



Research Article

Doküman Tabanlı Veri Paylaşım Platform Mimarisi

Abdülkadir Karabacak^{1*}, Ergün Okay², Mehmet S. Aktaş³

¹Intellica Ar-Ge Merkezi, İstanbul, Türkiye, abdulcadir.karabacak@intellica.net

²Intellica Ar-Ge Merkezi, İstanbul, Türkiye, ergun.okay@intellica.net

³Yıldız Teknik Üniversitesi, İstanbul, Türkiye, aktas@yildiz.edu.tr

Sorumlu Yazar: Abdülkadir Karabacak

(First received October 02, 2022 and in final form December 19, 2022)

2nd International Conference on Design, Research and Development
December 14 - 17, 2022

Reference: Karabacak, A., Okay, E., Aktaş, M. S. Doküman Tabanlı Veri Paylaşım Platform Mimarisi. Orclever Proceedings of Research and Development,1(1), 339-348.

Özet

Büyük ölçekli kurum/kuruluşlar için Büyük Veri yeni iç görü potansiyeli sağlamak için önemli bilgiler içermektedir. Endüstri 4.0 ile birlikte gelişen yeni teknolojik gelişmeler nedeniyle veriler artan hacimlerde üretilmektedir. Bu verilerin hacimleriyle başa çıkmak ve verilerin değerli bilgilere dönüştürülebilmesi için Veri Paylaşım Platformlarına ihtiyaç duyulmaktadır.

Bu ihtiyaç doğrultusunda, bu araştırma kapsamında, bir doküman tabanlı veri paylaşım platformu yazılım mimarisi önerilmektedir. Önerdiğimiz Veri Paylaşım Platformu Mimarisi; verileri analitik amaçlarla uygun ölçekte işlemek üzere tasarlanmış bir doküman tabanlı veri yönetim platformu için tasarlanmıştır.

Önerilen çalışmada; elde edilen büyük hacimli verilerin karmaşılaşması, kullanılamaması ve bataklığa dönüşmesini engellemek üzere meta veri yönetimi kullanılmaktadır. Önerilen mimari yapısında, verilerin sahibini belirleyebilmek, sürüm ve köken bilgilerinin saklayabilmek için zenginleştirilmiş araç seti içeren bir meta veri deposu bulunmaktadır. Çalışmada; doğru verilere kolay ulaşımın sağlanabilmesi için ihtiyaç duyulan verilerin konumları kullanıcılara detaylı şekillerde gösterilmektedir. Verilerin en uygun kalitede temizlenebilmesi için sistem üzerinde kullanıcının veri ön işleme yapabilmesini sağlayacak ek geliştirme çalışmaları entegre edilmiştir.



Verilerin güvenli şekilde kullanabilmesi amacıyla işlevsel bir güvenlik kontrolü bulunmaktadır. Önerilen yazılım mimarisinde, işletim sistemleri özelinde değişkenlik gösterebilecek seviyede standart bir kullanıcı grubu yönetimi entegre edilmiştir. Yine, önerilen yazılım mimarisinde, stokastik veri kümeleri içerisinde etiketleme yapılarak verilerin kategorize edilir ve sonraki kullanıcıların rollerine göre öneri yapabilecek şekilde öneriler sunabilmektedir. Ek olarak zaman içerisindeki değişiklikler ile başa çıkmak için de sürüm ve kural uyarlama yöntemi sağlanmaktadır. Sistemin her kullanıcının özel gereksinimlerine yanıt vermek ihtiyacını karşılamak için kişiselleştirilmiş kural uyarlama yöntemi önerilmektedir.

Geliştirmekte olduğumuz doküman tabanlı veri paylaşım platformu yazılım mimarisinin detaylarını, bu konferans bildirisi kapsamında aktarıyoruz.

