

KULLANICI ODAKLI MEKANSAL VERİ KALİTESİ

H. Demirel

İTÜ, İstanbul Teknik Üniversitesi, Jeodezi ve Fotogrametri Müh.Bölümü, Fotogrametri Anabilim Dalı Maslak İstanbul, hande.demirel@itu.edu.tr

ÖZET

Kullanıcıların karar verme sırasında kullanabilecekleri farklı sistemler ve mekansal veriler olması mevcut sistemlerin veri kalitesi açısından ele alınması gerekliliğini ortaya koymaktadır. Verilerin kalitesi kullanıcıların kararlarını etkilemektedir. Veri kalitesi ile ilgili güncel araştırmalar; veri kalitesi göstergeleri, kullanıcı ihtiyaçlarının belirlenip, kullanıcıya uygun kalitede verilerin sunulması ve karar verme aşamasında veri kalitesinin etkileridir. Bu çalışmada hedeflenen yaygın olarak kullanılan mekansal verilerin kalitesine ilişkin güncel problemleri ortaya koymak, kararlara etkisini incelemek ve çözümü için mevcut yaklaşımları sunmaktır.

Anahtar Sözcükler: CBS, Mekansal Veri, Kalite, Kullanıcı Merkezli, Meta Veri

ABSTRACT

USER CENTRIC SPATIAL DATA QUALITY

Since various systems and spatial data are available for decision-makers at the decision making process, quality of spatial data should be handled. The quality of spatial data is affecting users decisions. The current studies for data quality focused on; indicators for spatial data quality, determining the user requirements and providing the required information at required level and impacts of spatial data quality on decision –making analyses. The aim of this study is to present the current spatial data quality issues, to analyse the impact on decisions and to review the current solutions.

Keywords: GIS, Spatial Data, Quality, User Oriented, Metadata

1. GİRİŞ

Veri giriş hataları ve veri toplama, modelleme, sunma standartlarının olmaması karar verme sırasında kullanılan bilgilerin %25 oranında hatalı olmasına neden olmaktadır. (Gartner, 2009) Kurumlar ve kuruluşlar stratejik kararlarını ve günlük işlerini doğru olmayan ve tamamlanmamış veriler üzerinde gerçekleştirmektedir. Veri kalitesinden kaynaklı kayıplar artmakta, olası problemlere dikkat çeken araştırmaların sayısı hızla artmaktadır. (DWI, 2009; Devillers vd., 2007; Goodchild, 2007) Veri kalitesi verinin kendisi kadar eski bir konu olmasına rağmen, karar verme aşamasında paydaşların katılımı arttıkça önemi hızla artmaktadır.

Uydu teknolojilerinin gelişmesi ve mekansal teknolojilerin popüler hale gelmesi ile beraber, kullanıcıların karar verme sırasında kullanabilecekleri farklı sistemler ve mekansal veri setleri olması, mevcut sistemlerin veri kalitesi açısından ele alınması gerekliliğini ortaya koymaktadır. Son on yıl içerisinde gözlenen en büyük eğilim mekansal verinin “demokratikleştirilmesi”dir. (Devillers vd., 2007) Buna en iyi örneklerden birisi Google Earth olarak verilebilir. Bu tür olanaklar mekansal verinin popülerleşmesini sağlamış ve yeni uygulamalara olanak sağlamıştır. Mekansal veri miktarı olarak artmış, mekansal veriye erişim kolaylaşmış ve mekansal verinin toplanması, analizi ve sunulması için teknoloji sunulmuştur. Mekansal veri günümüzde farklı kullanıcı grupları tarafından sıklıkla kullanılmaktadır. Genel olarak olumlu olarak değerlendirilen bu gelişmeler verilerin toplanırken tasarlanan kullanım amacı dışında kullanılmasına, farklı kurum ve kuruluşların veri toplamasına ve farklı teknoloji, yöntem ve standartlarda veri toplanmasına neden olmaktadır. Bu bağlamda, mekansal verilerin hatalı, yanlış ve kötü kullanımından bahsedilmektedir. Verilerin bu şekilde kullanımının belirgin sosyal, politik ve ekonomik etkileri farklı örneklerde bulunabilir. (Beard 1989; Monmonier 1994; Agumya ve Hunter 1997; Gervais 2006)

