

TOPOGRAFİK VERİ ÜRETİMİNDE VERİ MODELLERİ ARASINDA DÖNÜŞÜM

D. Boyacı, M. Erdoğan, A.Torun, H.Ardıç, Z. N. Aksoy, A. Diri, Z. Can, E. Önal

Harita Genel Komutanlığı, Ankara, Turkey – (dijle.baysal, mustafa.erdogan, abdulvahit.torun, hayrullah.ardic, nejat.aksoy, ahmet.diri, zabit.can, erdal.onal)@hgk.msb.gov.tr

ANAHTAR SÖZCÜKLER: Kavramsal model, mantıksal model, semantik model, topografik veri, topografik veritabanı, veri modeli, veri modeli dönüşümü.

ÖZET:

Topografik veritabanlarının oluşturulması, veri üretiminin öncesinde tutarlı bir semantik model, kavramsal model ve mantıksal veri modelinin oluşturulmasını gerektirmektedir. Amaca uygun olarak hazırlanan mantıksal veri modelinin fiziksel veritabanı modeline dönüşümünün ardından topografik veri üretimi gerçekleştirilebilir. Topografik veritabanının doldurulması olarak da tanımlanan topografik veri üretimi, söz konusu hiyerarşik modellere uygun olarak belirlenen kaynaklardan veri çıkarımı, bütünleştirilmesi ve kalite kontrol süreçlerini içermektedir.

Bu çalışmada; Harita Genel Komutanlığınca MGCP (Multinational Geospatial Co-production Program-Çok Uluslu Coğrafi Veri Ortak Üretim Programı) veri üretiminde Program içinde zamana bağlı olarak ihtiyaç duyulan veri modelleri arasında ve bunun fiziksel gerçekleştirimi olan veriler arasında dönüşüm için geliştirilen yaklaşım tartışılmaktadır. Ayrıca, mevcut veri modeli ile hedeflenen veri modeli arasındaki dönüşüm, veriler arası dönüşüm ve bunların üretim yapılan teknolojik platformda gerçekleştirimine dair deneyim paylaşılmaktadır.

ABSTRACT:

The creation of topographic databases, prior to the production of data requires the creation of semantic model, conceptual model and logical data model. Prepared in accordance with the purpose of logical data model, physical database model transformation can be performed after the production of topographic data. Defined as the topographic database filling, topographic data production includes the extraction from data sources, data integration, and quality control processes determined in accordance with the hierarchical models.

In this study, conversion of the data model and the physical implementation of this conversion in MGCP (Multinational Geospatial Co-production Program) program of data generation, which are needed by the General Command of Mapping with the time in the program, is discussed. In addition, experiences about the data model transformation between the existing data model and the targeted model, the conversion between the data and technological platform for the production are shared.

KEY WORDS: A conceptual model, logical model, semantic model, topographic data, topographic database, data model, data model transformation.

1. GİRİŞ

Ulusal, bölgesel ve küresel coğrafi kaplamalı topografik veritabanlarının oluşturulması, veri üretiminin öncesinde tutarlı bir semantik model, kavramsal veri modeli ve mantıksal veri modelinin var olmasını gerektirmektedir. Bu hiyerarşik tanımlamaların ardından mantıksal veri modelinin bir teknoloji yardımıyla veritabanı bütünlüğü, tutarlılığı ve koşullarını da içeren fiziksel veritabanı modeline dönüşümü gerçekleştirilmektedir. Topografik veritabanının doldurulması olarak da tanımlanan topografik veri üretimi, söz edilen hiyerarşik modellere uygun olarak belirlenen kaynaklardan veri çıkarımı, bütünleştirilmesi ve kalite kontrol süreçlerini içermektedir. Vektör veri modelleme ve kalite kontrolü konularında literatürde farklı çalışmalar mevcuttur. Erdoğan vd. (2012) tarafından yapılan çalışmada vektör verilerin kalitesinin yükseltilmesinde kullanılabilecek kontrol aşamaları ve elde edilen sonuçlar incelenmiştir. Brown vd. (2005) dört anahtar ilişkinin (kimlik, nedensellik, zamansal ve topolojik) coğrafi veriyi ve modeli nasıl etkilediğini araştırmıştır.