Anahtar Kelimeler: Veri Paylaşım Platformu, Büyük Veri, Doküman Tabanlı Veri Paylaşımı, Platform Yazılımı, Yazılım Mimarisi

Document Based Data Sharing Platform Architecture

Abstract

The Big Data contains essential information for large organizations to provide new insight potential. Due to the new technological developments that have developed with Industry 4.0, data is produced in increasing volumes. Data Sharing Platforms are needed to cope with the volumes of this data and to transform data into valuable information. In line with this need, a document-based data-sharing platform software architecture is proposed within the scope of this research. The Data Sharing Platform Architecture we recommend; is designed for a document-based data management platform designed to process data at scale for analytical purposes. In the proposed study, Metadata management is used to prevent the large volume of data obtained from becoming complex and unusable. The proposed architecture has a metadata store with an enriched toolset to identify the data owner and store the version and lineage information. In the study, to provide easy access to the correct data, the locations of the data needed are shown to the users in detailed figures. To clean the data in the most appropriate quality, additional development studies are integrated into the system that will enable the user to pre-process the data. There is an operational security control to use the data securely. A standard user group management, which may vary according to operating systems, is integrated into the proposed software architecture. Again, the proposed software architecture categorizes the data by tagging it in stochastic data sets. It can offer suggestions in a way that can make suggestions according to the roles of the following users. In addition, a version and rule adaptation method is provided to deal with changes over time. A



personalized rule customization method is proposed to meet the system's need to respond to the specific needs of each user. We present the details of the document-based data-sharing platform software architecture we are developing within the scope of this conference paper.

Keywords: Data Sharing Platform, Big Data, Document Based Data Sharing, Platform Software, Software Architecture

1. Giriş

Büyük ölçekli kurum/kuruluşlar için veriler yeni içgörü potansiyeli sağlamak için önemli bilgiler içermektedir. Endüstri 4.0 ile birlikte gelişen yeni teknolojik gelişmeler nedeniyle veriler artan hacimlerde üretilmektedir. Bu verilerin hacimleriyle başa çıkabilmek ve farklı veri türleri arasındaki korelasyon çarpışmasının giderek daha yoğun hale gelmesi sebebiyle veri ambarlarını etkili bir şekilde kullanılabilecek ve veri toplama ile paylaşımını birleştirebilecek hibrit veri platformlarına ihtiyaç duyulmaktadır.

Bu araştırma kapsamında, yukarıda özetlediğimiz ihtiyaca, bir çözüm önerisi getirilmiştir. Bu amaçla bir Veri Paylaşım Platformu yazılım mimarisi önerisi yapılmıştır.

Araştırmamız kapsamında, işletmelerin iş zekasını derin ve fonksiyonel şekilde kullanabilmeleri amacıyla doküman tabanlı veri paylaşımı yapabilecekleri yazılımın tasarımı ve geliştirilmesi için ar-ge faaliyetleri yürütülmüştür. Önerilen Veri Paylaşım Platformu; verileri analitik amaçlarla uygun ölçekte işlemek üzere tasarlanmış bir Doküman Tabanlı bir Veri Yönetim Platformudur. Kuruluşların kendi iç bünyelerinde oluşturdukları heterojen kaynaklardan toplanacak veriler için merkezi yönetimi sağlayacak şekilde tek bir veri platformu üzerinden verilerin çeşitliliğini ve dinamiklerini destekleyecek şekilde yazılımın tasarım gerçekleştirilmektedir. Verilerin depolanması ve analiz edilmesi sonrasında platformun veri bataklığına dönüşmesini engellemek üzere meta veri yönetimi kullanılmaktadır. Meta veri yönetimi ile verileri depolamak ve işlemek için birçok fonksiyon kullanılmıştır. Verilerin sahibini belirleyebilmek, sürüm ve köken bilgilerinin saklayabilmek için zenginleştirilmiş araç seti içeren bir meta veri deposu bulunmaktadır. Çalışmada; doğru verilere kolay ulaşımın sağlanabilmesi için ihtiyaç duyulan verilerin konumları kullanıcılara detaylı şekillerde gösterilmektedir. Verilerin en uygun kalitede temizlenebilmesi için sistem üzerinde kullanıcının veri ön işleme yapabilmesini sağlayacak ek geliştirme çalışmaları entegre edilmiştir. Verilerin güvenli şekilde kullanabilmesi amacıyla işlevsel bir güvenlik kontrolü bulunmaktadır. İşletim sistemleri özelinde değişkenlik gösterebilecek seviyede standart bir kullanıcı grubu yönetimi entegre edilmiştir. Bu alan, Her kullanıcı veya gruba bu kontrol paneli



üzerinden Veri Platformu'na erişim, okuma veya yazma izinlerinden birini veya birkaçını verebilmektedir.

