



Bibliyometrik Analizde Dijital Çözüm : BiBLoX

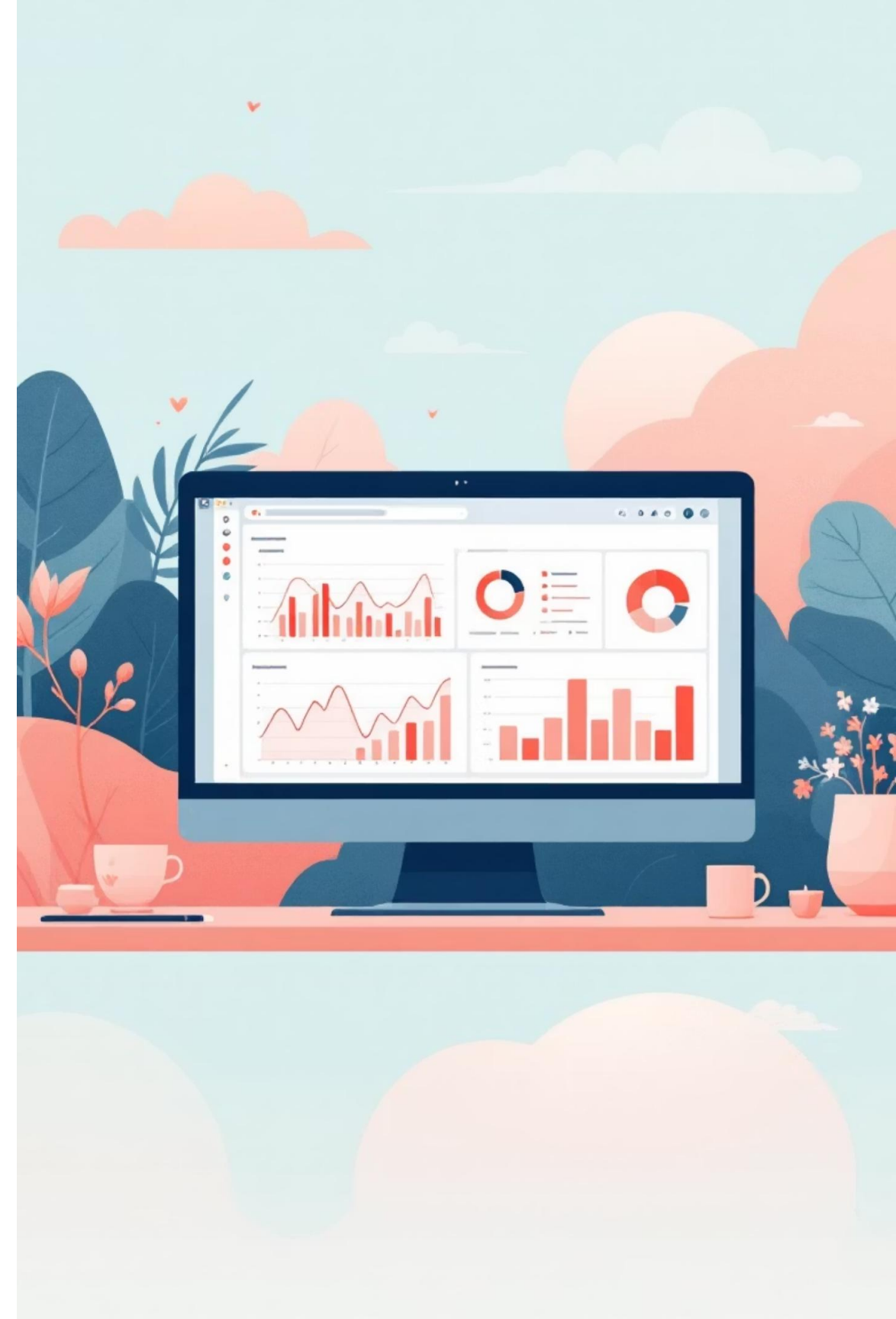
Bu çalışma, fen eğitimi alanında bibliyometrik analiz yapmak için geliştirilen BiBLoX yazılım programının tasarım ve geliştirme süreçlerini detaylı olarak ele almaktadır. Yazılım, Web of Science (WoS), Scopus ve TRDizin gibi akademik veri tabanlarından otomatik veri çekimi yaparak, kapsamlı bibliyometrik analizler gerçekleştiren yenilikçi bir platformdur.

