

Büyük Veri ve Büyük Veri Yaşam Döngüleri

Sadi Evren SEKER

Istanbul Medeniyet University, Department of Business

Özet

Bu yazının amacı, günümüzde hızla gelişmekte olan büyük veri kavramına genel bir giriş yapmak ve ardından büyük veri dünyasında var olan teknoloji firmalarının bu dünya için geliştirmiş olduğu farklı veri modelleri ve yaşam döngüsü modellerini karşılaştırarak okuyucuya genel bilgileri sunabilmektir. Yazı kapsamında toplanan kaynaklar akademik kaynakların yanı sıra büyük veri alanında faaliyet gösteren Oracle, SAS, ve EMC gibi firmaların kendi yayınlarından derlenmiştir. Ne yazık ki bu konuda yeterli Türkçe yayın bulunamamasından dolayı bu kaynaklarda bulunan görsellere atıf verilerek Türkçelenmiş ve Türkçe olarak yeniden üretilen görsellere makalede yer verilmiştir. Makale bundan sonra yayınlanacak olan yazılım proje yönetim teknikleri ile devam edecek olan serinin ilk makalesidir ve bu seri kapsamında son olarak büyük veri dünyasında yeni proje yönetim tekniklerine yer verilecektir.

Anahtar Kelimeler: Büyük Veri, Veri Bilimi, Yaşam Döngüleri

Summary

Aim of this paper is introducing the big data concept and comparing some life cycle models from the big data industry leaders. Besides the academic literature, this paper also refers to some documentation from industry like Oracle, SAS or EMC. Unfortunately, Turkish literature on the topic is limited. Because of the limited Turkish documents on the literature, the life cycle models in this paper are both referring to the original sources and also recreated in Turkish. This paper is the first paper of a three paper series ,which will continue with a general introduction to software project management techniques and finally on the third paper the new trends and approaches to the project management in a big data world will be discussed.

Keywords: Big Data, Data Science, Life Cycles

1. Büyük Veri

Bilgisayarların işleyemeyeceği kadar büyük olan verilerdir. Bu veri sürekli büyümektedir. Bunun daha güzel bir tanımı 5V olarak yapılan tanımdır. 5V volume, velocity, variety, veracity ve value'dir. Büyük veri denilince aslında akla ilk gelen boyut olarak büyük (volume), yani kapladığı yer olarak büyük verilerdir. İkinci V, Velocity yani hızdır. Örneğin facebook'ta her gün çok sayıda yeni mesaj yazılmakta, bunların işlenebilmesi için belirli bir işlem hızının yakalanması gerekir

Ayrıca verinin çeşitliliğinden de bahsedilebilir, yani 5V kavramının üçüncüsü olan variety'den. Kısaca değişik kaynaklarda değişik içeriklerden değişik özelliklere sahip veriler gelmektedir. Dolayısıyla büyük verinin hacmi, hızı ve kaynak çeşitliliği vardır. Dördüncü olarak verinin güvenilirliği (veracity) yani hangi kaynağına ne kadar itimat edeceğiniz? Sorusu büyük veri dünyası için önemlidir.. Bütün bunların neticesi olarak, değer (value) beşinci V olarak görülebilir... Yani bir verinin ne kadar büyük olduğu, bir verinin ne kadar hızlı değiştiği, ne kadar

çeşitli olduğu ve ne kadar güvenilir olduğuna göre bir değerden bahsedilebilir. İşte büyük veride bütün bunlar birer problem olarak karşımıza çıkar. Büyük veri düşünüldüğünden daha büyüktür. Hacim olarak büyüktür, hız olarak büyüktür, çeşitlidir ve güvenle ilgili kaynaklar ile ilgili netlik yoktur veya bu konularda sorunlar vardır denilebilir. Dolayısıyla verinin değeri bizim için çok yüksektir. Ama bununla birlikte değerini düşüren unsurlar da vardır. Şayet bunları düzgün bir şekilde yönetebilirsek verimiz gerçekten değerli hale gelecektir.