Verilerin kalitesi kullanıcıların kararlarını etkilemektedir. Veri kalitesi ile ilgili güncel yaklaşımlar, veri kalitesinin kullanımına ilişkin araştırmalar, kullanıcı ihtiyaçlarının belirlenip, kullanıcıya uygun kalitede verilerin sunulması ve karar verme aşamasında veri kalitesinin etkilerinin araştırılmasıdır. Veri kalitesi, veri toplama sırasında belirlenebilirken, verilerin işlenmesi ve analizi sonrası kalitede farklılaşmalar oluşabilmektedir. Sistem bileşenlerinden mekansal veri; farklı güncellik, doğruluk, duyarlılık, güvenilirlik, tutarlılık ve tamlıktaki çeşitli kaynaklardan, farklı yöntemler kullanılarak toplanmakta ve entegre edilerek kullanıcıya sunulmaktadır.

2. MEVCUT DURUM

Mekansal veri kalitesi uluslar arası standartlar organizasyonu ISO'nun teknik komisyonu TC 211 göre mekansal veri kalitesi mekansal verinin nasıl kullanılması gerektiğini tanımlayan bileşendir. (Tastan ve Altan, 1999) Mekansal veri kalitesini iç ve dış olarak iki bölümde ele alınabilir. İç veri kalitesi; tamlık, konumsal doğruluk, zamansal doğruluk, tematik doğruluk ve kavramsal tutarlık (Guptill and Morrison 1995; ISO/TC 211, 2002) olarak incelenebilirken, dışsal doğruluk kavramı verinin kullanım amacına uygunluğunun araştırılmasıdır. (Veregin 1999; Devillers and Jeansoulin 2006) İç ve dış doğruluk Şekil 1'de sunulmaktadır. Kullanıcı açısından, iç doğruluk üretici tarafından meta veriler ile birlikte verilmektedir. İç veri kalitesi genellikle veri üreticisi tarafından yönetmelikler ve ulusal, uluslararası standart tarafından belirlenirken, dış veri kalitesinin belirlenmesi araştırma konusudur.



Şekil 1: Mekansal veri kalitesi

Mekansal verinin değeri, karar verme aşamasında etkili olduğu sürece artmaktadır. Mekansal verinin değeri kullanıcı açısından zaman faktörü (hızlı karar verebilme, hızlı cevap alabilme), veri ve bilgiye dayanan tartışmasız ve sürekli karar verebilme ve yanlış karar verme riskinin azaltılması ile ölçülebilir. Bilgi Kalitesi veri kalitesi ve kullanıcı arasındaki ilişkiyi tanımlamaktadır. Bilgi kalitesinin genel tanımları “kullanım amacına uygunluk”, “kullanıcı ihtiyaçlarını karşılayan bilgi” olarak verilebilir. Bilgi kalitesinin amacı uygulama, veri ve kullanıcı arasındaki ilişkinin tanımlanmasıdır. Mekansal veri kalitesini arttırmak için, örneğin mekansal bilgi teknolojilerinde geometrik nesnelere için hata modelleri, belirsiz veri ve bilgiler ile mekansal sorgulamalar ve bunların verilen kararlara etkileri, belirsiz verilerin kullanıcıya sunulması ve metaveriler içerisinde veri kalitesinin sunulması gibi, pek çok farklı yöntem kullanılmakta ve araştırmalar devam etmektedir. Tüm bu araştırmalara rağmen mekansal veri kalitesi kullanıcılar ve veri üreticileri için hala bir sorun olarak değerlendirilmektedir. Üreticiler veri kalitesi hakkında gerekli bilgiyi sağlamamaktadır. (Östman, 1997; Jakobsson ve Vauglin, 2001) Kullanıcı açısından en önemli problem ise farklı kaynaklardan gelen verilerin ortak kullanımınıdır. Kullanıcılar, değerlendiriciler ve veri üreticileri arasındaki görüş farklarını inceleyen modeller mevcut araştırma alanlarının başındadır. (Demirel ve Altan, 2005) En önemli problem; sistemin kurulum aşamasında ve güncelleştirilmesi sırasında hızlı ve ekonomik veri girişi ve veri toplanması, maliyetleri ve sistemin yürürlüğe konulması için gereken zamanı kısaltırken, veri kalitesini düşürmektedir.

