıklarla temsil edilmektedir.
- **Process (Süreç)** kategorisi, literatürün **%22,5**'lik bölümünü kapsamaktadır. Özellikle *Deep Learning, Decision-Making, Scheduling Optimization* ve *Performance Effect* gibi süreç iyileştirmeye odaklı teknik temalar ön plana çıkmaktadır.
- **Sustainability (Sürdürülebilirlik)** başlığı, toplam içeriğin yaklaşık **%9,9**'unu oluşturmaktadır. *Sustainable Development, Energy Efficiency* ve *Circular Economy* gibi çevresel ve verimlilik odaklı temalar bu gruba dâhildir.
- **Manufacturing (Üretim)** kategorisi ise **%9,1**'lik oranla daha sınırlı bir alanı temsil etmektedir. Özellikle *Smart Manufacturing, Additive Manufacturing* ve *Agri-Food Production* gibi üretim süreçlerine ilişkin temalar bu başlık altında yer almaktadır.

Gözlemlenen İstatistiksel Yoğunluklar:

Sunburst görselleştirmesi, temaların literatürdeki ağırlığını oransal segmentler aracılığıyla temsil etmektedir. En büyük segmentlerden biri olan *Cyber-Physical Systems* ve *Software Systems* temaları, BB alanındaki teknolojik araştırmaların öncül konularını oluşturmaktadır. Diğer yandan, *Agri-Food Production* ve *Additive Manufacturing* gibi alt temalar, daha sınırlı oranda çalışılmış ve literatürde daha küçük yer kaplamıştır.

Bu değerlendirme, BB alanında yürütülen araştırmaların yalnızca teknik bileşenlerle sınırlı kalmadığını; aynı zamanda süreç yönetimi, sürdürülebilirlik ve yönetsel boyutlara da yayıldığını ortaya koymaktadır. Böylece, hem tematik çeşitliliğin hem de araştırma eğilimlerinin yöneldiği eksenlerin çok boyutlu bir çerçevede analiz edilmesi mümkün hâle gelmiştir.

Değerlendirme

Bu çalışmada elde edilen bulgular, BB literatüründe son on yılda yaşanan tematik dönüşümleri kapsamlı biçimde ortaya koymuştur. Bibliyometrik ve içerik temelli analizler ışığında, BB araştırmalarının yalnızca teknolojik altyapılarla sınırlı kalmayıp, süreç yönetimi, sürdürülebilirlik ve disiplinler arası uygulamalar gibi alanlara da yayıldığı görülmüştür. Özellikle LDA yönteminin sağladığı kavramsal kümelenme

sonuçları, literatürdeki ana odak alanlarını ve henüz yeterince çalışılmamış potansiyel temaları net bir biçimde belirleme olanağı sunmuştur.

Bu bağlamda, çalışmada uygulanan LDA tabanlı analizler, yalnızca metinsel frekansları değil, aynı zamanda temalar arası kavramsal örtüşmeleri, araştırma alanlarının evrimsel eğilimlerini ve literatürdeki yapısal boşlukları da görünür kılabilmiştir. Özellikle kelime bulutu ve kümelenme analizlerinde ortaya çıkan içerik çeşitliliği, BB alanının disiplinler arası karakterini ve sürekli genişleyen araştırma spektrumunu doğrulamaktadır. Anahtar kavramlar arasında belirgin yoğunlaşmalar ("service", "resource", "security" vb.) gözlemlenirken; bazı stratejik kavramların ("learn", "model", "blockchain" vb.) daha düşük frekanslarla temsil edilmesi, bu alanlarda hâlen gelişime açık boşluklar bulunduğunu düşündürmektedir.

Frekans analizleri sonucunda, “Emerging Technologies in Cloud Computing” ve “Distributed Applications and Frameworks” gibi temaların literatürde baskın konumda olduğu tespit edilmiştir. Buna karşılık, "Blockchain and Decentralized Systems" veya "Patient Health Monitoring Systems" gibi alanların daha düşük oranlarda temsil edilmesi, bu temaların potansiyel gelişim alanları olarak değerlendirilebileceğine işaret etmektedir. Ayrıca yıllık eğilim analizleri, enerji verimliliği gibi klasik konuların düşüş eğilimi gösterirken, IoT ve ağ iletişimi gibi alanların yükselen bir dinamizm sergilediğini ortaya koymuştur.

Özellikle BB alanında içerik temelli konu modelleme yöntemleriyle gerçekleştirilen çalışmaların sınırlı sayıda olması, bu çalışmanın literatüre özgün bir katkı sunmasını sağlamıştır. Bu bağlamda gerçekleştirilen LDA tabanlı analiz, BB araştırmalarında kavramsal kümelenmelerin ve tematik eğilimlerin sistematik biçimde haritalandırılmasına yönelik önemli bir adım niteliğindedir. Literatürdeki mevcut tematik yoğunluklar, evrimsel eğilimler ve araştırma boşlukları bu çalışma sayesinde daha görünür hâle getirilmiştir.

Bununla birlikte, çalışmanın bazı metodolojik sınırlılıkları da bulunmaktadır. İlk olarak, analiz yalnızca Scopus veri tabanındaki İngilizce yayınlarla sınırlı kalmış, diğer veri tabanları ve farklı dillerde yayımlanan çalışmalar kapsam dışında bırakılmıştır. İkinci olarak, kullanılan LDA modeli, kavramlar arasındaki ilişkileri tematik düzeyde başarılı bir şekilde ortaya koysa da bağlamsal derinlik açısından sınırlı kalabilmektedir. LDA, kelimeler arasındaki doğrudan ilişkiyi dikkate almadığı için, bağlamsal farklılıkları tam anlamıyla yansıtamamaktadır. Son olarak, tematik değerlendirmeler yalnızca anahtar kelime frekanslarına dayalı olarak yapılmış olup, metodolojik

yönelimler veya sektörel uygulama çeşitlilikleri gibi içeriksel ayrımlar çalışma kapsamına alınmamıştır.

Bu doğrultuda, ilerleyen çalışmalarda farklı veri tabanlarının entegre edilmesi, çok dilli literatürlerin analize katılması ve LDA dışında Contextualized Topic Models (CTM) veya BERTopic gibi bağlam duyarlı modelleme yöntemlerinin kullanılması önerilmektedir. Ayrıca disiplinler arası bağlantıların ve sektörel uygulamaların daha ayrıntılı analiz edilmesi, BB alanındaki kavramsal evrimin daha yüksek çözünürlükle anlaşılmasına katkı sağlayacaktır.

Literatüre Katkı

Bu çalışma, BB literatüründe daha önce genellikle atıf sayıları, yayın hacmi veya işbirlikçi ağlar gibi yüzeysel ölçütlerle yapılan analizlerin ötesine geçerek içerik temelli bir yaklaşım sunmuştur. LDA modeli kullanılarak gerçekleştirilen konu modelleme süreci, BB alanındaki araştırma kümelerinin yalnızca niceliksel değil, niteliksel boyutlarını da görünür kılmıştır.

Özellikle öne çıkan temalar, kavramsal yoğunluk bölgeleri ve zaman içerisindeki eğilimler, araştırmacılara BB literatürünün dinamik yapısını çok daha sistematik bir çerçevede değerlendirme imkânı sağlamaktadır.

Bu yönüyle tez çalışması, BB alanındaki akademik gelişimi haritalandırarak; gelecekte yapılacak araştırmalar için hem teorik hem de uygulamalı bir zemin oluşturmuştur.

Çalışmanın Sınırlılıkları

Bununla birlikte, bu çalışmanın bazı metodolojik sınırlılıkları da bulunmaktadır. İlk olarak, analiz yalnızca Scopus veri tabanındaki İngilizce yayınlarla sınırlandırılmıştır. Bu durum, farklı veri tabanlarında veya farklı dillerde yayımlanmış çalışmaları kapsam dışı bırakmış olabilir.

İkinci olarak, kullanılan LDA modeli, kavramlar arasındaki ilişkileri tematik düzeyde başarılı şekilde ortaya koysa da; bağlamsal veya semantik derinlik açısından sınırlı kalmaktadır. LDA, kelimeler arasındaki doğrudan ilişkiyi dikkate almaz; bu nedenle bağlamsal farklılıkları yakalamada zayıf kalabilir.