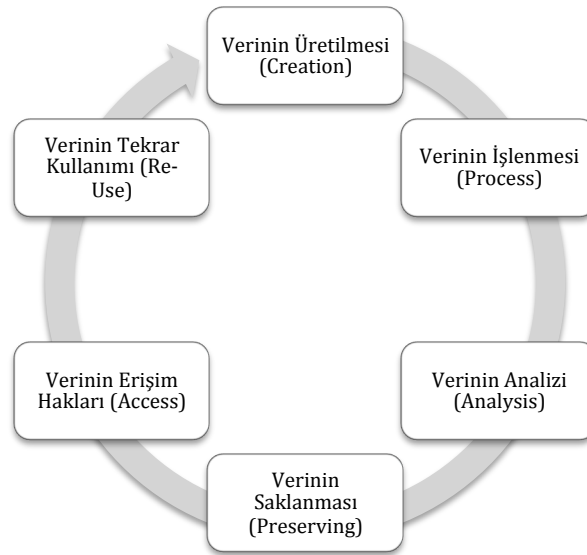
Büyük veri kavramı ve büyük veri çalışmaları, veriyi daha değerli hale getirmek için belirttiğimiz ilk 4V problemlerinin üzerine giden ve buradaki problemleri çözmeye çalışan çalışmalardır denilebilir. Büyük veri çalışmalarının derdi, çok büyük ölçekteki verileri, çok hızlı değişen, çok çeşitli kaynaklardan gelen ve çok farklı itimat değerlerine sahip olan verileri yönetebilmektir.

2. Büyük Veri Yaşam Döngüsü Modelleri

Büyük verinin tanımının ardından büyük verinin yaşam döngüsünden bahsedilebilir. Her varlık yaşar, her varlığın bir hayatı olduğu söylenebilir, bilgisayar dünyasındaki kavramlar da yaşarlar. Yani hiçbir varlığın tek başına bir kere var olduğunu ve sonsuza kadar var olacağını söylenemez. Dolayısıyla kavramlar da değişmektedir, değerler değişmektedir. Büyük veri değeri de değişmektedir ve büyük verinin de bir yaşam döngüsü vardır. Büyük veri her ne olursa olsun, örneğin; facebook'ta toplanan veriler, gmail de atılan maillerin tamamı, bir ülkedeki vatandaşların nüfus kayıtları veya aklınıza gelen herhangi bir büyük veriyi ele alın, aslında bir yaşam döngüsü vardır (SEKER, 2014). Bu yazı kapsamında, farklı firmalardan (SAS, IBM, Oracle gibi) bu yaklaşımları toplanmıştır. Büyük verinin hayat döngüsü ile ilgili. bu yaklaşımlar sunulacaktır.

2.1. Verinin Yaşam Döngüsü (Data Life Cycle)

İlk anlatılacak olan yaklaşım akademik bir yaklaşım ve Essex Üniversitesi tarafından literatüre kazandırılmış (Corti, Eynden, Bishop, & Woollard, 2014).



Şekil 1 Verinin Yaşam Döngüsü (Data Life Cycle)