Kullanıcılar genellikle veri kalitesini göz ardı etmekte ve risk almaktadır. Yapılan araştırmalar ve geniş katılımlı anket çalışmaları sonucunda kullanıcıların verinin uygun olmasını isteyenler, en uygun alternatifi seçecekler ve kalite bilgisini kullanmayanlar olarak üç gruba ayrıldığı görülmüştür. (Agunya ve Hunter, 1999) Kaliteyi önemsememenin alınan kararlarda riski arttırması kalitesi araştırmalarına hız kazandırmıştır. (Oort ve Bregt, 2005) Araştırmaların bu gün geldiği nokta, farklı kalite seviyelerinin olduğu ve bunun kullanıcı tarafından belirlendiğidir. Mutlak bir kaliteden bahsetmek mevcut uygulamalar düşünüldüğünde mümkün olmamaktadır. Mevcut mekansal veri teknolojilerinde en önemli nokta kullanıcı, sistem ve veriler arasında ilişkinin bulunması, belirlenmesi ve uygulanmasıdır. Mekansal bilgi teknolojileri gibi farklı kullanıcı grupları ve uygulamaların olduğu bir sektörde farklı ihtiyaçlar için farklı çözümler geliştirilmesi gerekliliği açıktır. Kalite kavramına daha esnek olarak yaklaşmak mekansal veri üreticileri için yeni olanaklar meydana getirecektir.

Veri kalitesi ile ilgili son yıllarda pek çok kitap yayınlanmıştır. Bunlardan başlıcaları Wnzhong Shi, Peter Fisher ve Michael F. Goodchild tarafından 2002 yılında yazılan “Mekansal Veri Kalitesi” ve Rodolphe Devillers ve Robert Jeansoulin tarafından 2006 yılında yazılan “Mekansal Veri Kalitesinin Temelleri”dir. (Shi vd., 2002; Devillers ve Jeansoulin, 2006) Ayrıca konu ile ilgili bilimsel makaleler fotogrametri, uzaktan algılama ve mekansal bilgi bilimleri konusunda yayınlanan periyodik yayınlardan (International Journal of Geographical Information Science, International Journal of Applied Earth Observation and Geoinformation, ISPRS Journal of Photogrammetry and Remote sensing,

Geoinformatica vd.) takip edilebilir. Mekansal veri kalitesi ile ilgili düzenli olarak organize edilen International Symposium on Spatial Data Quality ve Spatial Data Handling toplantıları son yıllarda yoğun ilgi görmektedir. Ayrıca konu ile ilgili uluslar arası kuruluşların mekansal veri kalitesi ile ilgili çalışma grupları bulunmaktadır. Bunların başlıcaları Open Geospatial Consortium veri kalitesi grubu (<http://www.opengeospatial.org/projects/groups/dqwg>) ve Uluslar arası Fotogrametri ve Uzaktan Algılama Birliği (ISPRS), ikini teknik komisyonu altında bulunan dördüncü çalışma grubu "Belirsizliğin modellenmesi ve mekansal verinin kalite kontrolü"dür(<http://www.isprs.org/commission2/>)

