

Συλλογή – Επεξεργασία Δεδομένων

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Αγρονόμων & Τοπογράφων Μηχανικών - Μηχανικών Γεωπληροφορικής
Εργαστήριο Χαρτογραφίας



Ελένη Τομαή, Δρ. ΕΔΙΠ

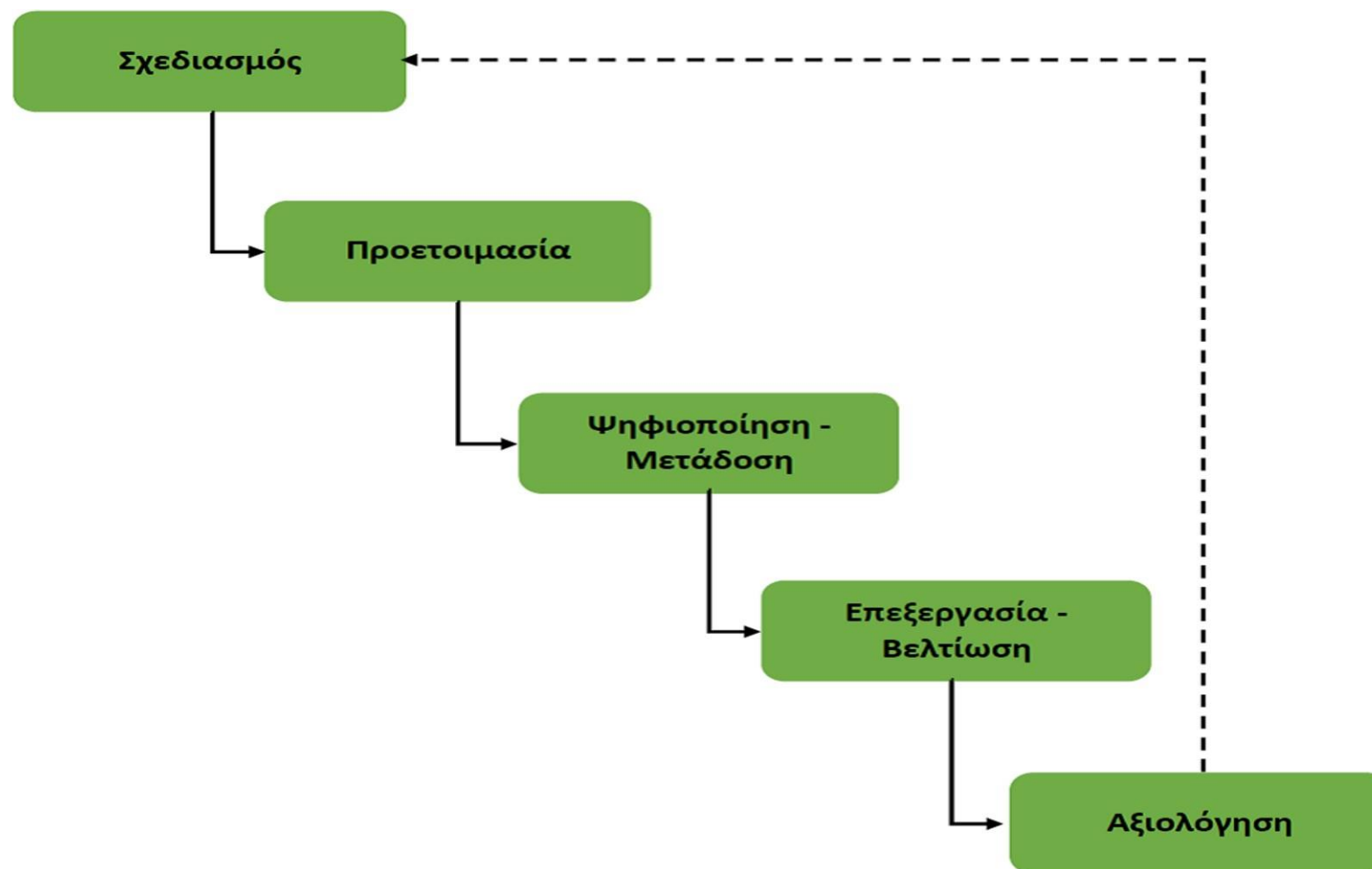
Περί Δεδομένων...