Son olarak, temalar yalnızca anahtar kelime frekanslarına dayalı olarak yorumlanmıştır. Yazar niyetleri, metodolojik çeşitlilikler veya sektörel uygulama farkları gibi detaylı içerik analizleri çalışmanın kapsamı dışında tutulmuştur.

Gelecek Araştırmalar İçin Öneriler

İlerleyen çalışmalarda, birden fazla veri tabanının entegre edilmesi ve çok dilli literatürlerin analize katılması önerilmektedir. Böylece, BB literatürünün küresel kapsamı daha doğru ve kapsayıcı şekilde yansıtılabilir.

Ayrıca LDA yerine veya LDA ile birlikte Contextualized Topic Modeling (CTM), BERTopic gibi daha gelişmiş bağlamsal modelleme yöntemleri kullanılabilir. Bu tür yöntemler, kelimelerin kullanıldığı bağlamı da dikkate alarak daha zengin temalar elde edilmesini sağlayabilir.

Son olarak, tematik analizlerin yanı sıra metinlerin metodolojik yönelimlerinin (nicel, nitel, karma yöntem vb.) veya sektörel uygulama alanlarının da incelendiği derinlemesine çalışmalar yapılması, BB literatürünün yapısal evrimi hakkında daha detaylı çıkarımlar yapılmasına olanak tanıyacaktır.

Bu çalışma, BB literatüründeki tematik yapıları ve eğilimleri daha önce sınırlı biçimde incelenmiş alanlara özgün bir katkı sağlamayı amaçlamıştır. LDA yöntemi kullanılarak gerçekleştirilen konu modelleme süreci, klasik bibliyometrik analizlerin ötesine geçerek, kavramsal kümelenmeleri, temalar arası ilişkileri ve zamansal değişimleri daha detaylı bir biçimde ortaya koymuştur. Elde edilen sonuçlar, BB araştırmalarının büyük ölçüde yeni teknolojilere, dağıtık sistem mimarilerine ve operasyonel optimizasyona odaklandığını, ancak sağlık sistemleri, blokzincir ve sürdürülebilirlik gibi stratejik niş alanların da yükselmekte olduğunu göstermektedir.

Literatürde önceki çalışmalarda tematik yoğunlukların ve örtüşmelerin bu denli detaylı analizi nadir görülürken, bu çalışmada LDA'nın sağladığı içeriksel ayrıştırma ile alandaki kavramsal çeşitliliğin daha net haritalandırılması mümkün olmuştur. Bununla birlikte, elde edilen bulgular BB alanında özellikle disiplinler arası entegrasyon, güvenlik uygulamaları ve veri analitiği gibi konularda henüz yeterince işlenmemiş önemli boşluklar bulunduğunu da ortaya koymaktadır.

Bu doğrultuda, gelecekte yapılacak çalışmalarda LDA ve benzeri içerik madenciliği tekniklerinin daha geniş veri kümeleriyle uygulanması ve zamansal-dinamik analizlerin derinleştirilmesi, BB'nin alanının evrimini daha yüksek çözünürlükle anlamaya katkı sağlayacaktır.

KAYNAKÇA

- Aggarwal, C. C., ve Zhai, C. (2012). Mining text data. Springer
<https://doi.org/10.1007/978-1-4614-3223-4>
- Albalawi, R., Yeap, T. H., ve Benyoucef, M. (2020). *Using topic modeling methods for short-text data: A comparative analysis. Frontiers in Artificial Intelligence*, 3, 42.
<https://doi.org/10.3389/frai.2020.00042>
- Alemayehu, E., ve Fang, Y. (2024). Supervised probabilistic latent semantic analysis with applications to controversy analysis of legislative bills. *Intelligent Data Analysis*, 28(1), 161–183. <https://doi.org/10.3233/ida-227202>
- Alfallah, F., ve Veri, J. (2024). The impact of cloud computing on the growth of technology startups. *International Journal of Education, Social Studies, and Management (IJESSM)*, 4(3), 1219–1229.
<https://doi.org/10.52121/ijessm.v4i3.523>
- Alghamdi, R., ve Alfalqi, K. (2015). A survey of topic modeling in text mining. *International Journal of Advanced Computer Science and Applications*, 6(1), 147–153. <https://doi.org/10.14569/IJACSA.2015.060121>
- Al-Hajri, S., Echchabi, A., Ayedh, A. M., ve Omar, M. M. S. (2021). The cloud computing systems' adoption in the higher education sector in Oman in light of the COVID-19 pandemic. *International Journal of Higher Education*, 10(5), 122–134. <https://doi.org/10.11591/ijere.v10i3.21671>
- Alhashmi, S. M., Hashem, I. A. T., ve Al-Qudah, I. (2024). *Artificial intelligence applications in healthcare: A bibliometric and topic model-based analysis. Intelligent Systems with Applications*, 21, 200299.
<https://doi.org/10.1016/j.iswa.2023.200299>
- Ali, O., Shrestha, A., Soar, J. ve Wamba, S. F. (2018). *Cloud computing-enabled healthcare opportunities, issues, and applications: A systematic review. International Journal of Information Management*, 43, 146–158.
<https://doi.org/10.1016/j.ijinfomgt.2018.07.001>
- Almaiah, M. A., Al-Khasawneh, A., ve Althunibat, A. (2020). *Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic. Education and Information Technologies*, 25(5), 5261–5280.
<https://doi.org/10.1007/s10639-020-10219-y>

- Almorsy, M., Grundy, J., ve Müller, I. (2016). *An analysis of the cloud computing security problem*. arXiv preprint arXiv:1609.01107. <https://arxiv.org/abs/1609.01107>
- Alouffi, B., Ameen, A., Rawashdeh, M., ve Baker, T. (2021). A systematic literature review on cloud computing security threats and mitigation strategies. *Journal of Cloud Computing*, 10(1), 1–30. <https://doi.org/10.1186/s13677-021-00247-5>
- Antonopoulos, N., ve Gillam, L. (2010). *Cloud computing: Principles, systems and applications* (Vol. 51, No. 7). Springer. <https://doi.org/10.1007/978-1-84996-241-4>
- Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., ... ve Zaharia, M. (2010). A view of cloud computing. *Communications of the ACM*, 53(4), 50-58. <https://doi.org/10.1145/1721654.1721672>
- Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., ... ve Martinis, J. M. (2019). Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779), 505–510. <https://doi.org/10.1038/s41586-019-1666-5>
- Ayaz, A., Celik, K., ve Ozyurt, O. (2021). Pattern detection in cloud computing: Bibliometric mapping of publications in the field from past to present. *COLLNET Journal of Scientometrics and Information Management*, 15(2), 469–494. <https://doi.org/10.1080/09737766.2021.2007038>
- Aydın, N. (2022). *Bulut bilişim e-ticaretin geleceğini nasıl etkileyecek?* *Journal of Social Humanities and Administrative Sciences (JOSHAS)*, 8(46), 613–623. <https://doi.org/10.29228/JOSHAS.67206>
- Baas, J., Schotten, M., Plume, A., Côté, G., ve Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377–386. https://doi.org/10.1162/qss_a_00019
- Badhe, M. G., ve Arjunwadkar, M. (2023). Decentralised storage systems for blockchain powered applications. *International Journal of Emerging Technologies in Multidisciplinary Sciences*, 7(6). <https://doi.org/10.46647/ijetms.2023.v07i06.046>
- Bai, P., Kumar, S., Kumar, K., Kaiwartya, O., ve Mahmud, M. (2022). GDPR compliant data storage and sharing in smart healthcare system: a blockchain-based solution. *Electronics*, 11(20), 3311. <https://doi.org/10.3390/electronics11203311>

