

Arařtırma Verilerinin Yönetimi ve Açık Bilim alıřtayı

11 Eylül 2018, Boğaziçi Üniversitesi, İstanbul



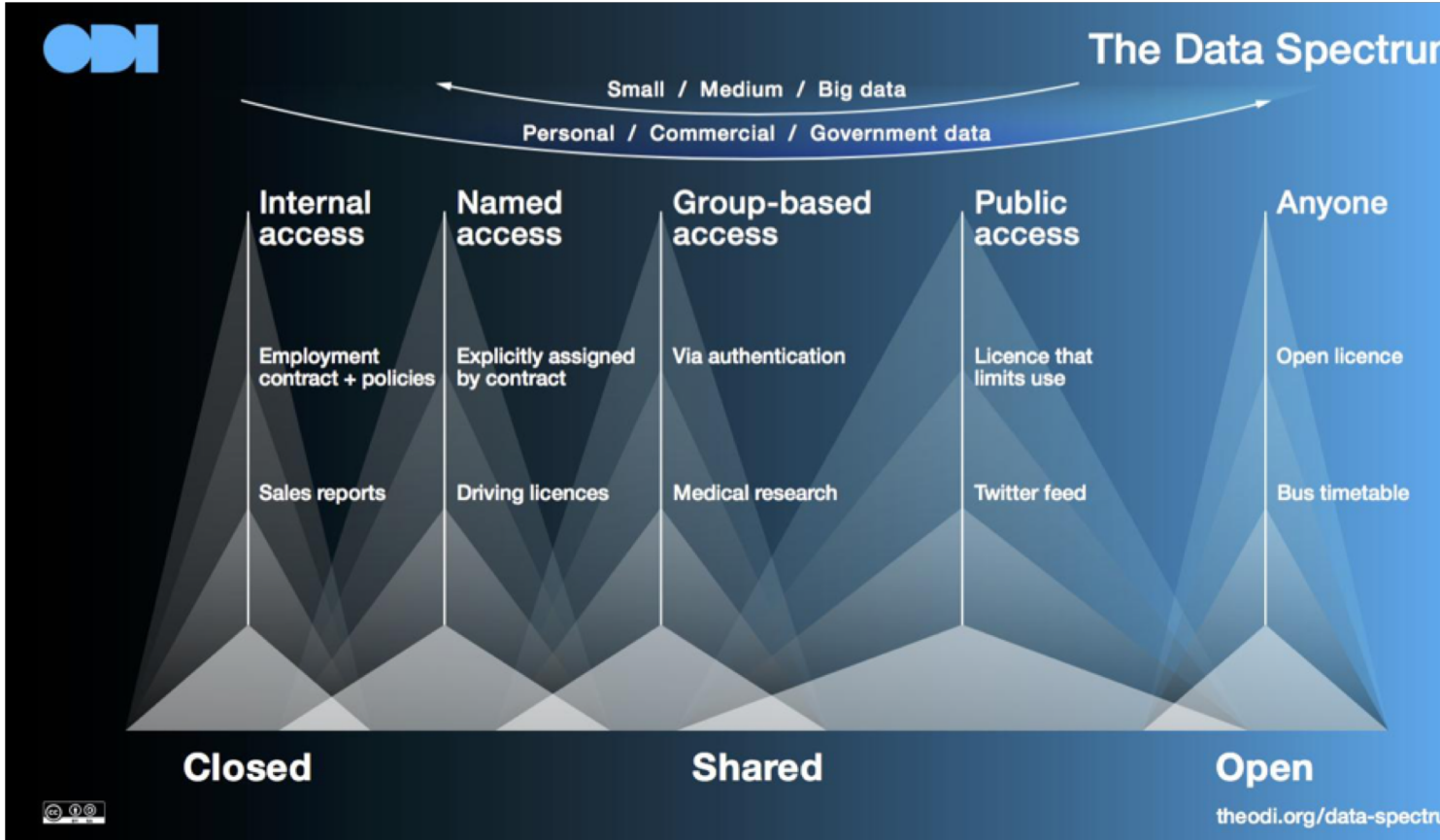
@ilkayholt

Açık Veri Yönetiminde Cerçe ve Yapılar ve Standartlar



İlkay Holt, Danışman

ih@ilkayholt.net



Açık veri, herkes tarafından serbest şekilde kullanılabilir, yeniden kullanılabilir (modifiye edilebilir) ve yeniden dağıtılabılır (paylaşılabilir) veridir (Open Knowledge International).

Verinin erişim ve paylaşımı, özelliklerine bağlı olarak kapalıdan açığa bir spektrumda yer alır.

Açık veri nedir?