- Τα Συστήματα Γεωγραφικών Πληροφοριών είναι αχόρταγα!
 - Ενσωματώνουν δεδομένα προερχόμενα από **πολλές και διαφορετικές πηγές** χρησιμοποιώντας ποικίλες μεθόδους και τεχνολογίες διαχείρισής τους.
 - Συχνά, η συλλογή δεδομένων είναι **αυτοσκοπός**.
 - Τα δεδομένα είναι το **δαπανηρότερο** στοιχείο οποιασδήποτε εφαρμογής.

Συλλογή Γεωγραφικών Δεδομένων

- Αναφέρεται στον **εντοπισμό** και την **παραγωγή** των γεωγραφικών δεδομένων απαραίτητων για την υλοποίηση μιας συγκεκριμένης εφαρμογής ΣΓΠ.
- Πρόκειται για μια διαδικασία ιδιαίτερα σημαντική η οποία επηρεάζει άμεσα την ποιότητα και την χρησιμότητα των αποτελεσμάτων που θα παράγει το ΣΓΠ.

Διαδικασία Συλλογής Δεδομένων



Προβλήματα

- Όταν γεωγραφικά δεδομένα από διαφορετικές πηγές εισάγονται σε μια ψηφιακή ΒΔ, οι πληροφορίες για την **πηγή** τους, την **τεχνική συλλογής και επεξεργασίας** τους, καθώς και την **ακρίβεια** τους είναι σπάνια διαθέσιμες.
- Ένα σοβαρό πρόβλημα στην πρωτογενή συλλογή δεδομένων είναι η **διάκριση και οριοθέτηση** των αντικειμένων στην πραγματικότητα.

Τεχνικά Ζητήματα Συλλογής και Επεξεργασίας - I

- Το **σύστημα αναφοράς**
 - Έτσι ώστε η αναπαράσταση των γεωγραφικών συντεταγμένων σε έναν επίπεδο χάρτη να είναι εφικτή
- Η **κλίμακα**
 - Τα ψηφιακά δεδομένα τυπικά **ανεξάρτητα** της κλίμακας
 - Χρησιμεύει ως δείκτης της **ακρίβειας** των δεδομένων

Τεχνικά Ζητήματα Συλλογής και Επεξεργασίας - II

- Η Βάση Δεδομένων η οποία περιλαμβάνει δεδομένα από διαφορετικές πηγές, χάνει κάθε πληροφορία για την αρχική τους ακρίβεια
 - Ανάπτυξη μέτρων ακρίβειας και ελέγχων αξιοπιστίας
 - Περιγραφή της πηγής και ακρίβειας των στοιχείων της Βάσης Δεδομένων

Τεχνικά Ζητήματα Συλλογής και Επεξεργασίας - III

- **Συσχέτιση** των διανυσματικών (vector) και κανονικοποιημένων (raster) δεδομένων
 - Εργαλεία για τη μετατροπή και τη σύνδεση των δεδομένων
- **Επαναδειγματοληψία**, επιλύει προβλήματα όπως:
 - διαφορετικό μέγεθος των pixels
 - της διαφορετικής αρχής, προβολής και προσανατολισμού των κανονικοποιημένων δεδομένων
 - ύπαρξη αλλαγών (επικαιροποίηση των δεδομένων)

Εύρεση Δεδομένων

Πού;

- Δευτερογενή: υπάρχοντα δεδομένα
 - Ήδη δημοσιοποιημένα/ διαθέσιμα
 - Σε οργανισμούς
- Πρωτογενή: από το μηδέν
 - Παράγονται για τη συγκεκριμένη εφαρμογή (Οι εργασίες πεδίου είναι αργές και ακριβές!)