- Bajaj, D., Bharti, U., Goel, A., ve Gupta, S. C. (2020). PaaS providers and their offerings. *International Journal of Scientific ve Technology Research*, 9(2), 4009–4015. <https://www.researchgate.net/publication/346788471>
- Batool, A., ve Byun, Y.-C. (2024). Enhanced sentiment analysis and topic modeling during the pandemic using automated Latent Dirichlet Allocation. *IEEE Access*, 12, 81206–81220. <https://doi.org/10.1109/ACCESS.2024.3411717>
- Beloglazov, A., ve Buyya, R. (2010, May). Energy efficient resource management in virtualized cloud data centers. In *2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing* (pp. 826–831). IEEE. <https://doi.org/10.1109/CCGRID.2010.46>
- Bernstein, D. J., ve Lange, T. (2017). *Post-quantum cryptography—Dealing with the fallout of physics success*. Cryptology ePrint Archive. <https://doi.org/10.1038/nature23461>
- Bhagwat, S., ve Pandit, D. (2024). SpaCy and NLTK NLP techniques for text summarization: A comprehensive comparison. In R. Dumbre, P. Ankalwar, S. Bhagwat, ve D. Pandit (Eds.), *Advancements in AI, IoT and Big Data* (pp. XX–XX). Springer. https://doi.org/10.1007/978-981-97-8160-7_5
- Bharati, V. K. (2019, April). *Mapping of scholarly publication on cloud computing during 2010–2019: A bibliometric study based on Web of Science*. In K. P. Singh, R. K. Sharma, R. K. Bhatt, M. Meera, ve J. N. Singh (Eds.), *Proceedings of the 3rd DLA-SRFLIS Summit 2019: Digital Age Strategies in Information Management for Sustainable Librarianship* (pp. 441–445). Delhi Library Association ve Satija Research Foundation. <https://www.researchgate.net/publication/338501540>
- Bharati, V. K., ve Singh, M. P. (2019). *Rejuvenating libraries from the cloud: A bibliometric analysis of cloud computing*. *Library Philosophy and Practice* (e-journal), 2552. University of Nebraska - Lincoln. <https://digitalcommons.unl.edu/libphilprac/2552/>
- Bianzino, A. P., Chaudet, C., Rossi, D., ve Rougier, J. L. (2012). A survey of green networking research. *IEEE Communications Surveys ve Tutorials*, 14(1), 3–20. <https://doi.org/10.1109/SURV.2011.113010.00106>
- Bilur, S. S., ve Salunkhe, M. S. (2024). Multi-tenant yet customizable cloud native SaaS web application leveraging AIML: Architecture ve strategies. *International Journal of Advances in Engineering and Management (IJAEM)*, 6(11), 118–127. <https://doi.org/10.35629/5252-0611118127>

- Blei, D. M., Ng, A. Y., ve Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
<http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Blei, D. M., ve Lafferty, J. D. (2006). *Correlated topic models* [Technical report]. Carnegie Mellon University.
<https://www.cs.cmu.edu/afs/cs/usr/lafferty/www/pub/ctm.pdf>
- Blei, D. M., ve Lafferty, J. D. (2007). A correlated topic model of Science. *Annals of Applied Statistics*, 1(1), 17–35. <https://doi.org/10.1214/07-AOAS114>
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
- Borra, P. (2024). *Comprehensive survey of Amazon Web Services (AWS): Techniques, tools, and best practices for cloud solutions*. *International Research Journal of Advanced Engineering and Science*, 9(3), 24–29.
<https://www.researchgate.net/publication/381956331>
- Breuninger, T. A., Wawro, N., Breuninger, J., Reitmeier, S., Clavel, T., Six-Merker, J., ve Linseisen, J. (2021). Associations between habitual diet, metabolic disease, and the gut microbiota using latent Dirichlet allocation. *Microbiome*, 9, Article 61. <https://doi.org/10.1186/s40168-020-00969-9>
- Butt, U. A., Mehmood, M., Shah, S. B. H., Amin, R., Shaukat, M. W., Raza, S. M., ... ve Piran, M. J. (2020). A review of machine learning algorithms for cloud computing security. *Electronics*, 9(9), 1379. <https://doi.org/10.3390/electronics9091379>
- Buyya, R., Pandey, S., ve Vecchiola, C. (2013). Market-oriented cloud computing and the Cloudbus toolkit. In *Large Scale Network-Centric Distributed Systems* (pp. 319–358). Wiley. <https://doi.org/10.1002/9781118640708.ch14>
- Buyya, R., Ranjan, R., ve Calheiros, R. N. (2010). InterCloud: Utility-oriented federation of cloud computing environments for scaling of application services. In C. H. Hsu, L. T. Yang, J. H. Park, ve S. S. Yeo (Eds.), *Algorithms and architectures for parallel processing (ICA3PP 2010)* (Lecture Notes in Computer Science, Vol. 6081, pp. 13–31). Springer. https://doi.org/10.1007/978-3-642-13119-6_2
- Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J. ve Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems*, 25(6), 599–616. <https://doi.org/10.1016/j.future.2008.12.001>

- Chen, L. K., Zhang, Y., Chen, L., Liu, S., Li, Z. D., Fan, J. Y., ... ve Pan, J. W. (2016). *Quantum secure cloud systems: Integrating quantum random number generators and homomorphic encryption*. *Science China Physics, Mechanics ve Astronomy*, 59(6), 1–7. <https://doi.org/10.1007/s11433-016-0037-0>
- Chen, P., Lin, S., ve Chu, Y. (2011). Using Google latent semantic distance to extract the most relevant information. *Expert Systems With Applications*, 38(6), 7349–7358. <https://doi.org/10.1016/j.eswa.2010.12.092>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., ve Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field. *Journal of Informetrics*, 5(1), 146–166. <https://doi.org/10.1016/j.joi.2010.10.002>
- Çam, H. (2012). *Türkiye'deki üniversitelerde bulut bilişim teknolojisinin uygulanabilirliğinin Teknoloji Kabul Modeli yaklaşımıyla belirlenmesi* (Yayımlanmış doktora tezi). Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü. <https://tez.yok.gov.tr/>
- Çullu, B. ve Arabelen, G. (2023). Pazarlamada konu modellemesi: Literatür taraması ve bilimetrik analiz. *Güncel Pazarlama Yaklaşımları ve Araştırmaları Dergisi*, 4(1), 58–89. <https://doi.org/10.54439/gupayad.1316544>
- De la Hoz-M., J., Fernández-Gómez, M. J., ve Mendes, S. (2021). *LDAShiny: An R package for exploratory review of scientific literature based on a Bayesian probabilistic model and machine learning tools*. *Mathematics*, 9(14), 1671. <https://doi.org/10.3390/math9141671>
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., ve Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41(6), 391–407. [https://doi.org/10.1002/\(SICI\)1097-4571\(199009\)41:6<391::AID-ASI1>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9)
- Devarajan, M. V. (2020). Improving security control in cloud computing for healthcare environments. *Journal of Science and Technology (JST)*, 5(6), 178–189. <https://doi.org/10.46243/jst.2020.v5.i06.pp178-189>
- Dissanayake, N., Jayatilaka, A., Zahedi, M., ve Babar, M. A. (2022). *Software security patch management—A systematic literature review of challenges, approaches, tools and practices*. *Information and Software Technology*, 144, 106771. <https://doi.org/10.1016/j.infsof.2021.106771>

