

MEKANSAL VERİ KALİTESİNDE GÜNCEL YAKLAŞIMLAR

H. Demirel, M.O. Altan

İstanbul Teknik Üniversitesi, Jeodezi ve Fotogrametri Mühendisliği Bölümü, Fotogrametri Anabilim Dalı, İstanbul,
hande.demirel@itu.edu.tr, oaltan@itu.edu.tr

ÖZET

Çalışmada hedeflenen yaygın olarak kullanılan mekansal verilerin kalitesine ilişkin güncel problemleri ve çözümü için yaklaşımları tartışılmaktadır. Mevcut durumda güncel yaklaşımlar; bilgi kalitesinin kullanımına ilişkin araştırmalar, kullanıcı ihtiyaçlarının belirlenip, kullanıcıya uygun kalitede verilerin sunulması, karar verme aşamasında veri kalitesinin etkilerinin araştırılmasıdır. Bu yaklaşımlar tartışılırken öncelik veri kalitesi konusunda kullanıcıların sistemden beklentileri ve karar-destek aşamasında veri kalitesinin etkilerine verilecektir. Çalışma kapsamında mekansal veri kalitesi temel kavramları, hata kaynakları, veri kalitesi kontrolü konuları incelenecek, karar-destek sistemlerinde veri kalitesinin önemi örnekler üzerinde tartışılacaktır. Bu amaçla uluslararası standartlar incelenecek ve kullanıcı ihtiyaçlarını ne ölçüde karşıladıkları ve olası çözüm önerileri tartışılacaktır.

Anahtar Sözcükler: CBS, mekansal veri, kalite, kullanıcı merkezli

ABSTRACT

CURRENT APPROACHES TO SPATIAL DATA QUALITY

The aim of this study is to discuss the recent problems of spatial data quality and approaches for solving them. Current approaches can be categorized as follows; research on the usage of data quality, determining user requirements and providing data in most "suitable, appropriate" quality and investigating impacts of data quality during decision making. Within this study, user expectations on data quality and impacts of data quality during decision making will be emphasized. In the frame of the study, basics of spatial data quality, error sources, data quality control is introduced within a decision making system and examples highlighting the importance of data quality is presented. International standards are studied and specifically within a user-oriented approach and possible solutions are discussed.

Keywords: GIS, spatial data, quality, user-oriented

1. GİRİŞ

Mekansal Bilgi Sistemlerinin kurumlarda yaygın olarak karar-destek sistemleri olarak kullanılması ile birlikte mekansal veri kalitesi hem araştırmacıların, hem de sistemin hizmetlerinden yararlanan kullanıcıların artan ilgisini çekmekte ve dünyada bu konuda pek çok araştırma yapılmaktadır. Kullanıcıların karar verme sırasında kullanabilecekleri farklı sistemler ve mekansal veri setleri olması mevcut sistemlerin veri kalitesi açısından ele alınması gerekliliğini ortaya koymaktadır. Yeterli kalitede olmayan veriler, mekansal bilgi sisteminden beklenen avantajlar yerine, pek çok hatalı kararın alınmasına sebep olmaktadır.

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. Mekansal veri kalitesini arttırmak için, örneğin mekansal bilgi teknolojilerinde geometrik nesneler için hata modelleri, belirsiz veri ve bilgiler ile mekansal sorgulamalar ve bunların verilen kararlara etkileri, belirsiz verilerin kullanıcıya sunulması ve metaveriler içersinde 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. Kullanıcılar, değerlendiriciler ve veri üreticileri arasındaki görüş farklarını inceleyen modeller mevcut araştırma alanlarının başındadır. 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. Veri toplama ve güncelleştirme maliyetleri ve veri kalitesi arasındaki ilişkiler bu konudaki en önemli araştırma alanlarından birisidir. Bir diğeri mekansal bilgi teknolojileri çerçevesinde farklı uygulama alanları için kullanılan verilerin verilen kararlara ne derece etki ettiği, veri kalitesi kabulleri ve sunumudur. Mevcut durumda veri üreticisinin mekansal bilgi kalitesi kabulleri ve kullanıcı ihtiyaçları, veri kalitesinin yanlış kullanımını doğurmaktadır. Fakat bilinmektedir ki, veri üreticilerinin veri kalitesini değerlendirmeleri ve veriler ile birlikte sunulması ürünün pazar payını arttıracaktır.