Öğr. Gör. Dr. Kadir KESGİN

Bandırma On Yedi Eylül Üniversitesi
Bilişim Güvenliği Teknolojisi Programı
kadir@bandirma.edu.tr

Danışman: Doç. Dr. Dilek ZEREN ÖZER

BUÜ Eğitim Fakültesi Fen Bilgisi Eğitimi ABD.
dzeren@uludag.edu.tr



Yazılım Geliştirme Yaşam Döngüsü (SDLC) Yaklaşımı

01

Planlama ve Gereksinim Analizi

Proje hedefleri ve kapsamı belirlendi. WoS, Scopus ve TRDizin veri kaynakları seçildi. Python, Flask, MySQL teknolojileri tercih edildi.

02

Sistem Analizi ve Tasarım

Veri akışı diyagramları oluşturuldu. Modüler yapı tasarlandı. Flask tabanlı web arayüzü planlandı. MySQL veri tabanı mimarisi geliştirildi.

03

Geliştirme Süreci

Python ile temel kodlar yazıldı. API entegrasyonları geliştirildi. Pandas ve NumPy ile veri temizleme süreçleri tasarlandı.

04

Test ve Optimizasyon

API bağlantıları test edildi. Büyük veri kümeleriyle performans testleri yapıldı. Kullanıcı arayüzü erişilebilirlik testleri gerçekleştirildi.

05

Dağıtım ve Bakım

Beta versiyonu test kullanıcılarına açıldı. Kullanıcı geri bildirimleri toplandı. Sürekli iyileştirme planları oluşturuldu.

Sistem Giriş ve Güvenlik Özellikleri

Güvenli Kullanıcı Erişimi

Sistem giriş ekranı kullanıcı adı ve parola bilgileri ile erişim sağlamaktadır. Bu ekran, kullanıcıların yetkilendirilmiş verilere ulaşmasını sağlamak ve güvenliği artırmak amacıyla tasarlanmıştır.

Güçlü bir kimlik doğrulama mekanizması, yetkisiz erişimleri engellemek ve kullanıcıların kişisel verilerini korumak adına kritik bir bileşendir. Modern sistemlerde iki faktörlü kimlik doğrulama ve şifreleme teknikleri sıklıkla kullanılmaktadır.

Akademik veri tabanlarında yer alan bilgiler, yalnızca yetkilendirilmiş araştırmacılar tarafından erişilebilir hale getirildiğinde, bilimsel çalışmaların güvenilirliği ve akademik etik ilkeler korunmuş olur.

Giriş Yap

Kullanıcı Adı:

Parola:

Giriş Yap

Hesabınız yok mu? [Kayıt Olun](#)

Arama teriminizi girin

Başlangıç Yılı (Opsiyonel)



Bitiş Yılı (Opsiyonel)



☐ Arama teriminin İngilizcesini de ara

Arama ve Veri Çekme Süreçleri

Sistem kullanıcı tarafından girilen arama terimlerini temel alarak akademik veri tabanlarında geniş kapsamlı bir tarama gerçekleştirmektedir. Bu süreçte, farklı veri tabanlarından (WoS, Scopus ve TRDizin) gelen veriler birleştirilerek analiz edilmektedir.

1

Anahtar Kelime Girişi

Kullanıcı arama terimlerini sisteme girer

2

API Sorguları

Sistem otomatik olarak veri tabanlarını tarar

3

Veri Birleştirme

Farklı kaynaklardan gelen veriler birleştirilir

4

Analiz Hazırlığı

Veriler bibliyometrik analiz için hazırlanır

Arama sürecinde kullanılan algoritmalar, yalnızca anahtar kelimeler üzerinden değil, aynı zamanda eşleşen başlık, özet ve atıf verileri üzerinden de ilgili akademik çalışmaları belirlemektedir.