Χρόνος
και
Κόστος

Σε ποια μορφή;

- ψηφιακή
- αναλογική

Καταλληλότητα
και
Δυνατότητα
Εφαρμογής

Γνωρίσματα των Γεωγραφικών Φαινομένων

- Η θέση, το σχήμα και το μέγεθός τους (**γεωμετρικά/χωρικά** χαρακτηριστικά)
- Οι **μη χωρικές θεματικές** ιδιότητες που ονομάζονται χαρακτηριστικά (μορφή κειμένου)
- Η **μεταβολή** τους μέσα στο χρόνο
- Η **συμβολική αναπαράστασή** τους σε διαφορετικά **επίπεδα λεπτομέρειας** (**γραφικά/ χαρτογραφικά** χαρακτηριστικά)

Τύποι Δεδομένων

- **Χωρικά:** Περιγράφουν την απόλυτη και σχετική θέση των χωρικών οντοτήτων
- **Περιγραφικά:** Περιγράφουν τα χαρακτηριστικά των χωρικών οντοτήτων. Μπορούν να είναι **ποσοτικά** ή **ποιοτικά**, και συνήθως συναντώνται υπό μορφή πίνακα.

Δομές Δεδομένων

- **Κανονικοποιημένη (Raster)**
 - Χρησιμοποιεί ένα πλέγμα φατνίων (cells ή pixels) καθένα από τα οποία καθορίζεται από τη γραμμή και στήλη όπου βρίσκεται στο πλέγμα
 - Το μέγεθος του φατνίου καθορίζει την ακρίβεια και το επίπεδο λεπτομέρειας
- **Διανυσματική (Vector)**
 - Χρησιμοποιεί διανύσματα, που αποτελούνται από ένα ζεύγος συντεταγμένων (X, Y) , για την απεικόνιση και αποθήκευση των γεωγραφικών οντοτήτων
 - Στηρίζεται σε μορφές όπως σημεία, γραμμές και επιφάνειες για να αναπαραστήσει τα δεδομένα

Πηγές Δεδομένων

Οι πηγές δεδομένων μπορεί να είναι:

1. **πρωτογενείς**, όπου τα δεδομένα μετρούνται άμεσα (π.χ. τοπογραφικές μετρήσεις, τηλεπισκόπηση)
2. **δευτερογενείς**, όπου τα δεδομένα προέρχονται από χάρτες, άλλες Βάσεις Δεδομένων, πίνακες χαρακτηριστικών, κ.λπ.

Εισαγωγή Δεδομένων στα ΣΓΠ



ΣΥΝΟΠΤΙΚΑ...

- Χάρτες / Σχέδια
- Τοπογραφικές Μετρήσεις/ GPS
- Εναέριες (ή άλλες) φωτογραφίες
- Δορυφορικές εικόνες
- Βάσεις CAD δεδομένων
- Χωρικές (και μη) βάσεις δεδομένων
- Έγγραφα και κείμενα

Κατηγοριοποιώντας...

ΧΩΡΙΚΑ ΔΕΔΟΜΕΝΑ		
	Raster	Vector
πρωτογενή	<ul style="list-style-type: none">Ψηφιακές δορυφορικές εικόνεςLIDAR (3D μοντέλα επιφανείας)	<ul style="list-style-type: none">Τοπογραφικές μετρήσειςGPS
δευτερογενή	<ul style="list-style-type: none">Σαρωμένες αεροφωτογραφίεςDEMs από υψομετρικές καμπύλεςΣαρωμένοι χάρτες	<ul style="list-style-type: none">Ψηφιοποιημένοι χάρτεςΥφισταμένες Βάσεις CAD Δεδομένων