Mevcut durumda güncel yaklaşımlar;

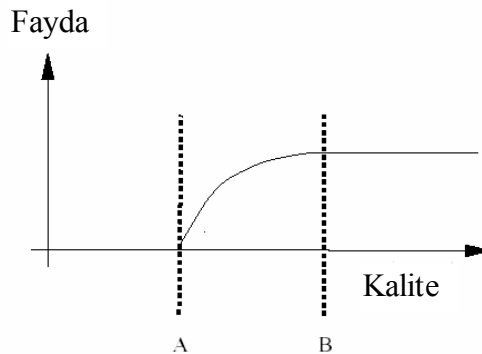
- Bilgi kalitesinin kullanımına ilişkin araştırmalar
- Kullanıcı ihtiyaçlarının belirlenip, kullanıcıya uygun kalitede verilerin sunulması
- Karar verme aşamasında veri kalitesinin etkilerinin araştırılmasıdır.

Dünyada gerçekleştirilen tüm bu araştırmalar, belirlenen standartların çeşitli kullanıcılar için geliştirilmesine yöneliktir. Ülkemizde ulusal düzeyde veri üretim standartları bulunmaması yakın gelecekte ve mekansal verilerin kullanımının artması ile beraber mekansal bilgi teknolojilerinde önemli problemler ile karşılaşılmasına neden olacaktır.

Çalışma kapsamında mekansal veri kalitesi temel kavramları, hata kaynakları, veri kalitesi kontrolü konuları incelenecek, karar-destek sistemlerinde veri kalitesinin önemi örnekler üzerinde tartışılacaktır. Mevcut çerçeve uluslararası standartlar incelenecek ve kullanıcı ihtiyaçlarını ne ölçüde karşıladıkları ve olası çözüm önerileri tartışılacaktır.

1. VERİ KALİTESİNDE GÜNCEL YAKLAŞIMLAR

Veri kalitesi ile ilgili herhangi bir tanımlama, ölçme ve değerlendirme yapmadan önce ‘Bilginin hangi özellikleri hakkında kalite bilgisine sahip olmalıyız?’ sorusunun cevabı aranmalıdır. Bu öncelikle araştırdığımız veri kalitesinin farklı kullanımlarda ne gibi bir önemi olduğunu ve sonuç ürünleri ne derece etkileyeceğinin belirlenmesini sağlayacaktır. Teknik olarak, bilgileri kalitesi hakkında veriler ile sunmak mümkündür; fakat bu kalite hakkındaki veriler bilginin hangi özellikleri hakkında kalite bilgisine sahip olmamız gerektiğine dair bilgi vermeyebilir. Veri hakkında kalite bilgilerinin sunulmasına veri toplama, veri modeli tasarımı ve sunumu aşamasında ek yükler getireceğinden hem veri üreticileri, hem de kullanıcıları tarafından gereken önem verilmemektedir. Burada esas problem bir optimizasyon problemi olup veri kalitesi kullanıcı ihtiyaçları ile ilişkilendirildiği oranda değer kazanacaktır. Geleneksel olarak veri üreticisi en yüksek teknik kalitede, yüksek doğrulukta veri setleri ve detaylar, veri üretimine önem verirken, bu politika yüksek üretim maliyetlerine ve yüksek fiyatlı mekansal veri üretimine sebep olmaktadır. Son kullanıcılar verileri satın alırken, her üründe olduğu gibi istenilen fiyatın adı geçen ürün için değer olup olmadığını değerlendirmek durumundadır. Diğer ülkeler ile karşılaştırıldığında henüz ülkemizde mekansal veri üretimi az sayıda özel kuruluş tarafından gerçekleştirilmektedir. Bu yüzden veri fiyatları konusunda henüz bir değerlendirme yapmak olanak dışıdır. Fakat yine bilinen bir gerçek, mekansal verileri kendileri toplayan kamu kuruluşları yüksek maliyetli veri toplama aşamasında çok çeşitli teknik ve mali problemler ile karşılaşmaktadır. Sürekli olarak güncellenen mekansal bilgi sistemleri sayısı oldukça azdır. Bu optimizasyon problemi veri üretiminin ve dolayısı ile veri kalitesinin kullanıcı merkezli bir yaklaşımla tekrar gözden geçirilmesi yöntemi ile çözülebilir. Veri üreticisi, mekansal veriden gerekli faydayı sağlayabilmek için verisini belirli kalitede sunmak zorundadır. Fakat mekansal verinin sağladığı faydalar çizgisel olarak artmamaktadır. Mekansal verinin kullanıcı için faydası ve kalite arasında ki ilişki Şekil 1’de sunulmaktadır (Krek, 2000). Şekildeki A noktasına kadar üretilen mekansal veri kullanıcı için çok düşük kalitede olduğundan herhangi fayda sağlamayacaktır. A ve B arasında mekansal veri potansiyel kullanıcı için ihtiyaç duyulan kalitededir. A ve B arasında kalan alanda kalitenin yükselmesi kullanıcı için faydanın artacağı anlamına gelmektedir. B noktasından sonra mekansal kalitenin artırılması kullanıcı için herhangi bir fayda sağlamamaktadır.