- Dokuz, A. Ş., ve Çelik, M. (2017). Bulut bilişim sistemlerinde verinin farklı boyutları üzerine derleme. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 6(2), 316–338. <https://doi.org/10.28948/ngumuh.338074>
- Dokuz, İ. ve Çelik, A. (2017). *Bulut bilişim altyapılarında performans yönetimi*. Bilgi Teknolojileri ve Uygulamaları Dergisi, 5(2), 14-28. <https://doi.org/10.28948/ngumuh.338074>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., ve Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Du, B. X., ve Liu, G. Y. (2021). *Topic analysis in LDA based on keywords selection*. *Journal of Computers*, 32(4), 1–12. <https://doi.org/10.53106/199115992021083204001>
- Dumais, S. T. (2004). Latent Semantic Analysis. *Annual Review of Information Science and Technology*, 38(1), 188–230. <https://doi.org/10.1002/aris.1440380105>
- Dwivedi, S. K., Yadav, J., Ansar, S. A., Khan, M., ve Pandey, D. (2023, November). *An investigation with emerging issues and preventive measures: Cloud computing perspective*. In *AIP Conference Proceedings* (Vol. 2821, No. 1). AIP Publishing. <https://doi.org/10.1063/5.0159071>
- El-Sayed, H., Sankar, S., Prasad, M., Puthal, D., Gupta, A., Mohanty, M., ve Lin, C.-T. (2018). Edge of things: The big picture on the integration of edge, IoT and the cloud in a distributed computing environment. *IEEE Access*, 6, 1706–1717. <https://doi.org/10.1109/ACCESS.2017.2780087>
- Falah, M. F., Panduman, Y. Y. F. ve Sukaridhoto, S. (2021). *Comparison of cloud computing providers for development of big data and internet of things application*. *International Conference on Information and Computer Technologies*, 22(3), 1723-1730. <https://doi.org/10.11591/ijeecs.v22.i3.pp1723-1730>
- Feldman, R., ve Sanger, J. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511546914>
- Fernandes, D. A., Soares, L. F., Gomes, J. V., Freire, M. M., ve Inácio, P. R. (2014). A quick perspective on the current state in cybersecurity. In B. Akhgar ve H. R. Arabnia (Eds.), *Emerging trends in ICT security* (pp. 423–442). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-411474-6.00025-6>

- Fernando, C. (2022). *Building enterprise software systems with microservice architecture*. In *Solution Architecture Patterns for Enterprise* (ss. 53–78). Springer. https://doi.org/10.1007/978-1-4842-8948-8_3
- Gandhi, H., ve Jain, D. A. (2025). Cloud cost optimization strategies using machine learning algorithms. *International Journal of Innovative Science and Research Technology (IJISRT)*, 9(11), 3573–3593. <https://doi.org/10.5281/zenodo.14831440>.
- Gandhi, V. A., ve Kumbharana, C. K. (2022). Comparative study of Amazon EC2 and Microsoft Azure cloud architecture. *International Journal of Advanced Networking Applications*, 13(6), 117–123. <https://www.ijana.in>
- Garg, D., Sidhu, J., ve Rani, S. (2019). Emerging trends in cloud computing security: A bibliometric analyses. *IET Software*, 13(3), 223–231. <https://doi.org/10.1049/iet-sen.2018.5222>
- Garg, M., ve Kanjilal, U. (2022). An exploratory analysis of online discussion of library and information science professionals in India using text mining. *Journal of Information Science Theory and Practice*, 10(3), 40–56. <https://doi.org/10.1633/JISTAP.2022.10.3.4>
- Genç, A., Yurtseven, A., Özyurt, H., ve Özyurt, Ö. (2024). STACKOVERFLOW’da “BIG DATA” ile ilgili gönderilerin konu modelleme ve birliktelik analizi ile özelliklerinin çıkarılması. *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi*, 32(1), 1257–1268. <https://doi.org/10.31796/ogummf.1375611>
- George, L. E., ve Birla, L. (2018, June). *A study of topic modeling methods*. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 109–113). IEEE. <https://doi.org/10.1109/ICCONS.2018.8663152>
- Gerli, S., Ascari, R., Migliorati, S., Cigna, T., ve Borrotti, M. (2024). Beyond human labelling: An automatic topic identification framework for big web data. *Electronic Journal of Applied Statistical Analysis*, 17(3), 545–571. <https://doi.org/10.1285/i20705948v17n3p545>
- Goyal, S. (2014). Public vs private vs hybrid vs community-cloud computing: A critical review. *International Journal of Computer Network and Information Security*, 6(3), 20–29. <https://doi.org/10.5815/ijcnis.2014.03.03>
- Göv, S. A. ve Erdoğan, D. (2020). Dördüncü Endüstri Devriminin (Endüstri 4.0) neresindeyiz? *İstanbul Gelişim Üniversitesi Sosyal Bilimler Dergisi*, 7(2), 299–318. <https://doi.org/10.17336/igusbd.661236>

- Griffiths, T. L., ve Steyvers, M. (2004). *Finding scientific topics*. Proceedings of the National Academy of Sciences, 101(Suppl. 1), 5228–5235.
<https://doi.org/10.1073/pnas.0307752101>
- Gupta, B. M., ve Gupta, R. (2015). *Indian cloud computing research: A scientometric assessment of publications output during 2004–13*. SRELS Journal of Information Management, 52(5), 315–326. <https://doi.org/10.17821/srels/2015/v52i5/79733>
- Gupta, D. N., ve Kumar, R. (2021). Networking in IoT: Technologies usage, security threats, and possible countermeasures. *International Journal of Sensors, Wireless Communications and Control*, 11(6), 619–626.
<https://doi.org/10.2174/2210327910666210106090826>
- Gupta, R. K., Agarwalla, R., Naik, B. H., Evuri, J. R., Thapa, A., ve Singh, T. D. (2022). Prediction of research trends using LDA based topic modeling. *Global Transitions Proceedings*, 3, 298–304. <https://doi.org/10.1016/j.gltp.2022.03.015>
- HaCohen-Kerner, Y., Miller, Y. ve Yigal, M. (2020). The bag-of-words approach for text classification: A critical review. *Expert Systems with Applications*, 161, 113738. <https://doi.org/10.1016/j.eswa.2020.113738>
- Hafeez, U. U., Karve, A., Dumba, B., Gandhi, A., ve Zeng, S. (2019, Ekim). *Towards automated patch management in a hybrid cloud*. In International Conference on Service-Oriented Computing (pp. 345–350). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-33702-5_26
- Hagerbaumer, C., et al. (2024). *OpenAQ: Harmonizing billions of air quality measurements into an open and FAIR database*. EGU General Assembly.
<https://doi.org/10.5194/egusphere-egu24-6810>
- Hamutoğlu, N. B. (2018). *Bulut bilişim teknolojileri kabul modeli 3: Ölçek uyarlama çalışması*. Sakarya University Journal of Education, 8(2), 8-25.
<https://doi.org/10.19126/suje.297586>
- Han, Y. J., ve Kim, G. H. (2020). A comparative study on decentralized storage platforms for self-sovereign data. *Asia-Pacific Journal of Convergent Research Interchange*, 6(5), 1–10. <https://doi.org/10.21742/apjcri.2020.05.01>
- Haq, A. U., Li, J., ve Hassan, R. (2019). *A systematic review and meta-analysis of cloud computing and big data in Industry 4.0: Trends, challenges and future research directions*. IEEE Access, 7, 114870–114886.
<https://doi.org/10.1109/ACCESS.2019.2950045>
- Haq, A. U., Zain, A. M., ve Wahab, D. A. (2019). Text mining techniques to capture facts

- for cloud computing adoption and big data processing. *IEEE Access*, 7, 113560–113573. <https://doi.org/10.1109/ACCESS.2019.2934312>
- Hashizume, K., Rosado, D. G., Fernández-Medina, E. ve Fernandez, E. B. (2013). An analysis of security issues for cloud computing. *Journal of Internet Services and Applications*, 4(1), 1–13. <https://doi.org/10.1186/1869-0238-4-5>
- Hassan, S. M. (2021). Determining factors of cloud computing for small and medium enterprises (SMEs). *International Journal of Modern Trends in Engineering and Research*, 8(4), 11–20.
<https://doi.org/10.21884/IJMTER.2021.8018.3Q5VS>
- Heilig, L., ve Voß, S. (2014). A scientometric analysis of cloud computing literature. *IEEE Transactions on Cloud Computing*, 2(3), 266–278.
<https://doi.org/10.1109/TCC.2014.2321168>
- Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42(1), 177–196. <https://doi.org/10.1023/A:1007617005950>
- Hofmann, T. (2017). Probabilistic latent semantic indexing. *ACM SIGIR Forum*, 51(2), 211–218. <https://doi.org/10.1145/3130348.3130370>
- Irshad, M. B. M. (2014). A review on cloud computing adoption: An exploratory analysis. In *Proceedings of the 4th International Symposium, South Eastern University of Sri Lanka (SEUSL)* (pp. 493–500). South Eastern University of Sri Lanka. <http://ir.lib.seu.ac.lk/handle/123456789/1448>
- Jan, R., Wani, W. R., ve Hafiz, O. (2015). Scientometric analysis of cloud computing. *Library Philosophy and Practice (e-journal)*, Article 1273. <http://digitalcommons.unl.edu/libphilprac/1273>
- Janmaijaya, M., Shukla, A. K., Muhuri, P. K., ve Abraham, A. (2021). Industry 4.0: Latent Dirichlet Allocation and clustering based theme identification of bibliography. *Engineering Applications of Artificial Intelligence*, 103, 104280.
<https://doi.org/10.1016/j.engappai.2021.104280>
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., ve Zhao, L. (2019). *Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey*. *Multimedia Tools and Applications*, 78, 15169–15211.
<https://doi.org/10.1007/s11042-018-6894-4>
- Jia, R., Yang, Y., Grundy, J., Keung, J., ve Hao, L. (2021). A systematic review of scheduling approaches on multi-tenancy cloud platforms. *Information and Software Technology*, 132, 106478. <https://doi.org/10.1016/j.infsof.2020.106478>