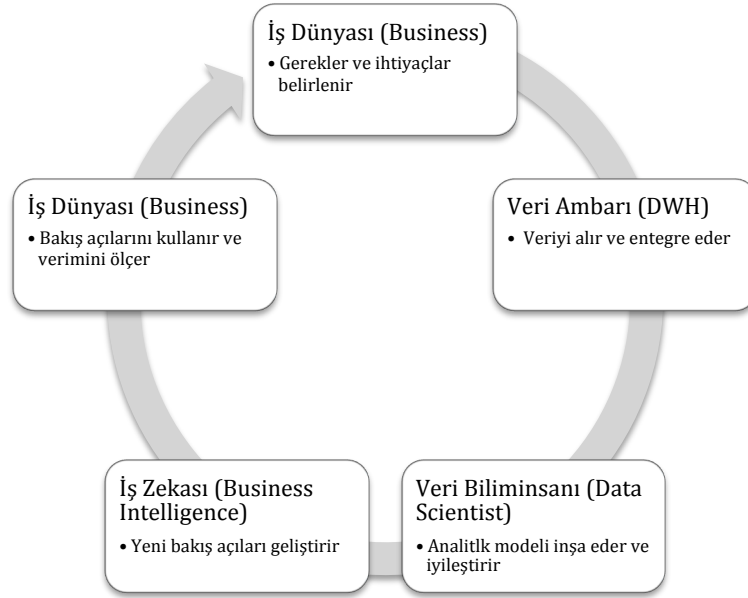
Bu yaklaşıma göre, verinin bir yaşam döngüsü var. Üretilmesi, işlenmesi, analiz edilmesi, bu verinin bozulmalara karşı saklanması, güvenlik açıklarına karşı saklanması veya bu veriye kimin erişeceği ve verinin tekrar kullanılması gibi giden bir döngüden bahsedilebilir. Bu aslında verinin yaşam döngüsü olarak görülebilir. Kısaca, verinin üretilmesi aşamasında (Creating) aslında ihtiyaçların belirlenmesi söz gerekir, veri üretiminin nasıl yönetileceği, verinin hangi kaynaklardan geleceği esas problemlerdir. Örneğin; bir satış zinciriniz var, marketler zinciriniz var; dükkanlarınıza giren çıkan müşteriler, yazar kasaların takip edilmesi gibi çeşitli kaynaklardan verileriniz geliyor. Öncelikle bu verilerin toplanması gerekiyor. Verinin işlenme aşamasında (process) ise bu veriler ön işlemeden geçirilir, anlamsız olan verilerin temizlenmesi, eksik olan verilerin toplanması, başka kaynaklardan beslenmesi, verinin başka kaynaklarda işlenebilecek şekle getirilmesi gibi aşamalar vardır. İşte bu ilk iki aşama verinin toplanması veya üretilmesidir. Bazen hesaplamalarla bizim ürettiğimiz veriler de olabilir bazen de doğrudan sahadan toplanan veriler olabilir. Verinin işlenmesi, temizlenmesinden, ön işlemeden sonra verinin analiz aşaması başlar. Bu aşama diğerlerine göre çoğu veri bilimcisi için en zevkli aşamadır. Burada sürprizler, yeni fikirler çıkar, bir nevi veriyle oyun oynanır. Mesela elinizde facebook'un verisinin olduğunu düşünün. Bu veriyi topluyorsunuz ve ön işlemeden geçirip temizlediğinizi düşünün, mesela Türkiye'deki insanlar en fazla hangi ülkeye yurtdışına mesaj atıyor. Veya diyelim ki en fazla mesajlaşma hangi ülkede gerçekleşiyor. Kaç tane sahte hesap var. Açılan sahte hesapların amaçları neler? Herhangi bir kanal, bir haber kanalı, bir resim paylaşım grubu, ile ilgili mesajlaşan kişilerin cinsiyet dağılımı nasıl yaş dağılımı nasıl demografik yapıları nasıl gibi bir sürü fikir ortaya atılabilir. Bu fikirlerin işlenmesi, çıkarılması mümkündür (SEKER, Sosyal Ağlarda Veri Madenciliği (Data Mining on Social Networks), 2015). Bu fikirler sonunda çıkan sonuçlar ise sürpriz olabilir mesela hiç beklenmeyen veya beklentilerin çok ötesinde sonuçlar çıkabilir.

Sonraki aşama bu analiz edilen verinin korunması aşamasıdır. Yani bu verinin yedeklenmesi hangi yapının, hangi formatın verinin saklanması için uygun olduğu gibi kararların verildiği aşamadır. Paylaşılması, arşivlenmesi hatta bu verinin güncel halde tutulması bile bir problemdir. Burada sürekli yenilenen bir veri söz konusu ve dolayısıyla bu verinin de bir şekilde güncellenmesi gerekir. Sonraki aşamalarda, da bu verinin erişim haklarının belirlenmesi, güvenlik ile ilgili kontrollerin yapılması. Bu veri erişim haklarının paylaşılması, dağıtılması, yasal çerçeveler içinde olup olmadığı, erişim izlenmesi kimler erişiyor, ne sıklıkla erişiyor, nerelere erişiyor, kimin ihtiyaç duyduğu bu veriye gibi problemlerle uğraşılıyor gibi soruların cevaplanması gerekir. Ve sonraki adımda da verinin tekrar kullanılması söz konusu, yani veriyi çıkarttık, ön işledik, analiz ettik ve sonuçlara ulaştık, peki bu veriyi çöpe mi atacağız? En azından veri değilse bile yöntemleri kullanamaz mıyız? Mesela bir firma için yaptığınız bir işi tamamen çöpe mi atacağız? Başka firmaların benzer yöntemleri kullanması söz konusu olamaz mı? En azından metodolojinin başka yerlerde kullanılması, hatta bazı durumlarda verinin bizzat başka yerlerde kullanılması mümkündür. Dolayısıyla verinin tekrar kullanılması büyük veri hayat döngüsü açısından önemlidir.