Χάρτης - I

- Η κυριότερη δευτερογενής πηγή δεδομένων
- Για να είναι χρήσιμος πρέπει να μετατραπεί σε ψηφιακή μορφή, μέσω της διαδικασίας που καλείται ψηφιοποίηση (digitization) η οποία είναι δύσκολη, χρονοβόρα και επιρρεπής σε σφάλματα
- Μειονέκτημα είναι η περιορισμένη και εξαρτημένη από την κλίμακα, ακρίβειά του

Χάρτης - II

- Επιπλέον, **δεν** αποδίδει με ακρίβεια τη θέση και το σχήμα των πραγματικών οντοτήτων, αφού αυτές έχουν υποστεί **διαδικασίες γενίκευσης και συμβολισμού**
- Η ποιότητα των δεδομένων επηρεάζεται από σφάλματα κατά την **επεξεργασία των δεδομένων, δηλαδή κατά τη συλλογή και αποθήκευση των δεδομένων**

Σάρωση

- Η διαδικασία μετατροπής αναλογικών εγγράφων σε ψηφιακή κανονικοποιημένη μορφή
- Επιτυγχάνεται μέσω μιας αυτόνομης περιφερειακής συσκευής, το σαρωτή

Σαρωτές

- Για χαρτογραφική χρήση, οι σαρωτές απαιτείται να έχουν ανάλυση τουλάχιστον **600 dpi** (μέγεθος **pixel 42 μm**)
- Οι περισσότεροι χάρτες σαρώνονται με ανάλυση **50 μm**
- Ο χρόνος σάρωσης εξαρτάται από την ανάλυση. Ως γενικός κανόνας ισχύει ότι ο χρόνος σάρωσης είναι ανάλογος του όγκου των δεδομένων

Διανυσματοποίηση

- Τα κανονικοποιημένα γραμμικά δεδομένα που προκύπτουν από σάρωση εικόνων χρησιμοποιούνται μόνο για αρχειοθέτηση, παρουσίαση και σχεδίαση
- Απαραίτητη η εξαγωγή γραμμικής πληροφορίας από τα κανονικοποιημένα δεδομένα, με μια μετατροπή που ονομάζεται **διανυσματοποίηση (vectorization)** ή μετατροπή raster - vector

Διανυσματοποίηση

- Η διανυσματοποίηση μπορεί να γίνει αυτόματα με το κατάλληλο λογισμικό χωρίς την ανθρώπινη επέμβαση.
- Η ποιότητα του παραγόμενου προϊόντος εξαρτάται από την ποιότητα και το περιεχόμενο του αρχικού εγγράφου.
- Απαιτείται κατάλληλη προ-επεξεργασία και μεθεπεξεργασία

Χειροκίνητη Ψηφιοποίηση - I

- Σκοπός
 - Η γρήγορη και ακριβής συλλογή ψηφιακής σημειακής και γραμμικής πληροφορίας από αναλογικά έγγραφα.
- Σύστημα συντεταγμένων
 - Τα δεδομένα δεν αποθηκεύονται με συντεταγμένες του συστήματος αναφοράς του ψηφιοποιητή, αλλά με τις πραγματικές επίγειες συντεταγμένες ⇒ **μετασχηματισμός**
 - Επιλογή, αναγνώριση και ψηφιοποίηση **σημείων ελέγχου**, γνωστών συντεταγμένων ⇒ **αγκίστρωση**

Χειροκίνητη Ψηφιοποίηση - II

- Σφάλματα της **χειροκίνητης καταχώρησης** δεδομένων συνδέονται με:
 - την **οπτική αντίληψη** και το **συντονισμό** ματιών-χεριού
 - την **υποκειμενικότητα** στην επιλογή των προς ψηφιοποίηση σημείων και
 - την **κούραση** του χειριστή
- Η επιλογή της κατάλληλης τεχνικής συλλογής εξαρτάται από:
 - την **φύση** των δεδομένων,
 - την απαιτούμενη **διακριτική ανάλυση**,
 - τις **διαθέσιμες πηγές δεδομένων** και
 - την **χρησιμοποιούμενη τεχνολογία**