Çerçeve yapılar: TBL 5 yıldız

- Verinin paylaşımı konusunda ilk çerçeve yapı, web'in kurucusu Tim Berners Lee (TBL) tarafından Bağlı Açık Veri (LOD) konsepti ile ortaya kondu ve 5 kümülatif ilkesi var:

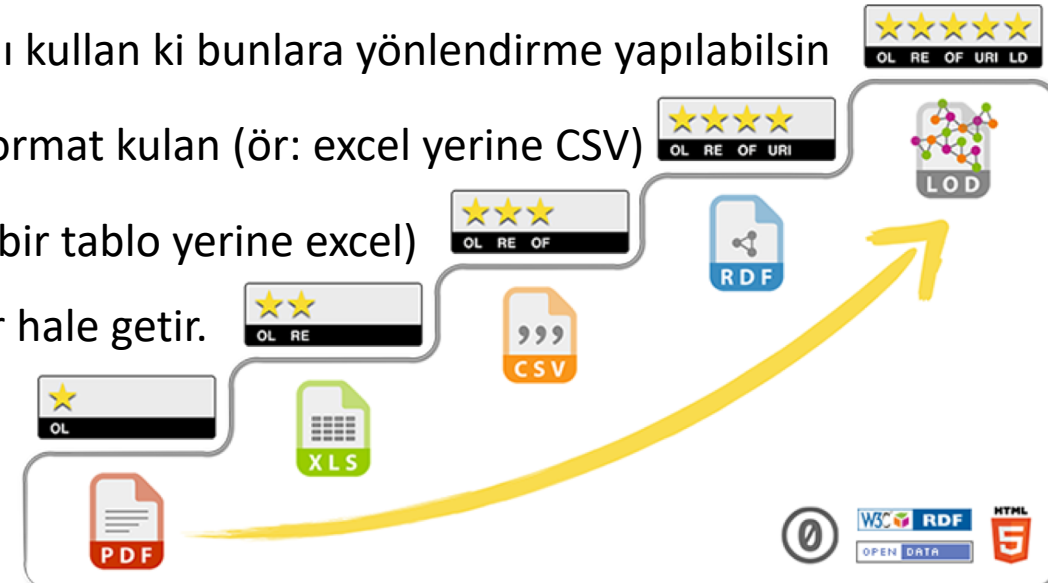
İçerik sağlamak için, verini diğer insanların verisine linkle

Objeleri tanımlamak için W3C (RDF ve SPARQL) açık standartlarını kullan ki bunlara yönlendirme yapılabilirsin

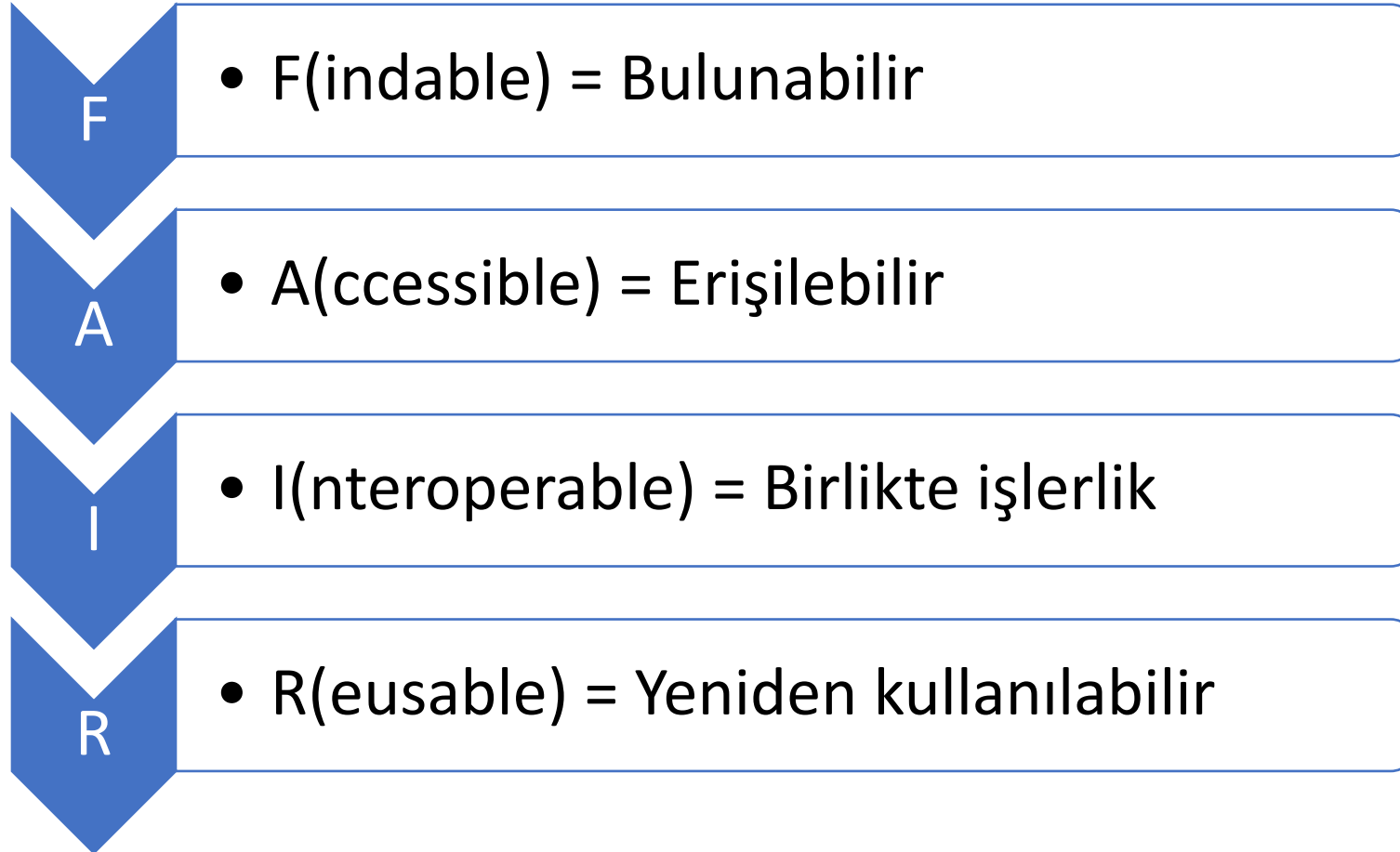
Korumasız bir format kulan (ör: excel yerine CSV)