Aslında bahsedilen şey yaşam döngüsü hayatın her alanında görülebilir. Herhangi bir varlığı ele alırsanız onun da bir hayatı vardır. Belli aşamalardan geçer ve bu hayat tekrar eden bir yapıya sahiptir. Örneğin, yağmur yağıyor, yağın yağmur buharlaşıyor tekrar bulut oluyor ve tekrar yağıyor. Böyle bir döngü söz konusudur. Bu hayatın farklı aşamalarında karşımıza çıkıyor. Yazılımda hayatın bir parçası ve yazılımın da kendine özgü bir hayatı var. Dolayısıyla verinin de buna benzer bir şekilde hayatı olduğundan söz edebiliriz.

2.2. Büyük Veri Yaşam Döngüsü

İkinci olarak EMC firması tarafından hazırlanan Büyük Veri Yaşam Döngüsü (Big Data Life Cycle) olarak geçen kavramı inceleyeceğiz (Schmarzo, 2012).

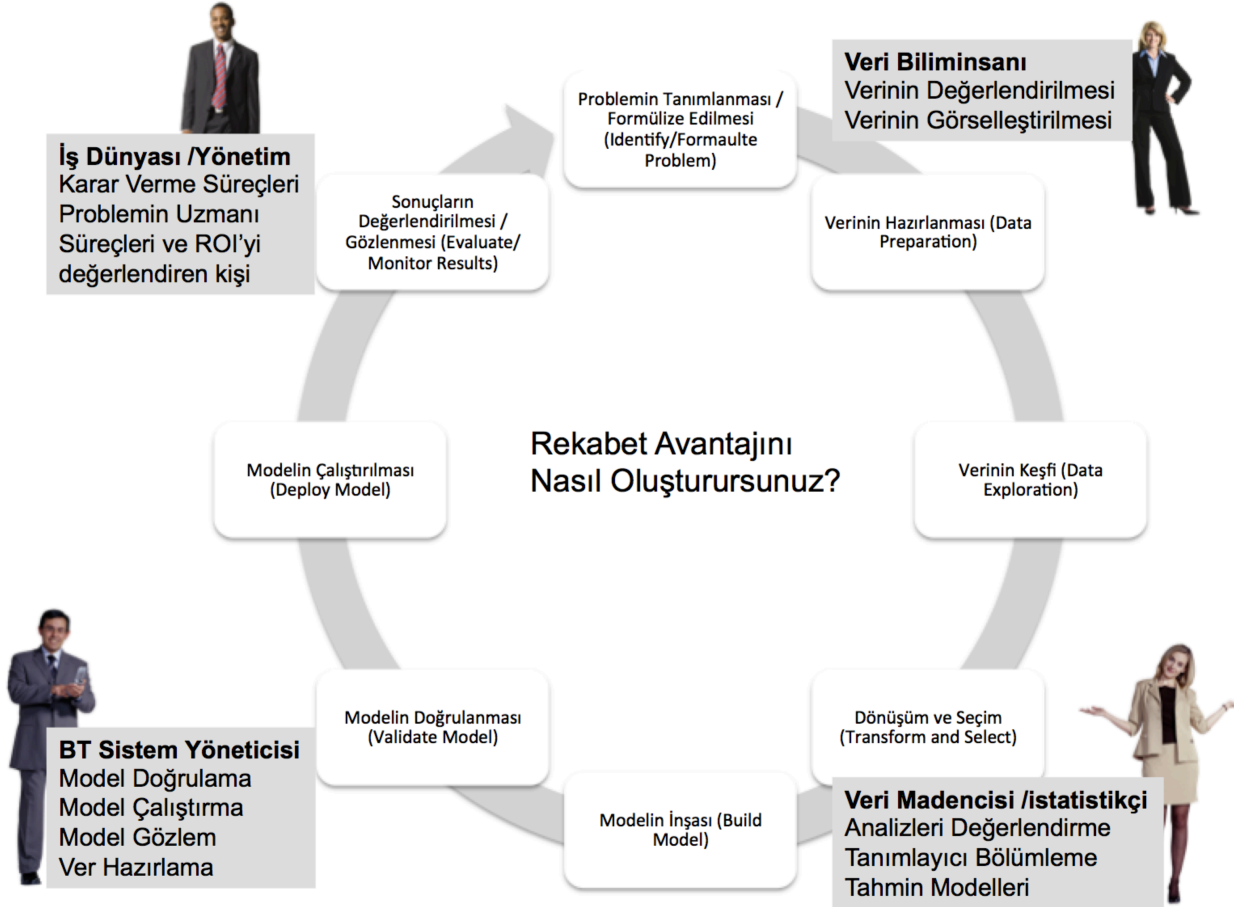


Şekil 2 EMC Firması tarafından geliştirilen Büyük Veri Yaşam Döngüsü

Big data life cycle kavramı henüz yeni ve oturmamış bir kavramdır. Birçok kişinin kullandığı bu terim, çok farklı anlamlara gelebilmektedir. Yeni modelde, iş dünyasıyla veri ambarları ve iş zekâsı süreçlerinin bir yaşam döngüsünde bağlanmaktadır. Öncelikle bir iş süreci var ve iş dünyasının belli ihtiyaçları ve sorguları bu modelde ele alınmış. İkinci adımda veri ambarı (Data Warehouse) süreci işin içine girmektedir. Burada görev yapan veri biliminsanından bahsedilebilir. Veri biliminsanı, veri ambarının çıkartmış olduğu farklı veri kaynaklarından topladığı, ön işlendiği ve analiz ettiği sonuçları alır ve bunların üzerinde kendine göre analiz/analitik modelleri geliştirir. Bu aşamada, veri madenciliği algoritmalarından bahsedilebilir. Belirli tekrarlar, belirli şablonlar yakalanmaya çalışılır. Sonraki adımda bir iş zekâsı süreci başlar. İş zekâsı sürecinde yeni fikirlerin ortaya çıkması çok önemlidir. Yani veri bilimcisi açısından baktığınızda, iş zekâsında