- Jirasatjanukul, K., Nilsook, P., ve Wannapiroon, P. (2019). Intelligent human resource management using latent semantic analysis with the Internet of Things. *International Journal of Computer Theory and Engineering*, 11(2), 23–26. , <https://doi.org/10.7763/ijcte.2019.v11.1235>
- Joulin, A., Grave, E., Bojanowski, P. ve Mikolov, T. (2017). Bag of tricks for efficient text classification. *European Chapter of the Association for Computational Linguistics Konferansı Bildirileri*, 2, 427–431. <https://doi.org/10.18653/v1/E17-2068>
- Jurafsky, D., ve Martin, J. H. (2020). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed., draft version, 2020). <https://web.stanford.edu/~jurafsky/slp3>
- Karakuş, L. (2021). *Eğitimde metin madenciliği: Türkçe metinlerde sözlük tabanlı duygu analizi* (Yüksek lisans tezi, Akdeniz Üniversitesi, Eğitim Bilimleri Enstitüsü). YÖK Ulusal Tez Merkezi. <http://tez.yok.gov.tr/>
- Kızılkaya, Y. M. (2018). *Duygu analizi ve sosyal medya alanında uygulama* (Doktora tezi, Uludağ Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstatistik Bilim Dalı). YÖK Ulusal Tez Merkezi. <http://tez.yok.gov.tr/>
- Kothapalli, M. (2018). *Cloud computing architectures: Comparing service models (IaaS, PaaS, SaaS) and deployment models (Public, Private, Hybrid, Community) – Uses and trade-offs*. *Journal of Scientific and Engineering Research*, 5(10), 334–341. <https://doi.org/10.5281/zenodo.12798259>
- Kratzke, N. (2018). A brief history of cloud application architectures. *Applied Sciences*, 8(8), 1368. <https://doi.org/10.3390/app8081368>
- Kropczynski, J., Oluwadare, S., Elrod, N., Mahajan, P., Asadi, A., Koch, J., ve Botsyoe, L. E. (2024, May). *Towards a standardized data framework for threat detection in smart homes*. In 2024 IEEE 24th International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW) (ss. 62–67). IEEE. <https://doi.org/10.1109/CCGridW63211.2024.00037>
- Kumar, S., ve Buyya, R. (2012). *Green cloud computing and environmental sustainability*. In S. Murugesan ve G. R. Gangadharan (Eds.), *Harnessing Green IT: Principles and Practices* (pp. 315–339). John Wiley ve Sons. <https://doi.org/10.1002/9781118305393.ch16>
- Kurban, H. (2021). Metin madenciliği ile tıbbi tedavi alanlarının yakınlıklarının ölçülmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (21), 518–526. <https://doi.org/10.31590/ejosat.833199>

- Landauer, T. K., Foltz, P. W., ve Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25(2–3), 259–284.
<https://doi.org/10.1080/01638539809545028>
- Lauriola, I., Lavelli, A., ve Aioli, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*, 470, 443–456. <https://doi.org/10.1016/j.neucom.2021.05.103>
- Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... ve Wolff, S. (2009). *A brief history of the Internet*. ACM SIGCOMM Computer Communication Review, 39(5), 22–31. <https://doi.org/10.1145/1629607.1629613>
- Leitner, P., ve Cito, J. (2016). Patterns in the chaos—A study of performance variation and predictability in public IaaS clouds. *ACM Transactions on Internet Technology (TOIT)*, 16(3), 1–23. <https://doi.org/10.1145/2885497>
- Li, W., Wu, J., Cao, J., Chen, N., ve Zhang, Q. (2021). *Blockchain-based trust management in cloud computing systems: A taxonomy, review and future directions*. Journal of Cloud Computing: Advances, Systems and Applications, 10, Article 33. <https://doi.org/10.1186/s13677-021-00247-5>
- Li, X., Lv, Q., ve Huang, W. (2015). Learning similarity with probabilistic latent semantic analysis for image retrieval. *KSII Transactions on Internet and Information Systems*, 9(4), 1581–1597. <https://doi.org/10.3837/tiis.2015.04.009>
- Li, Z., He, Z., Yang, S., Ye, H., Qiao, A., ve Zhang, X. (2024). Large language models for blockchain security: A systematic literature review. *arXiv*. <https://doi.org/10.48550/arXiv.2403.14280>
- Lisowska, R., ve Pamula, A. (2020). Changes in the business models of manufacturing companies in the SME sector after the implementation of cloud computing solutions. In *Proceedings of the International Conference on Business Informatics* (pp. 345–358). Springer. https://doi.org/10.1007/978-3-030-52294-0_20
- Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139084789>
- Liu, B., Hu, M. ve Cheng, J. (2005). *Opinion observer: Analyzing and comparing opinions on the Web*. 14. Uluslararası Dünya Çapında Web Konferansı (s. 342–351). <https://doi.org/10.1145/1060745.1060797>
- Lohachab, A., Lohachab, A., ve Jangra, A. (2020). A comprehensive survey of prominent cryptographic aspects for securing communication in post-quantum IoT networks. *Internet of Things*, 9, 100174. <https://doi.org/10.1016/j.iot.2020.100174>

- Luo, L., Meng, S., Qiu, X., ve Dai, Y. (2019). Improving failure tolerance in large-scale cloud computing systems. *IEEE Transactions on Reliability*.
<https://doi.org/10.1109/TR.2019.2901194>
- Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., ve Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, 12(2–3), 93–118. <https://doi.org/10.1080/19312458.2018.1430754>
- Manning, C. D., ve Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.
- Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J. ve Ghalsasi, A. (2011). Cloud computing – The business perspective. *Decision Support Systems*, 51(1), 176–189. <https://doi.org/10.1016/j.dss.2010.12.006>
- Mathur, P. (2024). Cloud Computing Infrastructure, Platforms, and Software for Scientific Research. In *Computing in Biomimetics: Modeling, Architecture, and Applications*. Springer. https://doi.org/10.1007/978-981-97-1017-1_4
- Mell, P., ve Grance, T. (2011). *The NIST definition of cloud computing* (NIST Special Publication 800-145). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-145>
- Merlec, M. M., ve In, H. P. (2024). Blockchain-based decentralized storage systems for sustainable data self-sovereignty: A comparative study. *Sustainability*, 16(17), 7671. <https://doi.org/10.3390/su16177671>
- Minh, Q. N., Nguyen, V. H., Quy, V. K., ve Chehri, A. (2021). *Edge computing for IoT-enabled smart grid: The future of energy*. *Energies*, 15(17), 6140. <https://doi.org/10.3390/en15176140>
- Miryala, N. K., ve Gupta, D. (2022). Data security challenges and industry trends. *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*, 11(11), 300–309. <https://doi.org/10.17148/IJARCCE.2022.111160>
- Modi, C., Patel, D., Borisaniya, B., Patel, A. ve Rajarajan, M. (2013). A survey on security issues and solutions at different layers of cloud computing. *The Journal of Supercomputing*, 63(2), 561–592. <https://doi.org/10.1007/s11227-012-0831-5>
- Mongeon, P., ve Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106, 213–228. <https://doi.org/10.1007/s11192-015-1765-5>