Semantik model, modelin çerçevesini oluşturan küçük dünya, bunun içinde var olan temel sınıflandırma ile bu sınıfların

tanımlarından oluşmaktadır. Daha anlamlı bir kavramsal veri modeli oluşturmak için zengin bir semantik modelin oluşturulması gereklidir (Peckham ve Maryanski, 1988). Kavramsal veri modeli sınıfların içinde yer alan varlıklar ile bunların ilişkilerinin tanımlandığı modeldir. Kavramsal modelleme veritabanının tasarımı, geliştirilmesi ve optimizasyonu için çok önemli bir adımdır (Franconi ve Scattler, 1999). Mantıksal model bir teknolojiye bağlı olmamakla birlikte tüm varlıkların özellikleriyle birlikte bütünlük, koşul ve tutarlılık ilişkilerinin tanımlandığı ve fiziksel modele dönüştürülebilir bir modeldir. Mantıksal model çoğunlukla, varlık, varlık tanım kümeleri, varlıklar arası ilişkiler ve koşulların yer aldığı bir veri sözlüğü şeklindeki gerçekleştirimdir. Son aşama olan fiziksel model ise mantıksal modelin, bir teknoloji üzerinde gerçekleştirimidir. Tüm bu tanımlamalarla birlikte veri modelleme, topografik veritabanı oluşturmak için temel rol oynar ve kullanıcının gerçek dünyayı nasıl algılayacağını kontrol eder (Goodchild, 1992).

Yukarıda tanımlanan semantik, kavramsal, mantıksal modelin, seçilen küçük dünyayı modellemede başarısı, fiziksel olarak gerçekleştirilen veritabanının doldurulmasındaki tutarlılık, doğruluk ve tamliğin sağlanmasındaki başarıyı etkilemektedir.

Şöyle ki; yeryüzündeki nesnelere, bu nesnelere sahip olabilecekleri öznitelikler ve bu özniteliklerin alabilecekleri değerler tanımlanan modellerde ne kadar iyi yansıtılırsa üretim sonucu elde edilen topografik veri de o derece yeryüzü ile uyumlu bir veri olacaktır. Veritabanını doldurmak için yapılan veri üretiminde tutarlılık, tamlık, doğruluk gibi veri kalitesini tanımlayan ilişki, kural ve koşullar ya otomatik ya da insan emeği ile kalite kontrol sürecinde sağlanmaktadır. Ayrıca verinin tutarlılığına ve kalite kontrolüne yönelik birçok tanımlamanın da yapılması gerekmektedir.

Topografik veri üretimi ve veritabanlarının doldurulması süreçleri, ülke ve kurumların gelenek ve yaklaşımlarına bağlı olarak farklılıklar göstermektedir. Veri üretiminde, temel girdi olarak stereo ya da mono görüntülerden fotogrametrik kıymetlendirme ile bilgi çıkarımı yapılabildiği gibi yersel tekniklerin kullanımı ya da diğer kaynak verilerin bütünleştirilmesinden de yararlanılmaktadır.

Oluşturulan veri modellerinde, üretim süreci içinde günün koşullarına göre amaç, ihtiyaçlar, teknoloji ve yaklaşımda yenilik gibi sebeplerden ötürü zamana bağlı olarak değişiklik ve tanımlanan yeni bir modele dönüşüm zorunlu hale gelmektedir. Mevcut veri modelinden, tanımlanan yeni bir modele dönüşüm yukarıda tanımlanan modeller hiyerarşisini izleyerek ve kayıpsız olarak gerçekleştirilebilir. Ancak bu durum gerçek uygulamalarda genellikle ya çok maliyetli olması ya da mümkün olmamasından dolayı büyük ölçüde başarılamamaktadır. İki model arasında tek yönlü dönüşümde tanım ve değer kümelerinin birebir olması gerekmezken, çift yönlü dönüşümlerde birebir ve tersinir koşulu sağlanmaktadır. Uygulamada, tek yönlü dönüşümde birebir koşulunu sağlamayan varlıklar için büyük oranda insan katkısı ve kararına ihtiyaç duyulmaktadır. Tek yönlü kayıpsız dönüşüm ancak semantik, kavramsal ve mantıksal modellemeler arasında paralellik olması halinde mümkün olmaktadır. Genellikle, iki model arasında her varlık için bire bir dönüşüm olmayabilir. Eski veri modelindeki tek bir varlık ya da varlık tanım kümesi yeni modelde birden çok varlığa karşılık gelebileceği gibi tersi durumlar da olabilir. Modelden bütünüyle çıkartılan ya da yeni eklenen varlıklar, öznitelikler ve tanım kümeleri de olabilmektedir. Tüm bu karmaşık durum dikkate alındığında veri modelleri ve veriler arası dönüşüm her durumun ayrı ele alındığı problem yönelimli bir çalışma haline almaktadır.