Bu araştırma kapsamında merkezi ortamda konumlanmış, dağıtık sistem altyapısı üzerinde kurgulanmış, doküman tabanlı veri paylaşımı gerçekleştiren hibrit veri paylaşımı gerçekleştirebilen bir yazılım platformu mimarisinin nasıl olabileceği ve hangi yöntemlerle geliştirilebileceği problemi üzerinde çalışılmaktadır. Yine araştırma kapsamında, bu yazılım mimarisinin prototip uygulamasının hangi kriterlere dayalı olarak değerlendirilebileceği önerilmektedir.

Bu bildirinin organizasyon yapısı şu şekildedir. Bölüm 2, temel kavramlar ve literatür özetini vermektedir. Bölüm 3, bu araştırma kapsamında, yaptığımız yazılım mimarisi önerisini, Bölüm 4'te prototip uygulama detaylarını ve önerdiğimiz performans değerlendirme kriterlerini vermektedir. Son bölüm ise, araştırmanın özetini ve gelecekteki çalışmaları aktarmaktadır.

2. Literatür Özeti

Veri Paylaşım Platformları özellikle veri ambarları ile karıştırılabilmektedir. Veri paylaşım platformları hem veri ambarı hem de veri havuzları olarak düşünülebilir. Ancak bunlar uygulamada farklı kavram ve yapılardan oluşabilmektedir. Veri ambarları iyi tanımlanmış düzenleyici işlevlere ve depolama kapasitesine sahip olmakla birlikte veri paylaşım platformlarında teorik olarak veri depolama limiti bulunmamaktadır. Her türlü veri her miktarda yüklenebilmektedir. Veri paylaşım platformu sağlayan mimariler kurum/kuruluşların veri türü ve yapısının ötesine bakmasını sağlayan ve istedikleri kadar veri yükleme olanağı olan yapılardır.[1] Veri ambarı ise yüksek düzeyde yapılandırılmış verilerin işlenmesinin tam aksine kullanıcılarına önceden oluşturulmuş olan sorgu yapısına göre veri özetleri sunmaktadır. Literatür araştırmalarında denk gelen çalışmaların aynı amaçla veri ekosistemleri üzerinden birden fazla kaynaktan alınabilen veriler üzerinde paylaşımlar yapılmanın önemi ve sistem tasarımının kritik noktaları vurgulanmıştır.[3] Verilerin paylaşılması esnasında hassas verilerin güvenliğinin sağlanması, veri alışverişi esnasından kümeleme sorunlarına otomatik çözümler sağlanması, verilerin bulut ortamında aktarım problemlerine çözümler sağlanması veri platformlarının mimarisinde çözülmesi gereken problemler olarak tespit edilmiştir.[2]

Ayrıca yine Veri paylaşım platformlarında meta veri yönetimi, veri modellenmesi ve veri yönetimi içeren bağlantılı bir düzüm süreci, rol ve kural tabanlı yapının mimarilerde yer alması gerekliliği görülmektedir. Bir veri paylaşımı yaşam döngüsü boyunca meta verilerin nasıl ele alınması ve sürdürülmesi gerektiğine ilişkin veri



yönetimine ilişkin standartlar ve politikalar içermelidir. Meta verilerin kontrollü aşamalarında kapsam öncelikleri belirlenmelidir. Özellikle her etki alanında tanımlanan en kritik veri öğelerinin sınıflandırılması gerekir.[5]

Veri Paylaşım Platformlarında Süreçleri, politikaları ve kuralları incelerken ve verilerin oluşturulması, kullanılması ve belgelenmesiyle ilişkili olarak, verilerin nasıl yönetilmesi ve kontrol edilmesi gerektiğine dair bir dizi süreç noktası ve beklenti olmalıdır. Örneğin, çözüm tasarım süreçlerinde ve veri paylaşım platformlarını destekleyen yönetim ekipleri ve iş odaklı ekiplerinden değişiklik kontrolü desteğiyle ilgilenen veri görevlileri aracılığıyla oluşturulan veri paylaşım platformu gereksinimleri ve standartları olmalıdır. Bu süreç noktalarından, aktivite gereksinimlere, kurallara, standartlara ve politikalara göre ölçülebilir. Bunların tümü, veri paylaşım yönetimi performansını ölçmek ve izlemek için operasyonel düzeydeki fırsatlardır. [6]