Teknik Altyapı ve Kullanılan Teknolojiler



Python Programlama Dili

Yazılımın temel kodları Python ile yazılmıştır. API entegrasyonları, veri işleme ve analiz süreçleri Python'un güçlü kütüphaneleri ile gerçekleştirilmiştir.



Flask Web Framework

Kullanıcı dostu web arayüzü Flask kullanılarak geliştirilmiştir. Dinamik web platformu HTML, CSS ve Flask entegrasyonu ile oluşturulmuştur.



MySQL Veri Tabanı

Büyük miktarda veriyi etkili bir şekilde saklama ve yönetme kapasitesi nedeniyle MySQL tercih edilmiştir. Kullanıcıya özel aramalar kaydedilmektedir.



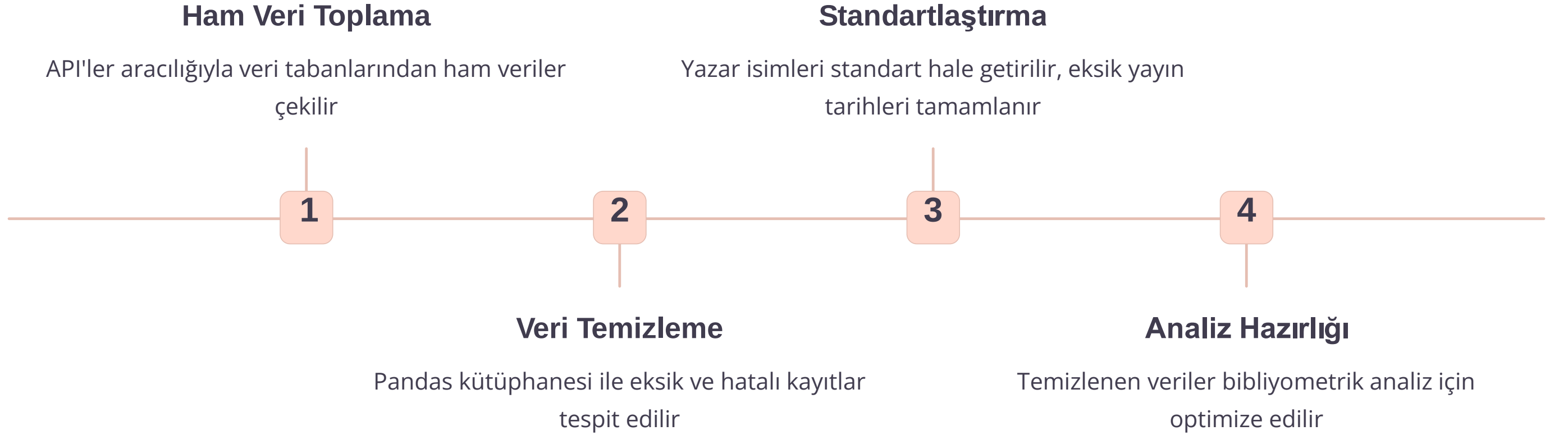
Pandas ve NumPy

Veri temizleme süreçleri Pandas ve NumPy kütüphaneleri kullanılarak tasarlanmıştır. Büyük veri kümeleri üzerinde esnek ve etkili işlemler yapılmaktadır.



Veri İşleme ve Temizleme Süreçleri

Veri işleme sürecindeki en önemli aşamalardan biri, veri temizleme ve eksik verilerin doldurulmasıdır. Veri analizi sürecinde kullanılan kaynaklardan elde edilen ham veriler çoğunlukla hatalı, eksik veya tutarsız kayıtlar içerir.