çıkanlar sadece rapordan ibarettir ama iş dünyası açısından baktığınızda o raporlar aslında iş dünyasına şekil veren iş dünyasını bazen komple değiştiren yeni bir sürece sürükleyen yeni fikirler getiren bir yapıya sahiptir. Ardından bu çıktıların iş dünyasındaki etkilerinin tekrar ölçülmesi, verinin anlaşılması, kararları etkilemesi, stratejik kararları yönetim sürecini etkilemesi, operasyonel anlamda çalışanların verimliliğini arttırması, kalitenin arttırılması gibi çok farklı yansımaları olur ve yeni ihtiyaçlar ortaya çıkar. Bu yeni ihtiyaçlar yeni veri işleme süreçlerini başlatır. Böylelikle, yeni büyük veri yaşam döngüleri başlar ve büyük veri yaşam döngüsü bu şekilde dönmeye devam eder. EMC firması tarafından hazırlanan bu model, daha çok iş dünyasını, veri işleme dünyasını ve veri bilimi dünyalarını birleştiren bir döngü. Bu yüzden önemlidir denilebilir.

2.3. Analiz Yaşam Döngüsü

Veri madenciliği, ve istatistik ile ilgili yazılımları ile tanınan SAS firması tarafından hazırlanan “The Analytics Life Cycle” olarak geçen verinin işlenmesi ile ilgili aşamaları içeren bir modelden de bahsedilebilir (SAS, 2013).



Şekil 3 Analiz Yaşam Döngüsü

Bu modeli incelemeye öncelikle kullanıcıları ve rollerimiz inceleyerek başlayalım.

İş dünyası olarak bir yönetici veya bir şirket sahibi, bir patron düşünülebilir. Bu kişinin kendi çalışma alanı, uygulama alanı ile ilgili uzmanlığı bulunmaktadır. Karar verme süreçlerinde sorumluluğu ve rolü vardır. Aynı zamanda yatırımın ne kadar geri dönüşü olacağı hangi konularda yatırım yapılması gerektiği ve bunun şirkete katkıları gibi değişik aşamalarda analizler yapmak isteyen buna göre de karar vermek isteyen bir yönetici olarak düşünülebilir. Bu yöneticinin doğru karar verebilmesi için doğru bilgilere ulaşması gerekir.