- Morris, K. (2020). *Infrastructure as Code: Dynamic Systems for the Cloud Age* (2nd ed.). O'Reilly Media. ISBN: 9781492075954.
- Muppala, G., ve Devi, T. (2023, September). Accurate recasting of giant text into charts using rapid automatic keyword extraction algorithm in comparison with bag of words algorithm. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 6, pp. 2548–2552). IEEE.
<https://doi.org/10.1109/IC3I59117.2023.10397804>
- Mustafa, F., ve Bukhari, S. (2023). *Cybersecurity in cloud-based financial systems: Protecting modern markets and digital assets*. ResearchGate.
<https://doi.org/10.13140/RG.2.2.20886.56641>
- National Institute of Standards and Technology – NIST. (2020). *Zero Trust Architecture (SP 800-207)*. U.S. Department of Commerce.
<https://doi.org/10.6028/NIST.SP.800-207>
- Newman, D., Asuncion, A., Smyth, P., ve Welling, M. (2009). Distributed Algorithms for Topic Models. *Journal of Machine Learning Research*, 10(Aug), 1801–1828.
<https://www.jmlr.org/papers/volume10/newman09a/newman09a.pdf>
- Newman, M. E. J. (2018). *Networks: An introduction*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199206650.001.0001>
- Ni, C.-F., Lundblad, R., Dykeman, C., Bolante, R., ve Łabuński, W. (2023). *Content analysis of psychological first aid training manuals via topic modelling*. *European Journal of Psychotraumatology*, 14(2), 2230110.
<https://doi.org/10.1080/20008066.2023.2230110>
- Noble, P. J. M., Appleton, C., Radford, A. D., ve Nenadic, G. (2021). Using topic modelling for unsupervised annotation of electronic health records to identify an outbreak of disease in UK dogs. *PLOS ONE*, 16(12), e0260402.
<https://doi.org/10.1371/journal.pone.0260402>
- Ofili, B. T., Erhabor, E. O. ve Obasuyi, O. T. (2025). *Enhancing federal cloud security with AI: Zero trust, threat intelligence and CISA Compliance*. *World Journal of Advanced Research and Reviews*, 25(2), 2377–2400.
<https://doi.org/10.30574/wjarr.2025.25.2.0620>
- Ogendi, E. G. (2025). Leveraging advanced cybersecurity analytics to reinforce zero-trust architectures within adaptive security frameworks. *International Journal of Research Publication and Reviews*, 6(2), 691–704.
<https://doi.org/10.55248/gengpi.6.0225.0729>

- Okatan, B., ve Çam, H. (2024). Analysis of customer reviews for digital banking applications with text mining methods. *Gümüşhane Üniversitesi Fen Bilimleri Dergisi*, 14(1), 45–60. <https://doi.org/10.17714/gumusfenbil.1361431>
- O'Loughlin, J., ve Gillam, L. (2015). *Addressing issues of cloud resilience, security and performance through simple detection of co-locating sibling virtual machine instances*. In CLOSER 2015- Proceedings of the 5th International Conference on Cloud Computing and Services Science, Lisbon, Portugal, 20–22 May, 2015 (pp. 60–67). SCITEPRESS. <https://doi.org/10.5220/0005485000600067>
- Onan, A., Yalçın, A. ve Atik, E. (2020). Üniversite Bilgi Yönetim Sistemi servis destek taleplerinin konu modelleme tabanlı analizi. *Avrupa Bilim ve Teknoloji Dergisi*, (Özel Sayı), 389–397. <https://doi.org/10.31590/ejosat.780642>
- Pakdil, M. E. ve Çelik, R. N. (2023). Bulut bilişimde sunucusuz mimariler ile coğrafi bilgi teknolojilerinin kullanımı üzerine bir inceleme. *Jeodezi ve Jeoinformasyon Dergisi*, 10(1), 1–15. <https://doi.org/10.9733/JGG.2023R0001>. T
- Paleyes, A., Urma, R. G., ve Lawrence, N. D. (2022). Challenges in deploying machine learning: a survey of case studies. *ACM Computing Surveys*, 55(6), 1–29. <https://doi.org/10.1145/3533378>
- Pamula, A., Gontar, Z., Gontar, B., ve Fesenko, T. (2023). *Latent Dirichlet Allocation in public procurement documents analysis for determining energy efficiency issues in construction works at Polish universities*. *Energies*, 16(12), 4596. <https://doi.org/10.3390/en16124596>
- Priyanka, E. B., Thangavel, S., ve Prabhakaran, P. (2024). *IoT with deep learning in pipeline and metro track risk estimation using smart cities development*. In *New Advances in Soft Computing in Civil Engineering: AI-Based Optimization and Prediction* (ss. 99–118). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-65976-8_5
- Qomariyah, S., Nur, I., ve Fithriasari, K. (2024). Topic modeling Twitter data using Latent Dirichlet Allocation and Latent Semantic Analysis. *AIP Conference Proceedings*, 2194, 040006. <https://doi.org/10.1063/1.5139825>
- Ramamoorthi, V. (2024). *AI-driven resource management in cloud computing: A review*. *International Journal of Scientific Research in Science and Technology*, 11(3), 225–231. <https://doi.org/10.32628/IJSRST24113260>
- Rasiwasia, N., ve Vasconcelos, N. (2013). Latent Dirichlet Allocation models for image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11), 2665–2679. <https://doi.org/10.1109/TPAMI.2013.69>

- Ravi, D. (2017). *Conceptual framework for evaluating and selecting PaaS solution for airline company* (Master's thesis, Helsinki Metropolia University of Applied Sciences). Theseus. <https://www.theseus.fi/handle/10024/137102>
- Ravi, V., ve Kamaruddin, S. (2017). Big data analytics enabled smart financial services: Opportunities and challenges. In *Big Data Analytics* (pp. 13–26). Springer. https://doi.org/10.1007/978-3-319-72413-3_2
- Ren, W., ve Han, K. (2014). Sentiment detection of web users using probabilistic latent semantic analysis. *Journal of Multimedia*, 9(10), 1194–1200. <https://doi.org/10.4304/jmm.9.10.1194-1200>
- Rijcken, E., Kaymak, U., Scheepers, F., Mosteiro, P., Zervanou, K., ve Spruit, M. (2022). *Topic modeling for interpretable text classification from EHRs*. *Frontiers in Big Data*, 5, 846930. <https://doi.org/10.3389/fdata.2022.846930>
- Rimal, B. P., Jukan, A., Katsaros, D., ve Goeleven, Y. (2011). Architectural requirements for cloud computing systems: An enterprise cloud approach. *Journal of Grid Computing*, 9(1), 3–26. <https://doi.org/10.1007/s10723-010-9171-y>
- Ristenpart, T., Tromer, E., Shacham, H., ve Savage, S. (2009). *Hey, you, get off of my cloud: Exploring information leakage in third-party compute clouds*. In *Proceedings of the 16th ACM Conference on Computer and Communications Security* (pp. 199–212). ACM. <https://doi.org/10.1145/1653662.1653687>
- Sadoughi, F., ve Erfannia, L. (2017). Health information system in a cloud computing context. In D. Hayn ve G. Schreier (Eds.), *Health Informatics Meets eHealth* (pp. 290–297). IOS Press. <https://doi.org/10.3233/978-1-61499-759-7-290>
- Sariman, G., ve Mutaft, E. (2020). COVID-19 sürecinde Twitter mesajlarının duygu analizi. *Euroasia Journal of Mathematics, Engineering, Natural ve Medical Sciences*, 7(10), 137–148. <https://doi.org/10.38065/euroasiaorg.149>
- Sarkar, S., Choudhary, G., Shandilya, S. K., Hussain, A., ve Kim, H. (2022). Security of zero trust networks in cloud computing: A comparative review. *Sustainability*, 14(18), 11213. <https://doi.org/10.3390/su141811213>
- Sefraoui, O., Aissaoui, M., ve Eleulldj, M. (2012). *OpenStack: Toward an open-source solution for cloud computing*. *International Journal of Computer Applications*, 55(3), 38–42. <https://doi.org/10.5120/8738-2991>
- Senyo, P. K., Addae, E., ve Boateng, R. (2018). Cloud computing research: A review of research themes, frameworks, methods and future research directions. *International Journal of Information Management*, 38(1), 128–139.