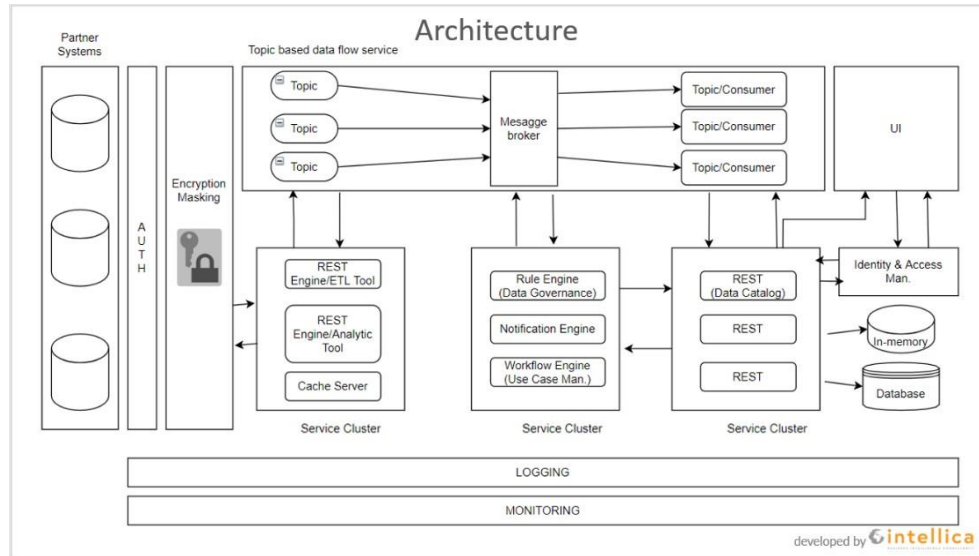
Bir veri müştereklerindeki verileri düzenlemek, içe aktarmak ve analiz etmek için Veri Paylaşım Platformlarının, veri kümelerini analiz etmek, uyumlu hale getirmek ve paylaşmak için ortak veriler ile çalışmak için doküman tabanlı olarak paylaşımının en basit yöntem olduğu vurgulanmıştır. [4] .

Önerilen mimaride sunulan veri paylaşım platformunun çözüm sağlamak istediği kritik yöntemlerden Süreç Madenciliği ile iş süreçlerinin sürekli olarak yeni gerçeklere uyum sağlayabilmesi hedeflenmektedir. Süreç madenciliği (PM), bu zorluğu ele alan, olgunlaşan bir veri odaklı süreç analizi teknikleri alanıdır. PM teknikleri, süreç modellerini otomatik olarak keşfetmek ve görselleştirmek gibi süreçle ilgili bilgileri çıkarmak için olay günlüklerini girdi olarak alır. PM uygulamalarının popüleritesi hem endüstride hem de akademide artıyor ve PM'nin makine öğrenimi, simülasyon ve diğer tamamlayıcı trendlerle entegrasyonu, Bir Kuruluşun Dijital İkizleri gibi, büyük ilgi görüyor. Bununla birlikte, PM'nin başarısı girdi olay günlüklerinin kalitesiyle doğrudan ilişkilidir, bu nedenle yüksek kaliteli olay günlüklerine olan ihtiyaç açıktır. On yıl önce PM manifestosu, olay verilerinin birinci sınıf vatandaşlar olarak ele alınması gerektiğini belirterek yüksek kaliteli olay günlüklerinin önemini vurgulamış olsa da, olay günlükleri genellikle hala bir IS'nin "yan ürünleri" olarak görülüyor. PM araştırma alanı içinde bile, olay günlükleri üzerine yapılan araştırmalar çoğunlukla geçici hazırlık tekniklerine odaklanır ve olay günlüğü yönetimine ilişkin araştırmalar kritik derecede eksiktir.[7] Veri paylaşımı esnasında gerçekleştirilecek olan olayların veri sağlayıcılar tarafından verilecek izinler ile doğru orantılı şekilde işlenebilir. Verilerin, çalışanların çoğunluğu tarafından bulunabilmesi, anlaşılabilmesi ve erişilebilmesi için demokratikleştirilmesini sağlamak üzere sözde veri sağlayıcılar, önce verileri yayınlamalı ve provizyon seçenekleri sunmalıdır. Ancak, veri sağlayıcıların verilerini paylaşma konusunda teşvik

eksikliği ve artan çaba, verilerin demokratikleşmesini engellemektedir.[8] Ayrıca bu paylaşılan verilerin paylaşım platformuna entegrasyonu ve belgelenmesi süreci, farklı veri kaynaklarını tek tip düzende sağlamak için kontrol ve hizalama hedefi içermektedir. Verilere erişim muhtemelen farklı veritabanı şemalarından gelecek veriler olacağı için bu veriler için veri tutarsızlıkları, anlamsa ve benzerlik temsilleri doğru şekilde uygulama ve doğru adreslemelerin yapılması gereklidir [9].