2. VERİ MODELİ DÖNÜŞÜMÜ

Bu çalışmada, Harita Genel Komutanlığınca MGCP (Multinational Geospatial Co-production Program-Çok Uluslu Coğrafi Veri Ortak Üretim Programı) veri üretiminde kullanılan veri modelinin zaman içinde meydana gelen değişiklikler nedeni ile yetersiz kalmasından dolayı yeni bir veri modeline dönüşümü için gerçekleştirilen çalışmalar anlatılmaktadır. Bu kapsamda çalışma; detay, öznitelik ve öznitelik değerleri bazında olan değişiklikler için ayrı ayrı değerlendirilmiş ve dönüşüm için yapılan çalışmalar anlatılmıştır.

2.1 Model Dönüşümünde Yaklaşım

Veri modelleri arasındaki dönüşüm genellikle tek atımda gerçekleştirilememekte ve çok sefer geri besleme, dönüşümün yeniden modellenmesini gerektirmektedir. Literatürde yer alan model dönüşümüne ilişkin çoğunlukla otomatik yapmayı sağlayan, bire-bir örtüşen teorik yaklaşımlar çoğunlukla uygulamada gerçekleştirilememektedir. Uygulamada veri dönüşümleri genellikle spiral geliştirme yöntemiyle

gerçekleştirilmektedir. Spiral teknikle; model dönüşümünün tasarlanması, planlanması, uygulanması, geri besleme alınarak yapılan iş paketlerinin güçlendirilerek tutarlılığının artırılması şeklinde bir yaklaşım benimsenmektedir. İlk döngüde genellikle alınan bir küçük küme için başarımlar sağlanmakta ve protip oluşturulmakta, kazanılan deneyim ile uygulama genişletilmektedir. Bildirinin konusu olan çalışmada, spiral teknik kullanılmıştır.

Kaynak ve hedef semantik modeller arasında, mini dünyaya bakışta değişiklik olmaması sebebiyle bir farklılık yoktur. Ancak, hedef kavramsal modelde yeni varlık sınıfları (detay kümesi) oluşmuş ve bazı detay kümeleri de modelden çıkarılmıştır. Hedef mantıksal modelde, kaynak modelden farklı olarak detay sınıfları arasında var olan ilişkiler, koşul ve kurallarda değişiklik gerçekleşmiştir.

Kaynak veri modeli ile hedef veri modeli arasında temel olarak semantik, kavramsal ve mantıksal model benzerlikleri korunmakla birlikte, veritabanı doldurmayı kolaylaştırmak ve kullanıcının veriden optimum yararlanmasını sağlayacak değişiklikler yapılmıştır.

Bu bölümde, öncelikle kavramsal modelde yeni ortaya çıkan detay kümeleri, bunların nitelikleri ve bu niteliklerin değer kümelerine ilişkin ortaya koyduğumuz yaklaşımımız tanıtılmaktadır. Ardından da mantıksal modelde yeni tanımlanan ya da yeni oluşturulan detay kümeleri sebebiyle ortaya çıkan yeni ilişkiler, koşullar ve kurallarda gerçekleştirdiğimiz değişiklikler sunulmaktadır.

2.2 Detay Dönüşümü

Detay dönüşüm çalışmasında eski ve yeni veri modelinde yer alan detaylar ve geometrileri incelenmiştir. Eski veri modelinde 174 tane detay, 224 tane detay alt türü mevcut iken, yeni veri modelinde 196 tane detay, 267 tane detay alt türü bulunmaktadır. Burada detaylarda ve detay geometrilerindeki değişim dikkate alınarak 5 farklı değişim sınıfı tanımlanmış ve aşağıdaki gruplar oluşturulmuştur.