Veri biliminsanı (data scientist) verinin araştırılması verinin görselleştirilmesi gibi veriyle uğraşan ama daha çok hazır araçları kullanan kişidir. Üçüncü olarak data miner yani istatistikçi rolünden bahsedilebilir. Veri ile ilgili değişik keşifler yapan, verilerin sınıflandırılması segmentlere , kategorilere ayrılması , tahmin modellerinin geliştirilmesi gibi klasik bazı veri madenciliği uygulamaları ile uğraşan roldür. Genelde iyi derecede istatistik bilmesi beklenir. İstatistik bilgisi yeterli olmasa bile veri madenciliği yapılabilir. Hazır çok sayıda araç kullanılarak, konu hakkında hiçbir fikre sahip olmadan çoğu sonuca ulaşılabilir. Ama sonuçların ne kadar doğru olduğu, ulaşılan sonucun doğru yorumlanması için başlangıç seviyesinde de olsa istatistik bilinmesinde fayda vardır. Çünkü istatistiksel modeller aslında istenen her sonucu verir. Ama o sonucun doğru olup olmadığı o sonucun ne anlama

geldiğini yorumlamak bu konularda az da olsa tecrübe gerektirir. Belki veri bilimi rolündeki kişi bu konularda bilgi sahibiyse ve elindeki araçları doğru şekilde kullanıyorsa, o zaman bir istatistikçiye ihtiyaç olmayabilir. Veya veri biliminsanı bu role de bürünmüş olabilir. Modeli geliştiren SAS şirketi açısından probleme bakacak olursak hazır araçları kullanan birisi belki bu aşamaları halledebilir ama SAS'ın araçları için bile çoğu durumda bazı modüllerin yeniden yazılması gerekebilmekte, dolayısıyla bu modüllerle ilgilenen bir rol daha modelde bulunmaktadır. Ayrıca modeldeki son rol bilişim teknolojileri uzmanıdır (IT Systems Management). Bu kişinin model doğrulaması (validation), modelin sistem üzerinde sorunsuz çalışması ile ilgilenmesi ve gözlemlemesi, çıkan problemleri çözmesi, verilerin ön hazırlanması, örneğin gerekli durumlarda veritabanından ön veriler çekmesi, farklı veritabanlarına bağlantı gerekmesi durumunda network üzerinden güvenli bağlantı kanalları kurması gibi bir dizi yaşam döngüsü için kritik öneme sahip sorumlulukları bulunmaktadır.

Bu dört temel rol, aslında şirketlere rekabet avantajı sağlamaya çalışır (Sadi Evren Seker, Cankir, & Okur, 2014). Temel soru şudur: “ bir şirket veriyi neden işlemek ister? Neden daha değerli veriye ulaşmak ister? “ ve sorunun cevabı : “çünkü ondan bir değer elde etmek, bir sonuç, bir kar elde etmek ister”. Bu da aslında en basit anlamda şirketlere rekabet avantajı getirir.

Şirket şayet verisini iyi bilirse iyi işlerse diğer şirketlere göre bir adım önde olabilir. Büyük veri yaşam döngülerinin temeldeki en büyük avantajı budur. Uygulama alanı olarak pazarlama alanında görülebilir. Müşterilere daha iyi, daha tercih edilebilir ürünleri, daha doğru zamanlarda, daha doğru şekilde sunana bir işletme, bunu yapamayan işletmelere göre rekabet avantajı elde eder. Günümüzde büyük verinin en çok kullanıldığı alan pazarlama alanıdır.