Şekil 1: Fayda ve kalite arasındaki bağıntı

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 (Krek, 2000). Bilgi Kalitesi (BK) veri kalitesi ve kullanıcı arasındaki ilişkiyi tanımlamaktadır. Bilgi kalitesinin genel tanımları “kullanım amacına uygunluk” (Tay, 1998), “kullanıcı ihtiyaçlarını karşılayan bilgi”, (Redman, 1996) olarak verilebilir. Bilgi kalitesinin

amacı uygulama, veri ve kullanıcı arasındaki ilişkinin tanımlanmasıdır(Wang, 1999). Bilgi kalitesi dört gruba ayrılmakta ve bu grupların tanımları (Naumann, 2002) ve ISO/FDIS 19115’de (URL 1, 2004) bulunabilmektedir. İçerik ile ilgili kriterler kapsamı mevcut veriler ve gerçek değerlerinin ne derece yansıtıldığıdır. Teknik kriterler; veri toplama kaynağının yazılım, donanım, kullanıcıya iletilmesi ve kullanıcıyı kapsamaktadır. Fikri kriterler veri kaynağının subjektif kısımları ile bağlantılıdır. Kullanıcıya ve geliştiricilere bağlı olduğundan güçlükçe ölçülebilir. Kurulum ile ilgili kriterler verinin sunumu ve buna bağlı olarak ergonomisi ve kullanılabilirliği ile ilişkilidir. Bu dört kategori ve ilgili kriterler Tablo 1’de sunulmaktadır (Jahn, 2004).

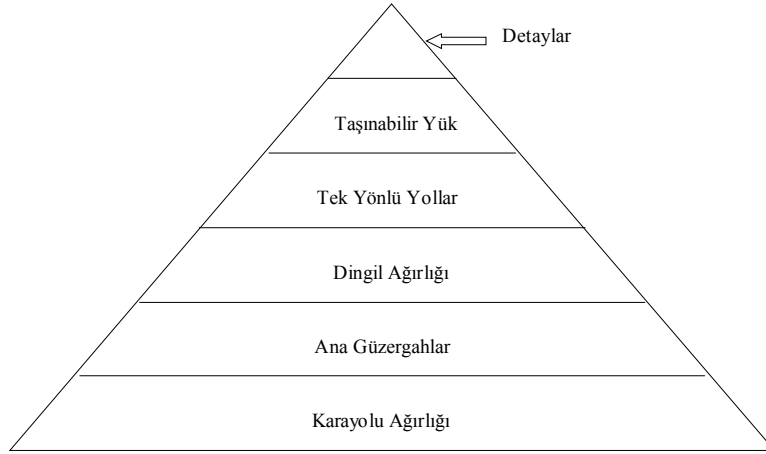
Tablo 1: Bilgi Kalitesi (BK) Kategori ve Kriterleri

Kategori	BK-Kriteri
İçerik	Doğruluk
	Tamlık
	Kullanıcı desteği
	Dokümantasyon
	Yorumlanabilirlik
	İlgi
	Değer-katılmış
Teknik	Ulaşılabilirlik
	Patent
	Fiyat
	Servis kalitesi
	Geri dönüş süresi
	Güvenlik
	Zamanlama
Fikri	İnandırıcılık
	Nesnellik
	İyi referans
Kurulum	Veri miktarı
	Sunuş hızı
	Sunuş bütünlüğü
	Anlaşılabilirlik
	Doğrulanabilirlik

Bilgi kalitesinin belirlenmesi aşamasında en önemli faktör kullanıcıdır. Kullanım amacına uygunluğun gerçekleşebilmesi için kullanıcının kriter oluşturma adımlarına dahil olması gerekmektedir. Kullanıcı grubundan gelen her tür isteğin bir temel verisine ve bilgi kalitesini oluşturabilmek için “veri kalitesi” bilgisine ihtiyaç bulunmaktadır.