3. Önerilen Yazılım Mimarisi

Bu araştırma kapsamında, giriş bölümünde özetlediğimiz ihtiyaçlara yönelik ve yine özetlediğimiz araştırma problemin çözümüne yönelik, dağıtık sistem tabanlı bir yazılım mimarisi önerisi yapılmaktadır. Aşağıda resmedilen yazılım mimarisi, merkezi yapıda konuşlanmış ve dağıtık sistem olarak kurgulanmış bir hesaplama ortamında, veri paylaşımı gerçekleştirilmesini sağlamak amacıyla tasarlanmıştır. Şekil 1' de verilen Doküman Tabanlı Veri Paylaşım Platformu mimarisinde önerilen çalışmanın temel modülleri yer almaktadır.



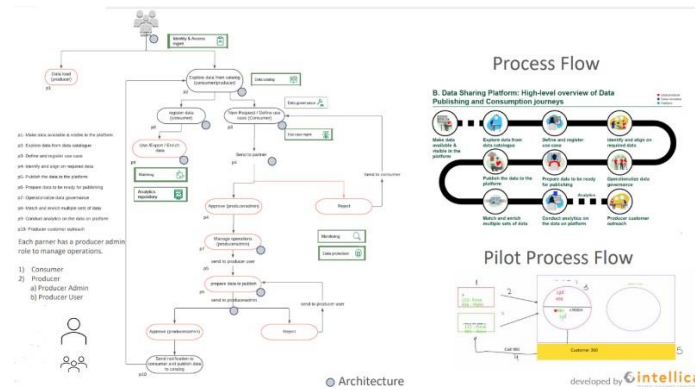
Şekil-1 Önerilen Yazılım Mimarisi

Ortak Veri Paylaşım Alanı: Şekil-1 'de, ortak veri paylaşım alanı resmedilmektedir. Burada, heterojen veri kaynaklarından alınacak veri kaynakları gösterilmektedir. Bu modül içerisinde yapılan veri işleme çalışmaları ile, veri semantiğinin oluşturulması, benzerlik tanımlamalarının yapılması ve farklı veriler arasındaki ilişkilerin ayrılması ve farklı birkaç kayıt tutulmasını sağlanmaktadır.

Çoklu Kullanıcı Erişimi: Kullanıcı giriş işlemlerinde veri girişi esnasında kurulan kontrol algoritması işletilerek, kurumsal kimlik denetimi yapılarak platforma giriş sağlanmaktadır. Geliştirilen veri paylaşım platformu “open authorization” yapısı ile birden fazla kullanıcının aynı tablolar üzerinde çalışma ve güncellemeleri gerçek zamanlı olarak gerçekleştirebilmektedir. Ayrıca yapılan tüm çalışmalar kullanıcıların ekranlarında gösterilebilmektedir.

Veri Maskeleye: Verilerin doğru, kaliteli, tekil ve ilişkisel olarak incelenmesi ve maskelenmesi ile platforma aktarımının gerçekleştirildiği alandır. Veri tabanları paket olarak şifreleme çözümleri kullanılarak hassas verilerin korunması sağlanmaktadır. Maskeleye işlemlerinin sistem performansını düşürmemesi için gerekli olan matematiksel hesaplamalar dahilinde sistemin depolama ve performans problemi olmadan çalışabilecek şekilde çalışabilmesi sağlanmıştır. Kullanılan hesaplama çözümleri şeffaf veri maskeleye yapılabilmesini sağlamaktadır. Bu işlemler için TPC-H karar destek ölçütü kullanılmaktadır.