Özellikle birden fazla veri kaynağından gelen bilgilerin birleştirilmesi aşamasında, farklı formatlar, eksik alanlar ya da tekrarlayan girdiler söz konusu olabilir. Bu nedenle, ham verilerin doğrudan kullanılması yerine, verilerin temizlenmesi ve eksik kısımların uygun yöntemlerle tamamlanması, elde edilen sonuçların daha güvenilir, tutarlı ve anlamlı olmasını sağlar.

Makine Öğrenmesi Algoritmaları ve Tahmin Modelleri



Doğrusal Regresyon (LR)

Yayın eğilimlerinin zamansal olarak doğrusal bir modelle ifade edilebileceği varsayımıyla tercih edilmiştir. Basit yapısı ve hızlı hesaplama yeteneğiyle büyük veri setleri üzerinde etkili sonuçlar sunmaktadır.



Destek Vektör Regresyonu (SVR)

Doğrusal olmayan veri setleri üzerinde güçlü performansı nedeniyle tercih edilmiştir. Fen eğitimi alanındaki yayın ve atıf verileri, doğrusal olmayan eğilimler içerebildiği için SVR'nin karmaşık ilişkileri modelleme kapasitesi önemli avantaj sağlamıştır.



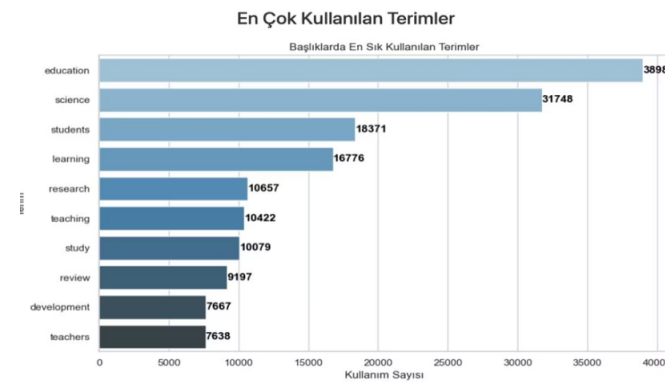
Karar Ağaçları Regresyonu (DTR)

Veri setlerindeki belirli kategorik faktörlerin etkisini anlamada etkili bir yöntemdir. Fen eğitimi yayınlarının bölgelere, dillere veya konu başlıklarına göre farklılıklarını modelleyerek değerli öngörüler sunmuştur.

Bu algoritmalar kullanılarak, fen eğitimi literatüründeki yıllık yayın eğilimleri ve atıf yoğunlukları tahmin edilmiştir. Makine öğrenmesi modelleri, bibliyometrik analizlerden elde edilen bilgileri geleceğe yönelik tahminler yapmak için kullanarak yazılımın işlevselliğini bir adım öteye taşımaktadır.

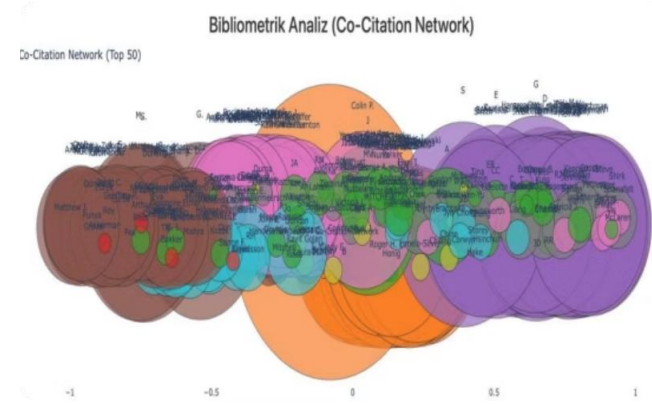
Veri Görselleştirme ve Analiz Sonuçları

Verilerin görselleştirilmesi aşaması, bibliyometrik analizlerin anlaşılabilir ve etkili bir şekilde sunulabilmesi için büyük önem taşımaktadır. Plotly ve Seaborn gibi güçlü görselleştirme kütüphaneleri kullanılarak hem kullanıcı dostu hem de etkileşimli görseller oluşturulmuştur.