2. KARAR VERME AŞAMASINDA MEKANSAL VERİ KALİTESİ

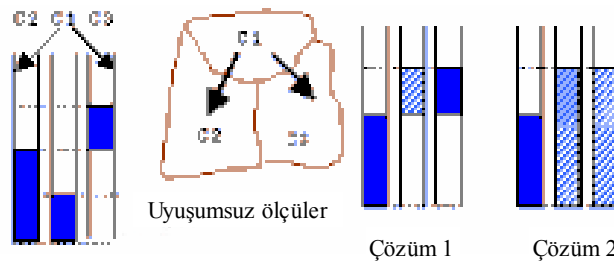
Kullanıcı merkezli veri kalitesi kavramı ile ilgili örnek Türkiye’de de hızla gündeme gelen araç bilgisayarları alanında verilebilir. Örnek olarak şehir içinde güzergahını araç bilgisayarı ile planlamak isteyen bir tır şoförünün ihtiyacı olan veriler nelerdir? Sorusunun cevabı aranmaktadır. Kullanıcı ihtiyaçları Maslow Piramidine önem derecelerine göre yerleştirilmiştir (Sheth, 2000). Maslow Piramidi Şekil 2’de sunulmuştur. Eğer ihtiyaçlardan bir tanesi veri setinde eksik olursa, örneğin ana güzergahlar, sadece altlık harita yeterli olmayacak ve haritanın kullanımı tır şoförü açısından ne kadar yüksek kaliteli olursa olsun mümkün olmayacaktır. Aynı sürücü kendisi için uygun yolu bulabilmek için belirli bir doğrulukta konum bilgisine ihtiyaç duymaktadır. 100 m’nin üzerindeki bir konum bilgisi sürücü için değersizdir. Bu doğrulukta bir veri ile yanlış yolu takip etme olasılığı yüksektir. Daha yüksek doğrulukta konum bilgisi sürücü için daha değerlidir ve yanlış yolu seçme riskini azaltacaktır.



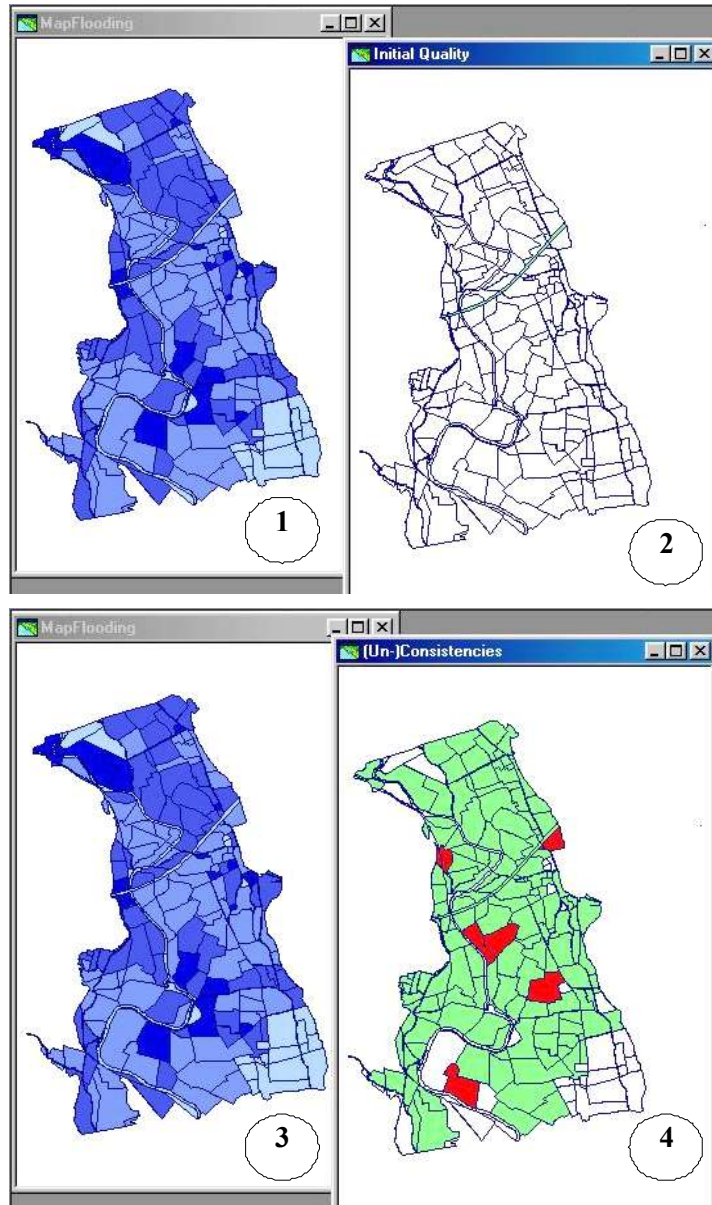
Şekil 2: Araç bilgisayarı için kullanıcı ihtiyaçları