<https://doi.org/10.1016/j.ijinfomgt.2017.07.007>

Shah, C., ve Sebastian, M. (2020). Sentiment analysis and topic modelling of Indian government's Twitter handle #IndiaFightsCorona. In *Social Media and COVID-19* (pp. 339–351). Springer. https://doi.org/10.1007/978-3-030-64861-9_30

Sharma, B., Halder, R., ve Singh, J. (2020, January). Blockchain-based interoperable healthcare using zero-knowledge proofs and proxy re-encryption. In *2020 International Conference on COMmunication Systems ve NETworkS (COMSNETS)* (pp. 1–6). IEEE.

<https://doi.org/10.1109/COMSNETS48256.2020.9027413>

Sharma, C., Batra, I., Sharma, S., Malik, A., Hosen, A. S. M., ve Ra, I.-H. (2022). Predicting trends and research patterns of smart cities: A semi-automatic review using Latent Dirichlet Allocation (LDA). *IEEE Access*, *10*, 121068–121083.

<https://doi.org/10.1109/ACCESS.2022.3214310>

Shen, Z., Yuan, M., Wu, H., Han, Y., ve Zhou, J. (2024). *Design and implementation of a web-based collaborative manufacturing system based on knowledge graph*. In 2024 IEEE 4th International Conference on Collaborative Manufacturing (CCM). IEEE. <https://ieeexplore.ieee.org/document/10778878>

Shi, B. ve Min, Y. (2023). Analyzing ESG news articles and research papers through LDA (Latent Dirichlet Allocation). *Asia-Pacific Journal of Convergent Research Interchange*, *9*(8), 73–85. <https://doi.org/10.47116/apjcri.2023.08.07>

Shi, W., Cao, J., Zhang, Q., Li, Y., ve Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, *3*(5), 637–646. <https://doi.org/10.1109/JIOT.2016.2579198>

Shrivastava, S., Saini, G., ve Agrawal, Y. (2020). Multi-cloud deployments and hybrid cloud architecture. *Journal of Critical Reviews*, *7*(4), 4754–4760. <https://pdfs.semanticscholar.org/59ee/a45e0ac3c0ac3dbd58ecd6915806e45fc9f7.pdf>

Silveira, R., Fernandes, C. G. O., Monteiro Neto, J. A., Furtado, V., ve Pimentel Filho, J. E. (2021). Topic modelling of legal documents via LEGAL-BERT. In *Proceedings of the RELATED Workshop at the International Conference on Artificial Intelligence and Law (ICAIL 2021)* (pp. 1–14). CEUR Workshop Proceedings. <https://ssrn.com/abstract=4539091>

Sivakumaren, K. S., Swaminathan, S., ve Karthikeyan, G. (2019). Growth and development of publications on cloud computing: A scientometric study.

- International Journal of Information Library ve Society*, 8(2), 37–43.
<https://www.researchgate.net/publication/336799479>
- Sriram, I., ve Khajeh-Hosseini, A. (2010). *Technical aspects of cloud computing: A review and research agenda*. arXiv preprint arXiv:1001. 3259.
<https://doi.org/10.48550/arXiv.1001.3259>
- Srivastava, A., ve Sutton, C. (2017). Autoencoding variational inference for topic models. In *Proceedings of the 5th International Conference on Learning Representations (ICLR 2017)*. arXiv. <https://arxiv.org/abs/1703.01488>
- Steyvers, M., ve Griffiths, T. (2007). Probabilistic topic models. In T. K. Landauer, D. S. McNamara, S. Dennis, ve W. Kintsch (Eds.), *Handbook of latent semantic analysis* (ss. 439–460). Psychology Press.
- Subashini, S. ve Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. *Journal of Network and Computer Applications*, 34(1), 1–11. <https://doi.org/10.1016/j.jnca.2010.07.006>
- Subramanian, N., ve Jeyaraj, A. (2018). *Recent security challenges in cloud computing*. *Computers ve Electrical Engineering*, 71, 28–42.
<https://doi.org/10.1016/j.compeleceng.2018.06.012>
- Sudolska, A., Lis, A., ve Błaś, R. (2019). *Cloud computing research profiling: Mapping scholarly community and identifying thematic boundaries of the field*. *Social Sciences*, 8(4), 112. <https://doi.org/10.3390/socsci8040112>
- Sultan, N. (2014). *Cloud computing for education: A new dawn?* *International Journal of Information Management*, 34(2), 177-184.
<https://doi.org/10.1016/j.ijinfomgt.2013.12.011>
- Sumathy, K. L., ve Chidambaram, M. (2013). Text mining: Concepts, applications, tools and issues – An overview. *International Journal of Computer Applications*, 75(6), 29–32. <https://doi.org/10.5120/13851-1685>
- Surbiryala, J., ve Rong, C. (2019). *Cloud computing: History and overview*. 2019 IEEE Cloud Summit. <https://doi.org/10.1109/CloudSummit47114.2019.00007>
- Syed, S., ve Spruit, M. (2018). Full-text or abstract? Examining topic coherence scores using latent Dirichlet allocation. *Proceedings of the 2017 International Conference on Data Science and Advanced Analytics*, 165–174.
<https://doi.org/10.1109/DSAA.2017.61>
- Şeker, S. E. (2015). Metin madenciliği (Text mining). *YBS Ansiklopedi*, 2(3), 30–32.
<https://www.ybsansiklopedi.com>

- Şirzad, N. (2023). İndirim kampanyalarının tüketicilerin fiyat algısı üzerindeki etkisi: Black Friday örneği. *Erciyes İletişim Dergisi*, 10(2), 913–928.
<https://doi.org/10.17680/erciyesiletisim.1255697>
- T.C. Cumhurbaşkanlığı Bilim, Teknoloji ve Yenilik Politikaları Kurulu. (2021, Aralık). *Büyük veri ve bulut bilişim teknoloji yol haritası: Sonuç raporu*.
https://tubitak.gov.tr/sites/default/files/2023-10/btypk_bvbb_tyh_sonuc_raporu.pdf
- Talley, E. M., Newman, D., Mimno, D., Herr, B. W., Wallach, H. M., Burns, G. A. P. C., ve McCallum, A. (2011). Database of NIH grants using machine-learned categories and graphical clustering. *Nature Methods*, 8(6), 443–444.
<https://doi.org/10.1038/nmeth.1619>
- Teh, R., ve Subramaniam, A. (2023). *A bibliometric analysis on cloud computing literature development*. International Journal of Innovative Computing, 13(2), 49–57. <https://doi.org/10.11113/ijic.v13n2.396>
- Teh, Y. W., Jordan, M. I., Beal, M. J., ve Blei, D. M. (2006). *Hierarchical Dirichlet processes*. Journal of the American Statistical Association, 101(476), 1566–1581.
<https://doi.org/10.1198/0162145060000000302>
- Tian, C., Zhang, J., Liu, D., Wang, Q., ve Lin, S. (2022). Technological topic analysis of standard-essential patents based on the improved Latent Dirichlet Allocation (LDA) model. *Technology Analysis ve Strategic Management*, 36(9), 2084–2099.
<https://doi.org/10.1080/09537325.2022.2130039>
- Tiwari, S., Dey, S., ve Sarma, W. (2021). Optimizing high-performance and scalable cloud architectures: A deep dive into serverless, microservices, and edge computing paradigms. International Journal of Innovative Research in Computer and Communication Engineering, 9(8), 9665–9680.
<https://doi.org/10.15680/IJIRCCE.2021.0908051>
- Toklu, R., Gök, M., Güven, E., ve Eren, T. (2024). Türkiye’de bulut bilişim üzerine yapılmış tezlerin bibliyometrik analizi. *Selçuk Turizm ve Bilişim Araştırmaları Dergisi*, 6, 37–55. <https://dergipark.org.tr/tr/pub/stbad>
- Torun, N. K., ve Şengül, A. (2022). Kripto para birimlerinin Twitter verileri ile metin madenciliği kapsamında incelenmesi. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 6(1), 54–65.
<https://doi.org/10.33461/uybisbbd.1089670>