Şimdi SAS firmasının sunduğu model olan “Analytics Life Cycle” döngüsünü incelemeye geçelim. İlk aşama da problemlerin şekillenmesi, tanımlanması ve yöneticinin ne istediğinin anlaşılmasıdır. Daha sonra verinin hazırlanması yani verinin ön hazırlanması olarak düşünülebilir. Verinin keşfedilmesi nerelerden bu verinin alınacağı, kaynaklarının belirlenmesi alınmış şekilleri bağlantı protokollerinin belirlenmesi, verinin dönüştürülmesi (transform) ve seçilmesi, işlenmesi (build) modelin inşa edilmesi nasıl bir model istiyoruz gibi sorular, veri madencisinin (data miner) ve veri biliminsanının (data scientist) devreye girdiği aşamalardır. Veri dönüştürüldükten (transform) sonra veri biliminsanı (data scientist) verilerin alınması ve veri madencisi (data miner) modelin inşası ile uğraşır Yani çıkan verinin isteklere uygun olup olmadığı, veride hatalar olup olmadığı bu aşamada incelenir. Elbette bilgi teknolojileri yöneticisinin (IT Manager), model doğrulaması (model validation) söz konusudur. Yani istenilen verilerin tamamına erişebilmiş mi? Veri erişiminde problem var mı? Sistem kayıtlarında hatalar olmuş mu? gibi sorularla uğraşma ve bu modelin çalıştırılmasına (deployment) ve gerçek hayata geçirilmese yönelik problemleri kontrol etmek bilişim teknolojileri yöneticisinin (IT Manager) sorumluluk alanıdır. Bir yazılım veya bir modülün çalıştırılacak hale getirilmesi ve çalıştırılmasından sorumludur. Daha sonraki aşamalarda da değerlendirilmesi (evaluate) ve sonuçların anlaşılması, işlenmesi ile iş süreçlerindeki yöneticiye (business manager) süreç tekrar dönmektedir. Yönetici, döngünün sonunda, kontrol eder ve çıktıları tutarsız bulabilir veya çıktılarına güvenerek karar süreçlerine dahil edebilir. Her durumda döngünün tekrar dönmesi ve yeni kararlara göre yeniden yapılanması veya sistemdeki hataların bulunması ve yeniden çalışması beklenmektedir.

2.4. Büyük Veri Mimarisi

Oracle tarafından geliştirilen Büyük Veri Mimarisi (Big Data Architecture) ile ilgili sundukları döngüde yenilikten (Innovation) çalışan bir modele (Execution) doğru giden bir boyut vardır (Oracle, 2015). Verinin çalışması aşamasına yaklaşıldıkça (execution) yenilik kavramını yavaş yavaş yitirmeye başlar. Çalışan verinin bağlanması ise ayrı bir aşama olarak görülebilir (Data Integration). Ayrıca laboratuvarı yeni fikirlerin ortaya çıkması gibi sürecin farklı bir boyutundan bahsedilebilir. Verilerin üzerinde işlem yapıldıkça yeni fikirlerin ortaya

çıkması kaçınılmazdır.. Verinin bağlanması (integration) aşamasında ise yeni fikirlerin dışında farklı kaynaklardan gelen, örneğin sosyal ağlardan, şirket kaynaklarından, haber kanallarında gelen veriler ve akla gelebilecek bütün diğer kaynaklar işlenir. Keşif laboratuvarı (Discovery Lab) kısmında değişik fikirlerin ortaya atıldığı laboratuvar ortamlarından bahsedebiliriz. Bu fikirlerin test edildiği, üretildiği ve veri bağlamasından (data integration) gelen verilerin işlendikten sonra verinin veri platformlarına (data platform) taşınma aşamasına geçilebilir. Verilerin tutulduğu veri rezervuarları (Reservoir), verinin üretildiği fabrikalar (factory) ve veri ambarları (warehouse) bu aşamada devreye girer. Aynı zamanda olaylar ve veri akışlarından bahsedilebilir (Events and streaming) örneğin haber akışları veya borsa ile ilgili finansal izlenimler pek çok farklı veri kaynağından akarak (stream) gelen veriler bu grupta değerlendirilebilir. Ardından bu akan veri kaynaklarının uygulamalara (application) taşınması aşamasına geçilebilir. Bu uygulamalar mesela cep telefonu, tablet veya internet sayfaları üzerinden sunulan raporlar olabilir. Veya analiz araçlarına (analytic tools) taşınması. Mesela basit analiz aracı olarak excel'i düşünebiliriz.