- Birinci sınıf, eski ve yeni veri modelinde değişmeden korunan detayları,
- İkinci sınıf, değişmeden kalan ve sadece yeni geometri eklenen detayları,
- Üçüncü sınıf, eski modelde olup yeni modelde başka bir detayın öznitelikleriyle ifade edilen detayları,
- Dördüncü sınıf, eski modelde olmayıp yeni modele eklenen detayları,
- Beşinci sınıf ise eski modelde olup yeni modelden çıkarılan detayları tanımlamaktadır.

Bu sınıflandırma sonucunda eski ve yeni modeldeki detay ve geometri karşılaştırma sonuçları Tablo 1'de verilmiştir. Yukarıda tanımlanan detay sınıfları için detay ve detay alt türü tanımlarının yapılması gerekir. Burada detay ifadesi ile veri sözlüğünde farklı detay kodları ile tanımlanan bağımsız nesnelere tanımlanmıştır. Detay alt türü ise bir detayın farklı geometrilerini ifade etmektedir. Örneğin, orman EC030 detay kodu ile tanımlanmış bir detaydır. 1/50 000 ölçeğine göre eğer kapladığı alan 15.625 m²'den büyük ise alan, küçük ise nokta, uzunluğu 500 m'den uzun ise çizgi geometrisi ile toplanmaktadır. Bu makalede bu tür detayla ilişkin geometriler detay alt türü olarak ifade edilecektir.

Değişim Sınıfı	Sayısal Değer
1	161

2	8
3	8
4	32
5	8

Tablo 1. Sınıf değerleri

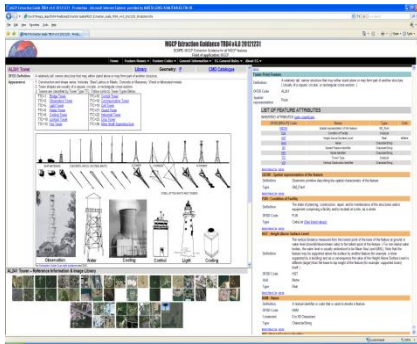
Tablo 1'de görüldüğü gibi 161 tane detay yeni veri modelinde değişmeden kalmıştır. Örneğin, AP010 yaz araba yolu detayı eski ve yeni modelde değişmeden korunan detaylar arasındadır. Bu detay öznelik, öznelik değerleri ve detay alt türünde hiçbir değişiklik olmadan yeni veri modeline aktarılmıştır.

İkinci sınıfta yeni geometri eklenen 8 tane detay tanımlanmıştır. Örneğin, AQ040 Köprü detayı eski veri modelinde çizgi geometrisi ile toplanırken yeni veri modeline alan geometrisi de eklenmiştir. Eğer köprü detayının kapladığı alan 5000 m²'den büyük ise alan geometrisinde toplanacaktır. Aynı şekilde DB160 kaya oluşumu detayı eski modelde alan ve nokta geometrisine sahipken, yeni modele çizgi geometrisi de eklenmiştir. Böylece sıra kayaların toplanması sağlanmıştır.

Üçüncü sınıfta tanımlanan ve yeni modelde başka bir detayın öznelikleri ile ifade edilen detayların sayısı 8'dir. Bunlar eski modelde bina içinde öznelik olarak tanımlıyken yeni modelde ayrılan 2 detay ve eski modelde ayrı 5 detayken kule detayı altında öznelik olarak birleştirilen detaylardan oluşmaktadır. Örneğin, AL240 iletişim amaçlı olmayan kule, AF030 soğutma bacası, AM080 su kulesi detayı eski veri modelinde ayrı birer detay olarak toplanmaktaydı. Ancak yeni veri modelinde bu detaylar tek bir detay adı altında (AL241 kule detayı) toplanmış ve detay türü TTC (Tower Type-Kule Tipi) özneliği ile belirlenmiştir.