3. TARTIŞMA VE ÖNERİLER

Mekansal bilgi teknolojilerinin gelişmesi, hızlı ve otomatikleştirilmiş veri toplama yöntemlerinin yaygınlaşması ile beraber, mekansal veri kalitesi ile ilgili araştırmalar hız kazanmıştır. Mekansal veri kalitesinin öneminin artmasının nedenleri (Longley et al.,1999)

- i) mekansal verinin kullanımı, ulaşılabilirliği ve değiş-tokuşu artmıştır.
- ii) mekansal veri kalitesi kavramının farkında olmayan kullanıcılar artmaya başlamıştır.
- iii) mekansal bilgi teknolojileri yardımı ile çok sayıda uygulamada kullanılan verilerin uygun olup olmadığı araştırılmamaktadır.
- iv) mevcut yazılımlar veri kalitesi ile ilgili herhangi bir araç sunmamaktadır.
- v) Son kullanıcı ve kalite hakkında en fazla bilgisi olması gereken üretici arasındaki mesafe artmaktadır.

Veri kalitesi ile ilgili güncel tartışma ve araştırma konuları aşağıda sıralanmıştır;

- Mekansal veriler gün geçtikçe paylaşılmakta, değiş-tokuş edilmekte ve üreticinin üretim amacı dışında kullanılmaktadır.
- Kullanıcılar, değerlendiriciler ve veri üreticileri arasındaki görüş farkları bulunmaktadır. Veri toplama ve güncelleştirme maliyetleri ve veri kalitesi arasındaki ilişkiler bu konudaki en önemli araştırma alanlarından birisidir.
- Kullanıcı açısından en önemli problem ise farklı kaynaklardan gelen verilerin ortak kullanımınıdır.
- Mevcut verilerin kullanıcının ihtiyaçlarına uygun olup olmamasının belirlenmesi ayrı bir problemdir ve deneyim gerektirmektedir.
- Mekansal bilgi teknolojileri çerçevesinde farklı uygulama alanları için kullanılan verilerin verilen kararlara ne derece etki ettiği önemli bir araştırma konusudur. Mevcut durumda veri üreticisinin mekansal bilgi kalitesi kabulleri ve kullanıcı ihtiyaçları, veri kalitesinin yanlış kullanımını doğurmaktadır. Veri üreticilerinin veri kalitesini değerlendirmeleri ve veriler ile birlikte sunulması ürünün pazar payını arttıracaktır.
- Ulusal ve uluslar arası veri altyapıları içerisinde meta verilerin toplanması, saklanması ve sunulması verinin kullanıma uygunluğu araştırmalarını destekleyen en önemli veri kaynağıdır. CBS'de meta veri olarak kalite bilgisinin sunulması risklerin belirlenmesi açısından ilk adım olarak önemlidir
- Konu üzerinde daha fazla teorik araştırma yapmak ve kullanıcıları bilgilendirmek için en iyi uygulama örneklerini hazırlayıp sunmak gerekmektedir.
- Mekansal verilerdeki belirsizlikleri anlamak gerekmektedir.
- Mekansal-zamansal verinin kalitesinin kontrolü yapılmalıdır.
- Mekansal analizde hata yayılımının modellenmesi gerekmektedir.
- Veri kalitesi problemi ile karşılaşan kurum ve kuruluşların bu problemi giderme yöntemleri bilinmemektedir.
- Zamansal-mekansal modellerin kalitesi araştırılmaktadır.
- Mevcut CBS yazılımlarının veri kalitesinin analizi ve sunumu açısından güçlendirilmesi gereklidir.