Servis Cluster Yapısı: Doküman Tabanlı Veri Paylaşım Platformu içerisinde bulunan servis yapısındaki en önemli özellik verilerin otomatik olarak haritalanması ile farklı kaynaklardan gelen verilerin doğrudan hedef veri modeline çevrilmesini sağlayan alandır. Platform içerisinde paylaşım işlem talebi olan verilerin, tablo kolon bazında incelenerek, üst-verilerinin ve kolon bazında veri deseninin çıkartılarak Rule-Notification-Workflow Engine yapısında işlenerek sistem yapısı işletilmesi sağlanacaktır.



Şekil-2 Önerilen Yazılım Mimarisi Üzerindeki Doküman Tabanlı İş Süreçleri

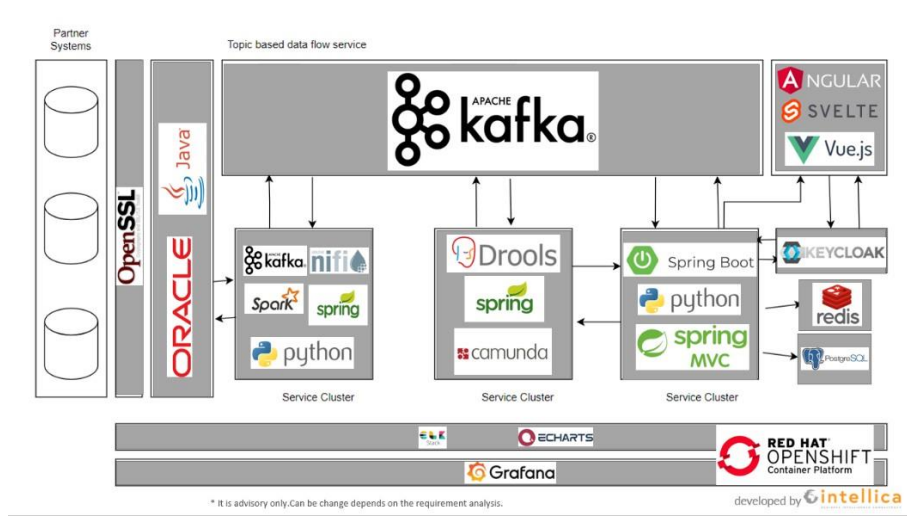
Önerilen yazılım mimari yapısının vermiş olduğu olanaklar (mesajlaşma yapısı, authentication authorization yapısı vb.) kullanılarak Şekil-2’de iş süreçleri doküman tabanlı olarak gerçekleştirilebilecektir. Bu iş süreçlerinden birisinin adımları ise aşağıda örneklenmiştir. Bu iş sürecinin aş adımları şu şekilde özetlenebilir: Identity & access



management: güvenlik ve erişim kontrolü ile rol tabanlı kullanıcı arabirimi ile platform erişimi sağlanmaktadır. Authorization ile birlikte kullanıcı tabalı yetkiler okuma,ekleme, silme, içe aktarma, dışa aktarma, kesme, kopyala/yapıştır olacak şekilde tanımlanır ve yetkilerin yapılandırılması için adresleme linki kullanıcıya sunulur. Main menü: Kullanıcıların erişebilecekleri veri katalogları, girdi/çıktı cevapları, kullanım alanları ve platform fonksiyonlarına erişim sağlanan süreç işletilir. Catalog: Kullanıcıların ortak veri paylaşımı yapabilme olanağını sağlandığı alandır. Bu alanda ortak veriler görüntülenir. Seçilen veri tipine göre iş akışları oluşturulur ve kullanım yetkisine göre veriler üzerinde çalışma imkanı sunulur. Explore Data: Kullanıcının üzerinde çalıştığı veri kataloguna özel keşifsel özetlemeleri görebileceği alandır. Registered Data Element: Kayıtlı data elementlerin görüntülediği ve kullanıcının kendi yaptığı çalışmaların versiyonlarının tutulduğu bölümdür. Bu bölüm kullanıcıya geçmiş çalışmaları hakkında bilgi aktarımı yapmaktadır. Use, Export, Enrich Data: Bu bölümde tutarlılık kontrolleri oluşturmak ve eşikleri ve uyarıları tanımlamak için geliştirmeler yapılmıştır. Tutarlılık kontrolleri, işler ve kural kombinasyonları tarafından oluşturulur. Oluşturmak için gerekli algoritmaları ve parametreleri içeren işler iş düzenleyici ile bir tutarlılık kontrolü oluşturulmasına olanak sunar. New Request: Kullanıcı üzerinde çalışmak istediği veriler için sorgu yapabilir şekilde tasarlanmıştır. Approval Mechanism: Onay mekanizması ile platformda birden fazla kullanıcının aynı veriler üzerinde bir arada çalışmasına olanak veren yapı geliştirilmiştir. Bu alanda manipüle edilmiş/zenginleştirilmiş veriler çapraz olarak kontrol edilerek komplikasyonlar önlenebilecektir. Manage Operations: Yönetim operasyonları bu alanda işlenir ve kullanıcıların yaptıkları işlemler görüntülenir. Prepare Data To Publish & Data Load (Producer): Veri ön işleme çalışmalarının kapsadığı alandır. Data Validation: Verilerin değerleri üzerinde çalışma fırsatları verir. Sütün türleri, otomatik değiştirme, sütunları otomatik doldurma, bağımlı tarih kontrolleri, otomatik kimlik oluşturma işlemleri bu bölümde yapılabilir. Trace of Data: Verilerin versiyonlarının takibinin yapılabilirdiği alandır.