Makinece okunabilir yapıda olsun (ör: taranmış bir tablo yerine excel)

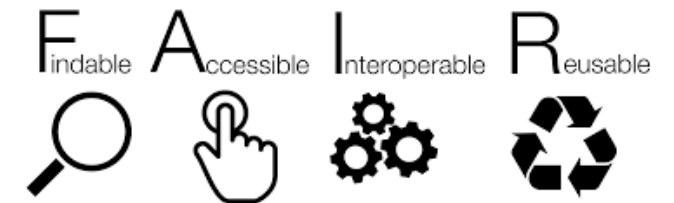
Veriyi (formattan bağımsız) açık lisans ile web'de erişilebilir hale getir.



Bağlı 'açık' veriden (LOD) \Rightarrow bağlı veriye (LD)

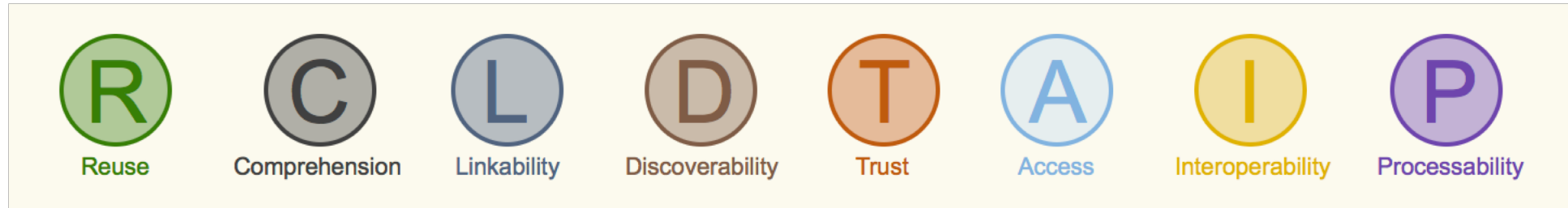


'Mümkün olduğunda açık, gerektiği kadar kapalı'



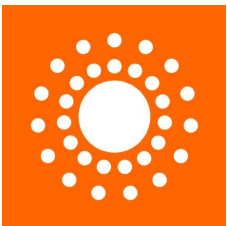
TBL 5 yıldız ve FAIR ilkelerinin farklılıkları

- Odak noktası verinin 'açık' olmasında ziyade erişim hakları üzerinedir.
- Verinin provenansına, atıf yapılmasına ve sürekliliğine dikkat çeker.
- İlkeler teknik spekeifasyon düzeyine inmez, verinin FAIR'lik düzeyini belirleyen genel bir rehber görevi görür.
- TBL 5 yıldız ilkelerinin aksine, FAIR ilkeleri kümülatif değildir. Her biri ayrı ayrı, farklı zamanlarda yerine getirilebilir.



