**5. VERi KALiTESi**

**5.1 GİRİŞ**

İnsanlar elde ettikleri verinin kalitesi üzerinde oldukça fazla yorum yaparlar. Bir doktor arkadaşım muayenehanesindeki kayıtlarını otomatikleştirmek istediğinden bahsetti. Bilgisayar çıktılarından bu bilgileri kimin girdiğini anlamak istiyordu. Bazı asistanlarının diğerlerinden daha duyarlı ve hassas olduğunu biliyordu. El yazısından, hangi asistanın veriyi kaydettiğini tanıyabiliyor ve girilen bilgiyi buna göre yorumluyordu. Veri kalitesine, veriyi giren asistanına göre karar veriyordu. Verinin kalitesi el yazısında kodlanmıştı. Benzer şekilde, uzun yürüyüşler yapanlar topoğrafik haritalarda patikaların bulunduğu yerlerin, ana yolların bulunduğu yerlere göre daha az doğru olduğunu farkederler. Patika ve yolların birbirlerine göre kalitesi konusunda verecekleri karar harita verisini kullanmalarında yol gösterir.

Veri kalitesini bilmek, ve buna göre uygulamalara karar vermek için kritik düzeyde önemlidir. Haritaları üstüste bindirerek konumsal analizler yapıldığında üstüste gelmesi gereken sınırları hizalamak için haritayı yavaşça kaydırmak gerekir. Üstüste binen haritalar tam doğru olarak çakıştırılamaz, ancak elle ayarlamalarla harita kaydırılarak herhangi bir bölgesel alan eldeki iş için yeterince yakın olarak çakıştırılabilir. Veri kalitesinin bu şekilde ve doğru olmayan bilgilere izin verilmesi otomatikleşmiş GIS kullanıldığında sona erer. Konumsal hatanın neden olduğu hizalanmama CBS verisini kullanırken göz önüne alınması gereken veri kalite konularından birisidir.

Veri kalitesini ölçme maliyeti gereken emek derecesiyle değişir. Veri kalitesi kontrolu ne kadar titizlikle yapılırsa, maliyeti o derece artar. Maliyet sadece bir kontrolün yapılmasının masrafì değil, fakat kontrolleri yapmak ve hataları düzeltmekten dolayı üretim işleminde neden olunan gecikmelerden kaynaklanır. Bu nedenle, kontrol seviyesi, daha az doğru veri veya daha az titiz olarak doğrulanmış veri kalitesinden ötürü doğacak sonuçların maliyetiyle dengelenmelidir. Gerçekte gerekenden daha yüksek düzeyde kalite talep etmek gereksiz bir harcama olur. Veri kalitesi değerlendirilmesi ve neticeleri arasındaki ilişkiyi açıklamak için genel maksatlı kalın sicim ve tırmanma ipi için sağlanan kalite değerlendirmelerini gözönüne alalım. İnsanlar genel amaçlı kalın sicimleri, gerilme kuvvetinin zamanla bozulacağını bilmeden evdeki ihtiyaçlar için kullanırlar. Kalın sicimin sağlamlığına yanlış hükmetmenin neticeleri genelde ciddi olmaz, insanlar şüphelenirlerse daha çok kalın sicim kullanırlar. Fakat bir tırmanıcının hayatı emniyet ipinin özelliklerine bağlı olabilir. Tırmanma ipleri çok titiz kalite kontrolünden geçer ve ayrıntılı kalite kontrol neticeleriyle satılır. Hatadan doğacak sonuçlar ciddi olduğu için ipin karakteristiklerinin titiz olarak kontrolü, kalite kontrolü masrafìna değer bulunmaktadır. Benzer şekilde, bir CBS'de veri kalitesinin kaydı ve kontrol masrafı, elverişsiz kullanılışından ötürü doğacak neticelere uymalıdır.

CBS'deki veri, sayısal olarak saklanmayan veriye göre daha yaygın analizlerde kullanılabilir. Gerçekten, CBS kullanılması için bahsedilen üstünlüklerden birisi budur. Daha önceden bir arada incelenemeyen ayrı veri kümelerini birleştirme yeteneği CBS'de bulunur. Fakat veri ücreticisi tarafìndan önceden düşünülmemiş şekillerde ve kullanıcıların uygulamanın elverişli olup olmayacağı konusunda karar verecek bilgi ve tecrübeye sahip olmadığı durumlarda kullanılabilir.