4. Protatip Uygulama Detayları

Yukarda yazılım mimari yaklaşımı ve örnek iş prosesi tanımlanan yazılımın, hangi teknolojilerle geliştirilebileceği konusunda kapsamlı bir çalışma yapılmıştır. Bu doğrultuda aşağıda resmettiğimiz prototip uygulama detay teknoloji diyagramı verilmektedir.



Şekil-3 Prototip Uygulamada Kullanılan Teknolojiler

5. Sonuçlar ve Gelecekteki Çalışmalar

Bu araştırma kapsamında, bir dağıtık sistem tabanlı veri paylaşım platformu yazılım mimarisi önerisi yapılmaktadır. Yapılan önerinin faydasının ortaya konulabilmesi için bir de prototip uygulama detayları, kullanılacak teknoloji detayları da verilerek anlatılmaktadır. Bu yazılım mimarisi merkezi ortamda konumlanmış dağıtık sistem altyapısı üzerinde kurgulanmıştır. Doküman tabanlı veri paylaşımı gerçekleştirebilen bir yazılım platformu mimarisidir. Yine araştırma kapsamında, bu yazılım mimarisinin prototip uygulamasının nasıl geliştirilebileceğinin detayları özetlenmektedir.

Gelecekteki çalışmalarda, prototip uygulamanın önerilen yazılım mimarisine dayalı geliştirilmesi ve prototip uygulamanın önerilen değerlendirme kriterlerine dayalı olarak, performans değerlendirmesinin yapılması planlanmaktadır.

References

- [1] Khine,P. , Shun Wang.Z. , Data lake: a new ideology in big data era , ITM Web ofConferences 17, 03025 (2018)
- [2] Dong, X., Li, R., He,H. , Zhou,W., Xue,Z., Wu, H., Secure Sensitive Data Sharing on a Big Data Platform, tsinghua science and technology 72-80, 2015.
- [3] Grossman, L., Data Lakes, Clouds, and Commons: A Review of Platforms for Analyzing and Sharing Genomic Data, Trends in Genetics 35(3), 2019.



- [4] Juan, P., Juan, C. Corchuelo, R., On exploring data lakes by finding compact, isolated cluster, Information Science pp 103-127, 2022
- [5] Eichler, R., Giebler, C., Gröger, C., Schwarza, H., Mitschanga, B., Modeling metadata in data lakes—A generic model Volume 136, 2021
- [6] Eichler, R., Gröger, C., Hoss, E., Schwarz, H., Data Shopping — How an Enterprise Data Marketplace Supports Data Democratization in Companies, Intelligence Information System pp 19-36, 2022
- [7] Cruchten, R., Weigand, H., Towards Event Log Management for Process Mining - Vision and Research Challenges Research Challenge Information Science pp 197-213, 2022
- [8] Eichler, R., Gröger, C., Hoss, E., Schwarz, H., From Data Asset to Data Product – The Role of the Data Provider in the Enterprise Data Marketplace, Symposium and Summer School on Service-Oriented Computing pp 119-138, 2022
- [9] Piantella, D., A Research on Data Lakes and their Integration Challenges, The 30th Italian Symposium on Advanced Database Systems, 2022