Sürücü için 50m’lik doğruluk sürücü için ihtiyaç duyulabilecek en yüksek doğruluktur. Daha yüksek doğruluk sürücü için herhangi bir ek değer taşımamaktadır. Daha yüksek doğruluk daha yüksek maliyet ve daha fazla zaman (haritada ki genelleştirmeden kaynaklanan farklılıklar) gerektirmektedir. Burada kullanıcı hem maliyet düşüklüğü, hem de kullanım kolaylığı açısından kendisine en uygunu seçecektir.

Kullanıcı merkezli kalite kavramı, Avrupa Birliği Bilgi Teknolojileri araştırma projesi, (FPP5)olan REVIGIS kapsamında incelenmiştir. (URL 2, 2004) Dört ayrı projede gerçekleştirilen çalışmadan, bu çalışma kapsamında sadece bir örnek verilecektir. Örnek sistem, sel felaketi sonucu oluşturulmuştur. REVIGIS projesi kapsamında su yüksekliğine, kısa dönemde selin hareketinin öngörülmesine dayalı (sonraki 12 saat için; su seviyesi alçalacak mı?, yükselecek mi?) mekansal analizler gerçekleştirilmiştir(Raclet, 2003)(Raclet, 2004). Araştırılan konular; (1) ekili arazideki her bir parsel için su seviyesinin belirlenmesi, (2) suyun yönünün kontrol edilebilmesi için görülebilir akıntıların gözlenmesidir. Su seviyesi bilinen en düşük seviye (bilgi sisteminden alınan) ve en yüksek seviye (hava fotoğraflarından görünen) yardımı ile belirlenmiştir. Uygulama için gerekli mekansal bilgi sistemi varlıkları; parseller, sınırlar, su, ekili alan, su seviyesi ve ekili alan arasındaki ilişki, akıntının yönü ve akıntı ve su seviyesi arasındaki ilişkidir. Bu analizler için kullanıcılardan temin edilen veriler; sayısal arazi yüksekliği, ekili alan yüksekliği (veri kalitesi açısından eksik ve yeterli doğrulukta değil), su seviyesi ve ekili alan ilişkisi (veri kalitesi açısından eksik) ve akıntı yönleri(eksik fakat mevcut veri yeterli doğruluktadır.) Çalışmada temel yaklaşım “istenilen kullanım için uygunluktur”. Kullanıcı istekleri ile mevcut verinin bağlantısı incelenmiş ve veri kalitesi kullanıcının beklentisi doğrultusunda kontrol edilmiştir. Çalışmadan elde edilmiş ilk veriler Şekil 3’de sunulmuştur. Soldaki harita, parsel başına ortalama su seviyesini göstermektedir. Sağdaki haritalar ise kalite ile ilgili bilgileri görsel olarak sunmaktadır. Bir numaralı harita sistemden alınmıştır. İki numaralı haritada ise herhangi bir işlem yapılmamış, kalite ile ilgili bilgiler bilinmemektedir. Dördüncü haritada ise yeşil kısımlar “doğru”, kırmızılar “uyuşumsuzluk”, beyazlar ise “bütünlüğün” kontrol edilemediği kısımlardır. Örnek olarak parsel “dağılım kuralının” uygulanmasına olanak vermeyecek kadar düz değil ise. Bir parsel için birden fazla uyşumsuzluk, bir uyşumsuzluk için birden fazla parsel söz konusu olabilir. “Uyşumsuzluklar” bir kez belirlendikten sonra, güvenilirlikleri çok yüksek olmayan veriler için bazı iyileştirmeler önerilebilir. Bu örnek için uyşumsuzlukların olduğu bölgeler için değer aralıkları değiştirilip uyşumsuzluklar giderilmiştir. Oluşturulan mantıksal model Şekil 4’de sunulmaktadır. Burada dikkat edilmesi gereken tekrar yeni uyşumsuzluklar oluşturulmamasıdır.