Facility Life Cycle Management isminde bir büyük veri (big data) yapısından daha bahsedilebilir. Burada da benzer bir yapı var. Desing, build, operate, decommission, plan desing şeklinde geçen yani planlama ile başlayan ve planla, tasarla, inşa et yap, operasyona geç ve bunu tekrar değerlendirmeye al şeklinde bir yapıdan bahsedilebilir. Ayrıca döngü içerisinde, programlama, kullanıcı talepleri, operasyon talepleri, kullanım alanları gibi birçok farklı alandan da bahsedilebilir. Döngünün aşamalarına bakılacak olursa, fonksiyonel ve sermaye ile ilgili planlama ve yönetim aşamalarından bahsedilebilir. Projenin maliyeti, projenin inşa edilmesi ve teslim edilmesi süreçleri, proje ile ilgili kaynakların yönetilmesi, projenin içerik ve malzeme yönetimi ile ilgili aşamalarından bahsedilebilir. Bunların alt aşamalarına bakıldığında kaynak planlaması bu kaynakların yönetilmesi, kaynakların maliyet hesapları ve maliyetin tahmin edilmesi ve projenin metriklerinin çıkartılması örneğin projenin süresinin çıkartılması kaç kişi gerektiğinin çıkartılması veya hangi kaynaklardan veri alınacak?, aynı veriye ulaşmak için başka alternatiflerimiz var mı? gibi sorulara cevaplar bulunması sayılabilir. Ayrıca kaynakların kullanımı ile ilgili ilave bir aşamadan bahsedilebilir. Bir otomasyon sisteminin inşa edilmesi ve bilgisayar tarafından kontrol edilen bilgisayar destekli bakım yapan bir bakım yönetiminden bahsedebiliriz. Burada big data'nın uygulama alanı olarak gerçek bir probleme uygulanması, ile ilgili bir yaşam döngüsü (life cycle) hazırlanmıştır. Ama bu yaşam döngülerinin ortak noktası hepsinin aşağı-yukarı aynı algıyı tekrarladığı görülebilir. Sürekli veri üretiliyor, gün içerisinde sürekli veri üretiliyor, bu yazıyı hazırlarken bile veri üretiliyor. Bu verinin işlenmesi, analiz edilmesi, saklanması, kimin erişeceğinin belirlenmesi, saklanabiliyor olması, tekrar tekrar bunun kullanılabilir bir veri haline getirilmesi ve bu verinin üzerinde bazı analizler yapıp sonuçlar çıkartılıp iş dünyasına sunulması, iş dünyasında bunun faydalı bir şekilde kullanılması gibi aşamalardan bahsediyoruz aslında. Bütün modellerde ortak olan bir yaşam döngüsü var diyebiliriz. Farklı firmalar farklı aşamalarda farklı detay seviyelerinde bahsetmiş olabilir ama aslında bahsedilen şey hemen hemen hepsinde aynı.

3. Sonuç

Bu yazı kapsamında, büyük veri nedir? büyük veri yaşam döngüsü nedir? sorularına cevap aradıktan sonra, büyük veri dünyasında proje yönetimi için biz neler yapabiliriz? Bize nasıl alternatifler sunmaktadır? Proje yönetimi ile ilgili neler değişmektedir? sorularına cevap aranmıştır. Makale bundan sonra yayınlanacak olan yazılım proje yönetim teknikleri ile devam edecek olan serinin ilk makalesidir ve bu seri kapsamında son olarak büyük veri dünyasında yeni proje yönetim tekniklerine yer verilecektir.

Kaynaklar

Corti, L., Eynden, V. V., Bishop, L., & Woollard, M. (2014). *Managing and Sharing Research Data , A Guide to Good Practice*. Essex: SAGE.

Oracle. (2015, May). An Enterprise Architect's Guide to Big Data Reference Architecture Overview. *ORACLE ENTERPRISE ARCHITECTURE WHITE PAPER* .

Sadi Evren Seker, Cankir, B., & Okur, M. E. (2014). Strategic Competition of Internet Interfaces for XU30 Quoted Companies. *International Journal of Computer and Communication Engineering* , 3 (6), 464.

SAS. (2013). Manage the Analytical Life Cycle for Continuous Innovation From Data to Decision SAS Institute Inc. *White Paper* .

Schmarzo, B. (2012, Sept). New Roles in the Big Data World. *EMC, Infocus* .

SEKER, S. E. (2014). Sosyal Ağlarda Akan Veri Madenciliği. *YBSAnsiklopedi* , 1 (3), 21-25.

SEKER, S. E. (2015). Sosyal Ağlarda Veri Madenciliği (Data Mining on Social Networks). *YBSAnsiklopedi* , 2 (2), 30-39.