Τοπογραφικές Μετρήσεις / GPS

- Άμεσος τρόπος συλλογής δεδομένων
- Τα δεδομένα που προκύπτουν περιγράφουν άμεσα ή έμμεσα την **θέση**, είτε με συντεταγμένες, είτε με άλλες μετρήσεις
- Σημαντική είναι και η συνεισφορά της τεχνολογίας του Παγκόσμιου Συστήματος Εντοπισμού (**GPS**) στη συλλογή δεδομένων, ιδιαίτερα για τις εφαρμογές που συμβαίνουν σε πραγματικό χρόνο

Τοπογραφικές Μετρήσεις



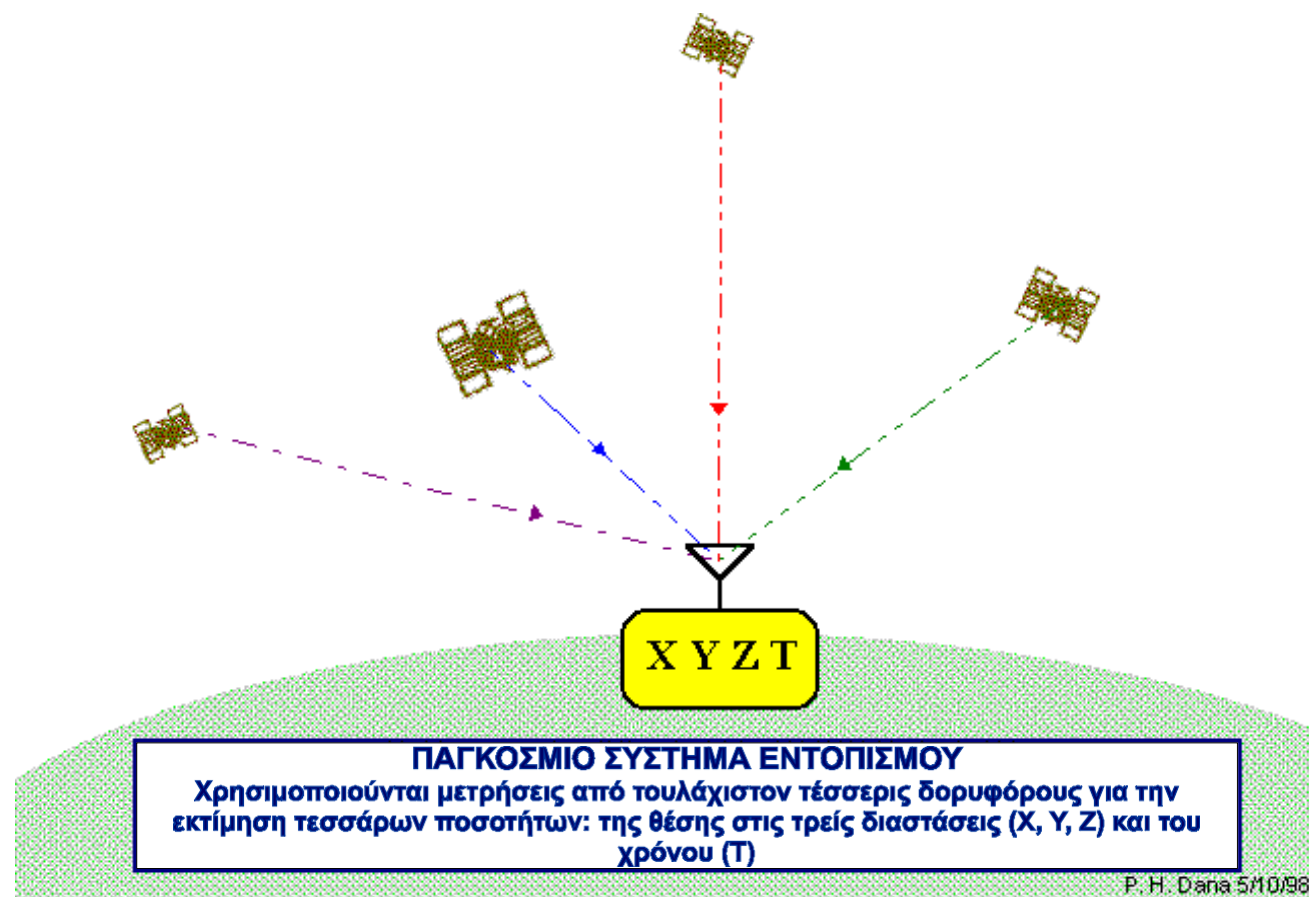
- Προσδιορίζεται η **θέση** (οριζοντιογραφικά και υψομετρικά) σημείων με μέτρηση διευθύνσεων και αποστάσεων
- Τα δεδομένα συνοδεύονται και από πληροφορίες για την **αναγνώριση και κωδικοποίηση των θέσεων**
- Τα δεδομένα εισάγονται στον υπολογιστή και επεξεργάζονται από ένα ΣΓΠ, ένα CAD ή άλλο πρόγραμμα