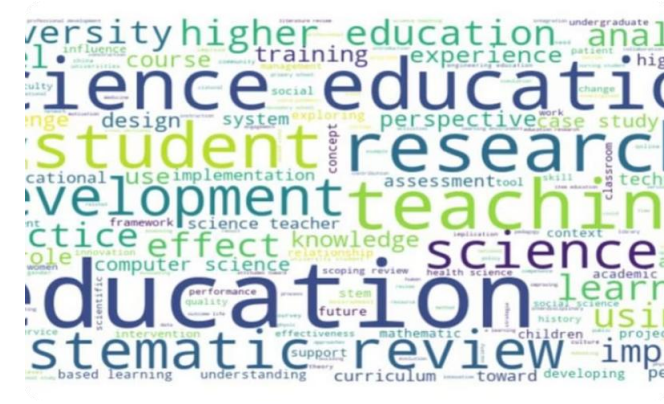
En Çok Kullanılan Terimler

Anahtar kelime analizi ile literatürde en sık kullanılan terimlerin görselleştirilmesi



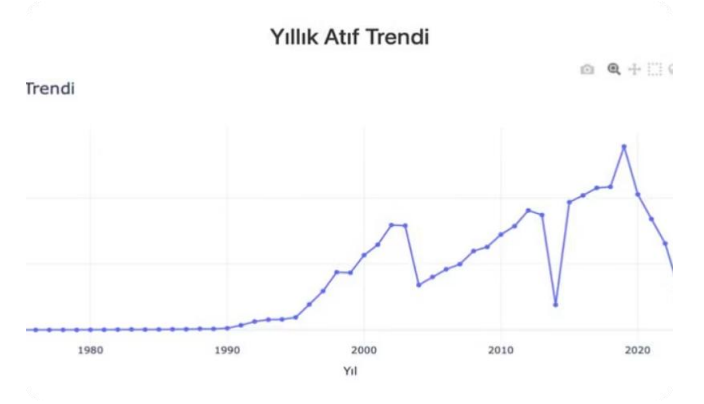
Ortak Yayın Ağları

Yazarlar arası işbirliklerinin ağ analizi ile görselleştirilmesi



Kelime Bulutu

Anahtar kelimelerin frekanslarına göre boyutlandırılmış görselleştirilmesi



Yıllık Atıf Trendleri

Zaman içerisindeki atıf yoğunluklarının değişiminin grafiksel analizi

Bu görselleştirmeler, yazar işbirlikleri, atıf trendleri ve yayın yoğunluğu gibi verilerin zamansal ve mekansal olarak analiz edilmesini kolaylaştırmıştır. Ülkeler bazında yayın yoğunluk haritaları, bilimsel çalışmaların coğrafi dağılımını anlamak için etkili bir araç olarak kullanılmaktadır.

Sistem Doğrulama ve Gelecek Perspektifleri

BiBLoX yazılımının doğruluğu, aynı sorguların farklı veri tabanlarında yapılması ve sonuçların karşılaştırılması ile test edilmiştir. "Distance education on science education" teması sorgulandığında:

2681

WoS Sonuçları

Web of Science veri tabanından dönen sonuç sayısı

4960

Scopus Sonuçları

Scopus veri tabanından dönen sonuç sayısı

100%

Doğruluk Oranı

BiBLoX ile orijinal veri tabanları arasındaki uyum

Geliştirilen bu yazılım programı ve kullanılan kütüphaneler, Python programlama dili ile açık kaynak olarak mevcuttur. Flask, Pandas, MySQL Connector, Scikit-Learn ve Plotly gibi kullanılan tüm araç ve kütüphaneler, ilgili resmi sitelerinden veya Python Paket Yöneticisi (PyPI) aracılığıyla indirilebilir.

 **Gelecek Geliştirmeler:** Yazılımın gelecekte daha fazla veri kaynağı eklenerek genişletilebilmesi için ölçeklenebilir bir altyapıya sahip olması sağlanmıştır. Bu sayede bilimsel çalışmalarda daha çok veri odaklı kararlar alınması mümkün olmaktadır.