- Varghese, B., ve Buyya, R. (2018). Next generation cloud computing: New trends and research directions. *Future Generation Computer Systems*, 79, 849–861. <https://doi.org/10.1016/j.future.2017.09.020>
- Varma, M. K., ve Choi, E. (2016). Comparative Study of Various Platform as a Service Frameworks. *International Journal of Cloud Computing: Services and Architecture (IJCCSA)*. <https://doi.org/10.5121/ijccsa.2016.6103>
- Voorsluys, W., Buyya, R., ve Broberg, J. (2011). *Service level agreements in cloud computing*. In R. Buyya, J. Broberg, ve A. Goscinski (Eds.), *Cloud Computing: Principles and Paradigms* (pp. 285–320). Wiley. <https://doi.org/10.1002/9780470940105.ch1>
- Wallach, H. M., Mimno, D., ve McCallum, A. (2009). Rethinking LDA: Why priors matter. In *Proceedings of the 26th Annual International Conference on Machine Learning (ICML 2009)* (pp. 1065–1072). Association for Computing Machinery. <https://doi.org/10.1145/1553374.1553515>
- Wang, N., Liang, H., Jia, Y., Ge, S., Xue, Y., ve Wang, Z. (2016). Cloud computing research in the IS discipline: A citation/co-citation analysis. *Decision Support Systems*, 86, 35–47. <https://doi.org/10.1016/j.dss.2016.03.006>
- Wang, X., Chen, M., Taleb, T., Ksentini, A., ve Leung, V. C. (2014). Cache in the air: Exploiting content caching and delivery techniques for 5G systems. *IEEE Communications Magazine*, 52(2), 131–139. <https://doi.org/10.1109/MCOM.2014.6736753>
- Weiss, S. M., Indurkha, N., ve Zhang, T. (2010). *Fundamentals of predictive text mining*. Springer. <https://doi.org/10.1007/978-1-84996-226-1>
- Weiss, S. M., Indurkha, N., Zhang, T., ve Damerau, F. (2010). *Text mining: Predictive methods for analyzing unstructured information*. Springer Science ve Business Media.
- Wiemer-Hastings, P., Wiemer-Hastings, K., ve Graesser, A. C. (2004). Latent Semantic Analysis. *Discourse Processes*, 37(2), 219–240. https://doi.org/10.1207/s15326950dp3702_4
- Xie, J., Tang, H., Huang, T., Yu, F. R., Xie, R., Liu, J., ve Liu, Y. (2019). A survey of blockchain technology applied to smart cities: Research issues and challenges. *IEEE Communications Surveys ve Tutorials*, 21(3), 2794–2830. <https://doi.org/10.1109/COMST.2019.2899617>

- Xue, J., Chen, J., Chen, C., Zheng, C., Li, S., ve Zhu, T. (2020). *Public discourse and sentiment during the COVID-19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter*. PLOS ONE, 15(9), e0239441. <https://doi.org/10.1371/journal.pone.0239441>
- Yadav, R. (2024). Harnessing Chaos: The Role of Chaos Engineering in Cloud Applications and Impacts on Site Reliability Engineering. *International Journal of Computer Trends and Technology*, 72(6), 25–30. <https://doi.org/10.14445/22312803/IJCTT-V72I6P104>
- Yadav, R., Zhang, W., Li, K., Liu, C. ve Laghari, A. A. (2021). Managing overloaded hosts for energy-efficiency in cloud data centers. *Cluster Computing*, 1–15. <https://doi.org/10.1007/s10586-020-03182-3>
- Yan, J., Zeng, J., Gao, Y., ve Liu, Z. Q. (2015). *Communication-efficient algorithms for parallel latent Dirichlet allocation*. Soft Computing, 19(11), 3141–3151. <https://doi.org/10.1007/s00500-014-1376-8>
- Yaseen, M. G., ve Albahri, A. S. (2023). Mapping the evolution of intrusion detection in big data: A bibliometric analysis. *Malaysian Journal of Big Data*, 1(3), [Makale sayfa numarası bilinmiyor]. <https://doi.org/10.58496/MJBD/2023/018>
- Yılmaz, E., ve Polat, E. Ç. (2023). #VEGAN: Veganizm üzerine Twitter sosyal ağ analizi. *Moment Dergi*, 10(1), 48–82. <https://doi.org/10.17572/mj2023.1.4882>
- Yiğit, K., ve Ayata, F. (2024). *Bulut bilişim mimarisi ve güvenliği*. Computer Science, 9(1), 62-71. <https://doi.org/10.53070/bbd.1448462>
- Young, T., Hazarika, D., Poria, S., ve Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, 13(3), 55–75. <https://doi.org/10.1109/MCI.2018.2840738>
- Younis, R., Iqbal, M., Munir, K., ve Javed, M. A. (2024). A comprehensive analysis of Cloud service models: IaaS, PaaS, and SaaS in the context of emerging technologies and trend. *IEEE Transactions on Computers*. <https://doi.org/10.1109/ICECCE63537.2024.10823401>
- Yu, J., Yang, Z., Zhu, S., Xu, B., Li, S., ve Zhang, M. (2018). A bibliometric analysis of cloud computing technology research. In *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (pp. 2353–2358). IEEE. <https://doi.org/10.1109/IAEAC.2018.8577750>
- Yu, J., Yang, Z., Zhu, S., Xu, B., Li, S., ve Zhang, M. (2018). A bibliometric analysis of cloud computing technology research. *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 348–353.

<https://doi.org/10.1109/ICCCBDA.2018.8386578>

Zhang, Q., Cheng, L., ve Boutaba, R. (2010). *Cloud computing: state-of-the-art and research challenges*. Journal of Internet Services and Applications, 1(1), 7–18.
<https://doi.org/10.1007/s13174-010-0007-6>

Zhang, M., Krintz, C., ve Wolski, R. (2020, Mart). *Stoic: Serverless teleoperable hybrid cloud for machine learning applications on edge devices*. 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 1–6. IEEE.

<https://doi.org/10.1109/PerComWorkshops48775.2020.9156239>

Zhou, W., Ning, P., Zhang, X., Ammons, G., Wang, R., ve Bala, V. (2010, Aralık). *Always up-to-date: Scalable offline patching of VM images in a compute cloud*. In Proceedings of the 26th Annual Computer Security Applications Conference (pp. 377–386). <https://doi.org/10.1145/1920261.1920317>

Zissis, D., ve Lekkas, D. (2012). Addressing cloud computing security issues. *Future Generation Computer Systems*, 28(3), 583–592.

<https://doi.org/10.1016/j.future.2010.12.006>

Zupic, I., ve Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429–472.

<https://doi.org/10.1177/1094428114562629>

ÖZGEÇMİŞ

Emrah ELBASDI, İlk ve ortaöğrenimini Ankara Gülhane İlköğretim Okulu'nda, lise öğrenimini ise İncirli Teknik Lisesi'nde tamamlamıştır. Lisans eğitimine Kocaeli Üniversitesi Teknik Eğitim Fakültesi Elektrik Öğretmenliği bölümünde devam etmiş ve mezun olmuştur. Ardından Karamanoğlu Mehmetbey Üniversitesi Elektrik-Elektronik Mühendisliği lisans programını da başarıyla tamamlamıştır.

Halen Gümüşhane Üniversitesi Lisansüstü Eğitim Enstitüsü Yönetim Bilişim Sistemleri Yüksek Lisans programında öğrenimine devam etmektedir.

2006 yılından bu yana İçişleri Bakanlığı'na bağlı bir kamu kuruluşunda görev yapmaktadır. Bu süreçte bilişim alanında çeşitli eğitim programlarına ve seminerlere katılmıştır. İngilizce dilinde B1 düzeyinde yeterliliğe sahiptir.