Şekil 4: Mantıksal Model



Şekil 3: Sel felaketi bilgi sistemi verileri analizi (Raclot, 2003)

3. SONUÇ

Araştırmaların bu gün geldiği nokta, farklı kalite seviyelerin 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.

Mekansal veri kalitesi araştırmalarının önerisi, üretilen mekansal verilerin kullanıcının amaçlarına ve isteklerine göre düzenlenmesi ve sunulmasıdır. Genellikle farklı disiplinler ile birlikte çalıştığımız mekansal bilgi sistemlerinde kullanıcılar kendi problemlerinin çözülmesini beklemekte ve verinin üretim şekli ve kalitesi ile ilgili problemleri göz ardı edebilmektedir. Örneğin turizm amacıyla oluşturulacak bir bilgi sisteminde konum doğruluğu düşük ancak öznitelik doğruluğunun yüksek olması beklenmektedir. Ancak imar planlarına yönelik bir bilgi sisteminde hem geometrik hem de tematik doğruluğun yüksek olması beklenmelidir. Mekansal verilerin sıklıkla kullanıldığı sektör çeşitliliği, konum bazlı sistemler, ulaşım, çevresel etkinin araştırılması, kadastro vb., düşünüldüğünde mekansal verilerin çeşitliliği, dolayısı ile kalitesindeki çeşitlilik daha iyi anlaşılabilir. Bu aşamada, kullanım amacına ilişkin araştırmalar yapılmalı, bu bilgilerin ulusal bazda olmaması halinde ulaşılabilir uluslararası standart ve/veya doğruluk ölçütleri kaynakları incelenerek öneriler getirilmelidir. Ülkemizde de diğer ülkelerde olduğu gibi mekansal

veri kullanan sektörlerde kullanıcı isteklerini belirlemek amacı ile analizler yapılmalı ve mekansal veri daha ekonomik ve işlevsel şekilde kullanıcılara sunulmasının yöntemleri araştırılmalıdır. Bu çalışmalara bir örnek, mekansal veri konusunda genellikle fazla deneyimli olmayan kullanıcılar için farklı veri tabakaları, nesnelerinin (ulaştırma ve çevresel etki, abone bilgisi ve sokak adı vb) entegre edilerek kullanıcıya sunulması ve kullanıcı istekleri doğrultusunda karar aşamasında destek vermek üzere gerekli veri kalitesi çalışmalarının yapılması olabilir. Hem üreticinin, hem de kullanıcının faal olarak yer aldığı mekansal veri üretimi ihtiyaç duyulmaktadır. Kalite kavramına daha esnek olarak yaklaşmak mekansal veri üreticileri için yeni olanaklar meydana getirecektir.

KAYNAKLAR

Krek, A., Frank, A. U., 2000, *The Economic Value of Geo Information*, Geo-Information-Systeme - Journal for Spatial Information and Decision Making 13(3): 10-12.

Tayi, G. K., Ballou, D. P., 1998, *Examining data quality*, Communications of the ACM. 41 (2): 54-57.

Redman, T., 1996, *Data Quality for the Information Age*, Boston, London, Artech House.

Wang, R. Y., Strong D. M., 1999, *An information quality assessment methodology*, Proceedings of the International Conference on Information Quality, Cambridge MA.

Naumann, F., 2002, *Quality-Driven Query Answering for Integrated Information Systems*, Berlin, Springer Verlag.

Jahn, M., Frank, A.U., 2004, *How to increase usability of spatial data by finding a link between user and data*, Proceedings AGILE 2004, p.653-661

Sheth, J. N., 2000, *Consumer Behavior, Marketing Best Practice*, M. R. Czinkota. Harcourt College, Harcourt College Publishers: 136-176.

Raclot D., Puech C., 2003, *What does AI contribute to Hydrology? Aerial photos and Flood levels*, Applied Artificial Intelligence 17(1): 71-86

Raclot D., 2004, *Méthodologie d'exploitation quantitative des photographies aériennes d'inondation de plaine*, UMR 3S Cemagref/ENGREF, Uni. Montpellier II, PhD thesis, <http://tel.ccsd.cnrs.fr/documents/achieves/00/00/32/57>

URL 1, Uluslar arası Standartlar Organizasyonu (ISO), 2004, <http://www.iso.org/iso/en/ISOOnline.frontpage>, 10 Mart 2005

URL 2, REVIGIS Projesi, 2004, <http://www.lsis.org/REVIGIS/Full/index.html>, 10 Mart 2005