TTC=1. Köprü	TTC=12. Yangın Kulesi
TTC=2. Gözlem Kulesi	TTC=16. Kontrol Kulesi
TTC=5. Navigasyon Kulesi	TTC=18. İletişim Kulesi
TTC=6. Su kulesi	TTC=19. GSM Kulesi
TTC=8. Soğutma Bacası	TTC=21. Nöbet Kulesi
TTC=10. İzleme Kulesi	

Tablo 2. AL241 Kule detayı ve TTC öznelik değerleri



Şekil 1. AL241 Kule detayı ve TTC öznelik değerleri

Dördüncü sınıfta tanımlanan eski modelde bulunmayıp yeni modele eklenen detayların sayısı 32'dir. Bu detaylar zaman içinde üretimde karşılaşılan eksiklikler neticesinde belirlenmiş ve yeni modele eklenmiştir. Örneğin, AT042 pylon, AQ065 menfez gibi detaylar yeni veri modeline eklenen detaylardır.

Son sınıfta tanımlanan 8 adet detay zaman içerisinde üretimde kullanılmadığı belirlenen ve modelden çıkartılan detaylardır.

Örneğin, AQ045 köprü ayak arası detayı bu sınıfta yer almaktadır.

2.3 Öznelik Dönüşümü

Detay kümelerini niteleyen öznelik dönüşümünde, her bir detayın eski ve yeni veri modelinde yer alan öznelikleri için tablolar oluşturulmuş ve detay bazında özneliklerde gerçekleştirilen değişim değerlendirilmiştir.

Yapılan çalışmada, öznelik bazında 2 farklı değişimin olduğu tespit edilmiştir. Bazı detayların bir ya da birden çok özneliği tamamen kaldırılırken, bazı detaylara da yeni öznelik ya da öznelikler eklenmiştir. Örneğin, GB055 pist detayı dışındaki tüm alan detaylardan LEN (uzunluk) ve WID (genişlik) öznelikleri kaldırılmıştır. Aynı şekilde AF010 baca detayının Tablo 3'de yer alan * işareti ile gösterilen 7 özneliği (AWP, LEN, NAM, NFI, NFN, SMC, WID) kaldırılmış ve detayın 4 öznelik ile tanımlanması sağlanmıştır.

GEOM	Özelliğin Konumsal Gösterimi (Spatial representation of the feature)
AWP*	Hava Uyarı Işığı (Air Warning Light Present)
FUN	Tesisin Durumu (Condition of Facility)
HGT	Yüzey Seviyesinden Yüksekliği (Height Above Surface Level)
LEN*	Uzunluk veya Çap (Length or Diameter)
NAM*	İsim (Name CharacterString)
NFI*	İsmlendirilmiş Özellik Tanımlayıcı (Named Feature Identifier)
NFN*	İsim Tanımlayıcı (Name Identifier)
SMC*	Yüzey Malzemesi Tipi (Surface Material Type)
VOI	Düşey Engel Tanımlayıcı (Vertical Obstruction Identifier)
WID*	Genişlik (Width)

Tablo 3. AF010 baca detayının öznelikleri

İkinci grup değişimde, örneğin, BH030 ark ve BH010 su kemeri detayına NAM (isim) özneliği eklenmiştir. İsim özneliği eski ve yeni veri modelinde diğer detaylarda yer alan özneliklerdendir. Yeni modelde ayrıca, eski modelde hiçbir detayda yer almayıp modele yeni eklenen MCC özneliği (Yapısal Malzeme Tipi - Structural Material Type) özneliği bulunmaktadır. Bu öznelik, BI020 baraj detayı gibi bazı detaylara yeni eklenmiş ve 22 farklı değer ile detayın yapı tipi tanımlanmıştır.

2.4 Öznelik Değerlerinin Dönüşümü

Detay kümesi özneliklerinin dönüşüm ilişkisi tanımlandıktan sonra, özneliklerin değer kümeleri (etki alanı - domain) arasındaki farklılık ve dönüşüm ele alınmıştır. Etki alanı, ilgili özneliğin alabileceği değerlerinin tümünü içeren bir değer kümesidir. Tax ve Duin (2004) iyi tanımlanmış bir değer kümesinin hedef verinin tamamını kapsaması ve hiçbir boşluğa yer vermemesi gerektiğini belirtmiştir. Son aşama olan öznelik değerlerinin dönüşümü çalışmasında, öznelik değerlerinin etki alanı ve iki model arasında bu değerlerde gerçekleştirilen değişimler incelenmiştir.