F(indable) = Bulunabilir

- Veri ve yardımcı materyaller yeterince zengin üstveri, tek ve kalıcı tanımlayıcıya sahiptir.
 - F1. üst(veri)ye global olarak tek ve sonsuz kalıcılıkta bir tanımlayıcı atanır
 - F2. veri zengin üstveri ile tanımlanır
 - F3. üst(veri) taranabilir bir kaynakta kayıtlıdır ve indeklenir
 - F4. üstveri veri tanımlayıcıyı nitelendirir, belirtir



Schema.org

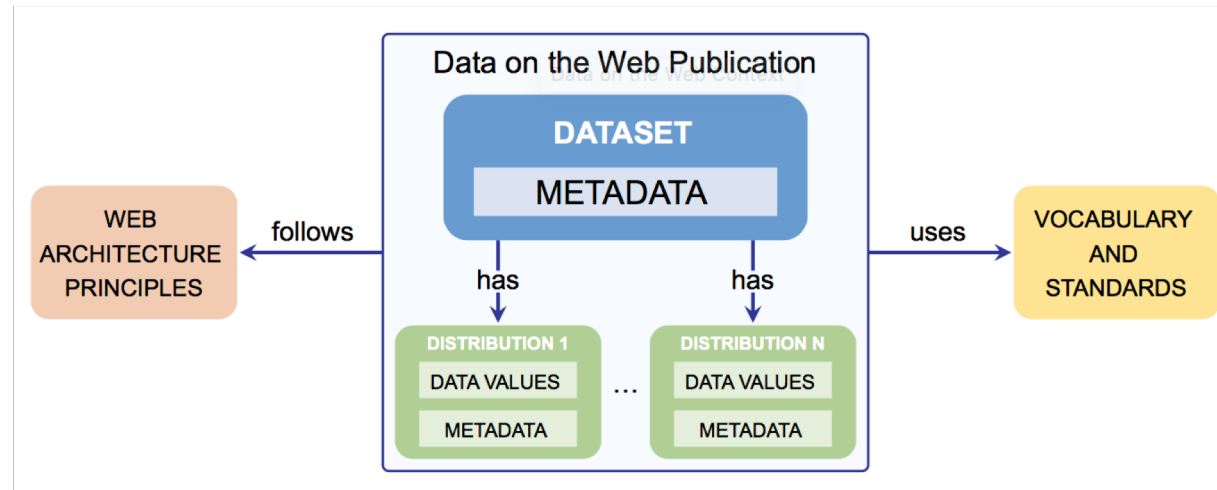
A(ccessible) = Erişilebilir

- Üstveri ve veri insan ve makine tarafından anlaşılır. Veri güvenilir bir arşivde depolanır.
 - A1 (üst)veri tanımlayıcı ID'si ile, standart bir iletişim protokolü kullanarak, bulunabilir
 - A1.1 bu protocol açık, ücretsiz ve global olarak uygulanabilir koşuldadır
 - A1.2 bu protocol yetkilendirmeye, gerektiğinde yetkilendirme prosedürüne izin verir
 - A2 üstveri, veri mevcut olmadığı zaman bile erişilebilirdir



I(nteroperable) = Birlikte işlerlik

- Üstveri bilginin temsili için standart, erişilebilir, paylaşımlı ve yaygın bir biçimde uygulanabilir bir dil kullanır
 - I1. Üstveri bilginin temsili için standart, erişilebilir, paylaşımlı ve yaygın bir biçimde uygulanabilir bir dil kullanır
 - I2. (üst)veri FAIR ilkelerini izlemek için kontrollü sözlükler kullanır
 - I3. (üst)veri diğer üst(veri)ye nitelikli referanlar verir



R(eusable) = yeniden kullanılabilir

- Veri ve koleksiyonlar, net biçimde belirtilmiş lisanslara sahiptir, ve tam ve doğru kaynak bilgisi içerir
 - R1. üst(veri) tam, doğru ve ilgili atıfların tümünü içerir
 - R1.1. (üst)veri net ve erişilebilir bir kullanım lisansı ile açılır.
 - R1.2. (üst)veri kaynağı ile ilişkilendirilir
 - R1.3. (üst)veri disipline özel, ilgili topluluğun ihtiyaçlarını karşılar

[Creative Commons CCZero \(CC0\)](#)

[Open Data Commons Public Domain Dedication and Licence \(PDDL\)](#)

[Creative Commons Attribution 4.0 \(CC-BY-4.0\)](#)

[Open Data Commons Attribution License \(ODC-BY\)](#)

[Creative Commons Attribution Share-Alike 4.0 \(CC-BY-SA-4.0\)](#)

[Open Data Commons Open Database License \(ODbL\)](#)