Wisconsin'de bir toprak sahibi, standard topoğrafik haritadaki bir göl kıyısında yüksek su işareti yanlış gösterildiği için eyalet aleyhine mahkeme açtı. Bu durumda (8. bölümde anlatılıyor) kullanıcı, bu çeşit bir topoğrafik haritanın, , toprak parsel sınırlarını yeterince doğru göstermediğini fark etmedi. Bunun bir neticesi olarak, arazi sahibinin topoğrafının parçası yüksek su işaretinin altında kaldı. Eyalet kanunlarına göre yüksek su işareti altında kalan arazi eyaletin malı oluyordu. Hatanın bunun ertesinde düzeltilmesine rağmen, arazi sahibi hasarıyla zararları için dava açtı.

Bu bilinen kalitede bir harita olan, A.B.D. Coğrafik Araştırma Dairesinin topoğrafik haritasının, bilinmeyen kalitede (yüksek su seviyesi) bir verinin gösterilmesinde kullanılmış bir örnektir. Birleşik veri kümesinin kalitesi konusunda doğru olmayan bir tahmin yapılmıştır. Coğrafik bilgi üreticileri, verilerine akla uygun bir güvenin mali kayıplara veya kişisel incinmelere neden olması durumunda, bundan artan ölçüde sorumlu tutulmaktadırlar.

Coğrafik verinin kalitesi, çoğu zaman, yanlış karalar verildikten maddi kayıplar ve kişisel zararlar olduktan sonra, incelenir. Coğrafik bilgi üreticileri, ürünleri hatalı bulundukça, ürün tasarımı kötüyse veya ürün tasarımcının aklına gelmemiş amaç ve şekillerde kullanılmışsa, sorumlu tutulmaktadır. Veri kalite standardları uygun bir şekilde tanımlanır, konrol edilir ve raporları çıkarsa, coğrafik bilgiyi hem üreten hem de kullanan kişiler korunabilir. Veri, standard formatda tanımlanmış ve kabul edilmiş kalite seviyesinde sunulursa, üretici uygun olmayan kullanışlardan ötürü sorumluluktan kurtulur. Bu standardlar aynı zamanda kullanıcıyı da elverişsiz bilgiye güvenmeye karşı korur.

Bir CBS sistemi, çeşitli kabiliyetleri olan kullanıcılar tarafìndan ve daha çeşitli uygulamalar için kullanılmak üzere coğrafi bilgi imkanları sağlar. Bu verilerin karar vermede kullanılabilmeleri için kaliteleri bilinmeli ve tahmin edilmelidir.

CBS veri kalitesinin tanımlanması, değerlendirilmesi ve raporlarının hazırlanması konusunda en büyük katkı Sayısal Haritacılık Veri Standardlarının Ulusal Komitesinin Veri Kümesi Kalite Çalışma Grubu tarafìndan yapılmıştır.

Bu bölüm, coğrafik bilgi kullanılırken göz önünde bulundurması gereken veri kalitesi etkenlerini özetlemektedir. Her etkenin önemi ve rapor edilmesi için gereken ayrıntı, CBS içinde kullanıldıkları kavrama bağlı olarak değişecektir. Veri girişi başlamadan önce, bir CBS veri tabanındaki veriden istenen kalite standardları ve kalite ölçümü için kullanılan metodlar açıkça tanımlanmalıdır. Bu veri kalite standardları kullanıcıların gereksinimlerine cevap vermelidir, böylece kullanıcılar, bir CBS veri tabanı için veri kalitesi standardlarının belirlenmesine ve bütçe, teknik olanaklar ve üretim hızı gibi konularda sözsahibi olmalıdırlar.

**5.2 VERİ KALİTESİNİN BİLEŞENLERİ**

Verinin kullanışlılığını etkileyen karakteristikler 9 bileşene ayrılabilir. Bunlar da kendi içinde mikro düzeyde bileşenler, makro düzeyde bileşenler ve kullanış bileşenleri olmak üzere 3 kategoride gruplanmışlardır.

**5.3 MİKRO DÜZEY BİLEŞENLER**

Mikro düzey bileşenler veri elemanlarına ilişkin veri kalitesi etkenlerdir. Bu bileşenler genelde daha yüksek kalitede bağımsız bilgi kaynağıyla istatistik kontrol edilerek değerlendirilir. Konum doğruluğu, nitelik doğruluğu, mantıksal kararlılık ve ayrım gücünü içine alìr.