Çalışmada, detayların bazı öznelik değer kümelerine eklenen ve çıkartılan değerlerin olduğu tespit edilmiştir. Ayrıca, bazı

özniteliklerin detay bazında sahip oldukları değer kümeleri farklılıklarının standart hale getirildiği tespit edilmiştir. TRS (Transportation System Type), WLE (Water Level Effect), ACC (Horizontal Accuracy Category), FUN (Condition of Facility) ve MCC (Structural Material Type) özniteliklerinin alabilecekleri değer kümeleri tüm detaylar için standart hale getirilmiştir.

Örneğin, AF010 baca detayının FUN özniteliğinin değer kümesine önceki veri modelinde yer alan, ancak bu detayda bulunmayan “terk edilmiş” değeri ve önceki veri modelinde hiçbir detayın öznitelik kümesinde yer almayan “tasfiye edilmiş” değeri eklenmiştir. Yeni veri modelindeki FUN öznitelik kümesi Tablo 4’de gösterilmiştir.

FUN	Unknown (Bilinmeyen)	0
	Under Construction (Yapım Aşamasında)	1
	Abandoned (Terk edilmiş)	2
	Destroyed (Harap Olmuş)	3
	Dismantled (Çıkarılmış-Tasfiye edilmiş)	4
	Fully Functional (Tamamen İşlevsel)	6
	Damaged (Hasarlı)	13

Tablo 4. AF010 baca detayına ait FUN öznitelik değer kümesi

3. SONUÇLAR

Veri modeli dönüşümünde veri kaybını önlemek ve en doğru dönüşümü sağlamak için yapılacak modelleme çok önemlidir. Dönüşüme detay, detay geometrileri, detay öznitelikleri ve öznitelik değerlerinin dönüşümü olarak yaklaşılması en uygun yöntem olarak görünmektedir. Şöyle ki; yeryüzündeki nesnelere, bu nesnelere sahip olabilecekleri öznitelikler ve bu özniteliklerin alabilecekleri değerler tanımlanan modellerde ne kadar iyi yansıtılırsa üretim sonucu elde edilen topografik veri de o derece yeryüzü ile uyumlu bir veri olacaktır.

Detayların dönüşümünde öncelikle geometri ve özniteliklere bakılmaksızın değişmeyen detaylar belirlenmelidir. İkinci adımda, yeni eklenen ve modelden çıkartılan detaylar tanımlanmalıdır. Sonraki adımda, yeni geometri tanımlamalarının belirlenmesi uygundur. Burada amaç, birebir detay dönüşümünü tanımlamaktır. Sonraki adımda her detayın sahip olduğu öznitelikler incelenmeli ve modelden çıkartılan, yeni eklenen ve etki alanı değişen öznitelikler belirlenmelidir. Son adımda ise öznitelik değerleri incelenerek yine çıkarılan, eklenen ve tanımı değişen öznitelik değerleri incelenmelidir. Bu şekilde adım adım yapılan bir model dönüşümünde hatalar minimuma indirgenmekte ve veri kayıpları önlenmektedir.

KAYNAKÇA

Brown, D. G., Riolo, R., Robinson, D. T., North, M., Rand, W., 2005. Spatial process and data models: Toward integration of agent-based models and GIS. *Journal of Geographical System*, 7:25-47.

Erdoğan, M., Torun, A., Boyacı, D., 2012. Revisiting The Procedures For The Vector Data Quality Assurance In Practice. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XXXIX, Part B4, pp. 23-27.

Franconi, E., Scattler, U., 1999. A DataWarehouse Conceptual Data Model for Multidimensional Aggregation. *Proceedings of*

the International Workshop on Design and Management of DataWarehouses (DMDW'99), Heidelberg, Germany.

Goodchild, M. F., 1992. *Geographical Data Modeling*. Computers and Geosciences, Vol. 18, No. 4, pp. 401-408.

Peckham, J., Maryanski, F., 1988. *Semantic Data Models*. ACM Computing Surveys, Vol. 20, No. 3, pp. 153-189.

Tax, D. M. J., Duin, R. P. W., 2004. Support Vector Data Description. *Machine Learning*, 54, 45-66.