Ülkemizde geçmişte veri kaynakları kısıtlı olduğundan pratik önemi az olan bu konu, hızla mekansal verilerin toplandığı şu aşamada önem kazanmaktadır. Ülkemizde de diğer ülkelerde olduğu gibi mekansal veri kullanan sektörlerde kullanıcı isteklerini belirlemek amacı ile analizler yapılmalı ve mekansal veri kalitesinin daha ekonomik ve işlevsel şekilde kullanıcılara sunulmasının yöntemleri araştırılmalıdır.(Demirel ve Altan, 2005) Aynı zamanda veri kalitesi ile ilgili riskler kullanıcılara, özellikle karar vericilere sunulmalıdır. Bilgi, kullanıcı ihtiyaçları ve özellikle karar verme aşaması ile ilişkilendirildiği oranda değerlidir. Özet olarak önerilen, üretilen mekansal verilerin kullanıcıların amaçlarına ve isteklerine göre toplanması, modellenmesi, analizi ve sunulmasıdır.

KAYNAKLAR

Agumya, A. and Hunter, G. J., 1997, Determining fitness for use of geographic information, ITC Journal 2 , pp. 109-113

- Beard, K., 1989, *Use error: the neglected error component*, Proceedings of AUTO-CARTO 9 pp. 808-817. — Baltimore, MD
- Data Warehousing Institute (DWI), 2009, *Data quality and the bottom line: archiving business success through a commitment to high quality data*, <http://www.dw-institute.com>, 14 Nisan 2009
- Demirel, H., Altan, O., 2005, *Mekansal veri kalitesinde güncel yaklaşımlar*, TMMOB Harita ve Kadastro Mühendisleri Odası 10. Türkiye Harita Bilimsel ve Teknik Kurultayı, 28 Mart - 1 Nisan 2005, Ankara
- Devillers, R., Bedard, Y., Jeansoulin, R., Moulin, B., 2007, *Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data*, International Journal of Geographical Information Science, Volume 21, Issue 3, syf. 261 – 282
- Devillers, R. and Jeansoulin, R. (eds), 2006, *Fundamentals of Spatial Data Quality* ISTE , London
- Gartner Inc., 2009, *Gartner predicts 2009*, <http://www.gartner.com/>, 14 Nisan 2009
- Gervais, M. Devillers, R. and Jeansoulin, R. (eds), 2006, *On the Importance of External Data Quality in Civil Law*, Fundamentals of Spatial Data Quality pp. 283-300. ISTE , London
- Goodchild M., 2007, *Beyond metadata: Towards user-centric description of data quality*, ISSDQ 2007, ITC, Netherlends, <http://www.itc.nl/ISSDQ2007/keynote.aspx>, 14 Nisan 2009
- Guptill, S. C. and Morrison, J. L. (eds), 1995, *Elements of Spatial Data Quality*, Elsevier Science , Oxford
- ISO/TC 211, 2002, *Geographic Information—Quality Principles, Report—* 19113
- Jakobsson, A., Vauglin, F., 2001, Status of data quality in European national mapping agencies. Proceeding of the 20th International Cartographic Conference , 4 , 2875–2883.
- Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., 1999, Introduction to Data Quality. In: Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W. (eds.), *Geographical Information Systems: Principles and technical issues*, New York: John Wiley & Sons, 2nd ed., vol. 1, pp. 175-176.
- Monmonier, M., 1994, A case study in the misuse of GIS: Siting a low-level radioactive waste disposal facility in New York State. pp. 293-303.
- Oort, P., Bregt, A., 2005, Do users ignore spatial data quality? A decision-theoretic perspective, Risk Analysis, Volume: 25 Issue: 6 Pages: 1599-1610
- Östman, A., 1997, The specification and evaluation of spatial data quality. In Proceedings of the 18th ICA International Cartographic Conference (pp. 836–847), Sweden
- Shi, W., Fisher, P. F. and Goodchild, M. F. (eds), 2002, *Spatial Data Quality*, Taylor & Francis , London
- Taştan, H. and Altan, M.O., 1999, *Spatial Data Quality*, Proceedings of the Third Turkish-German Joint Geodetic Days, June 1-4, Istanbul, pp.15-30
- Veregin, H. Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (eds), 1999, Data quality parameters, *Geographical Information Systems* pp. 177-189. Wiley , New York