Παγκόσμιο Σύστημα Εντοπισμού - Ι



- Το σύστημα που έχει επικρατήσει είναι το **Navigation Satellite Timing and Ranging (NAVSTAR) Global Positioning System (GPS)**
- Η θέση ενός σημείου (ή οχήματος) προσδιορίζεται μετρώντας **αποστάσεις προς 4** τουλάχιστον δορυφόρους GPS -επίλυση ενός συστήματος τεσσάρων εξισώσεων με τέσσερις αγνώστους (**γεωγραφικό μήκος, γεωγραφικό πλάτος, ύψος, χρόνος**)

Παγκόσμιο Σύστημα Εντοπισμού - II



Παγκόσμιο Σύστημα Εντοπισμού - III

- Ο εντοπισμός με GPS μπορεί να πραγματοποιηθεί και σε **πραγματικό χρόνο**
 - Ο εντοπισμός με GPS μπορεί να είναι:
 - **στατικός**, όπου καθορίζεται η θέση ενός σημείου ή
 - **κινηματικός**, όταν προσδιορίζεται η τροχιά ενός σημείου
 - Οι συντεταγμένες ενός μόνο σημείου προσδιορίζονται με
 - **απόλυτο** ή
 - **διαφορικό εντοπισμό**
- χρησιμοποιώντας ένα δέκτη/ δύο δέκτες και μετρώντας ψευδο-αποστάσεις από τέσσερις ή περισσότερους δορυφόρους

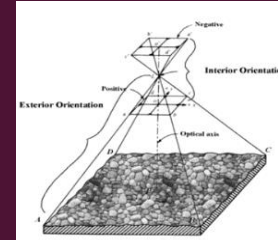
Παγκόσμιο Σύστημα Εντοπισμού - IV

- Το GPS αποτελεί μια ακριβή πηγή γεωγραφικών πληροφοριών με ιδιαίτερη χρησιμότητα για τα ΣΓΠ επειδή:
 - Λειτουργεί ως ένας **πραγματικός τρισδιάστατος ψηφιοποιητής** στο έδαφος
 - Μπορεί να χρησιμοποιηθεί για την παρακολούθηση **κινούμενων αντικειμένων**
 - Διευκολύνει τους **επίγειους ελέγχους**
 - Η τεχνολογία του GPS έκανε δυνατή την **ανάπτυξη εφαρμογών των ΣΓΠ σε πραγματικό χρόνο**, όπως είναι οι ηλεκτρονικοί χάρτες, η αποστολή οχημάτων και η διαχείριση στόλων.

Φωτογραμμετρία / Τηλεπισκόπηση

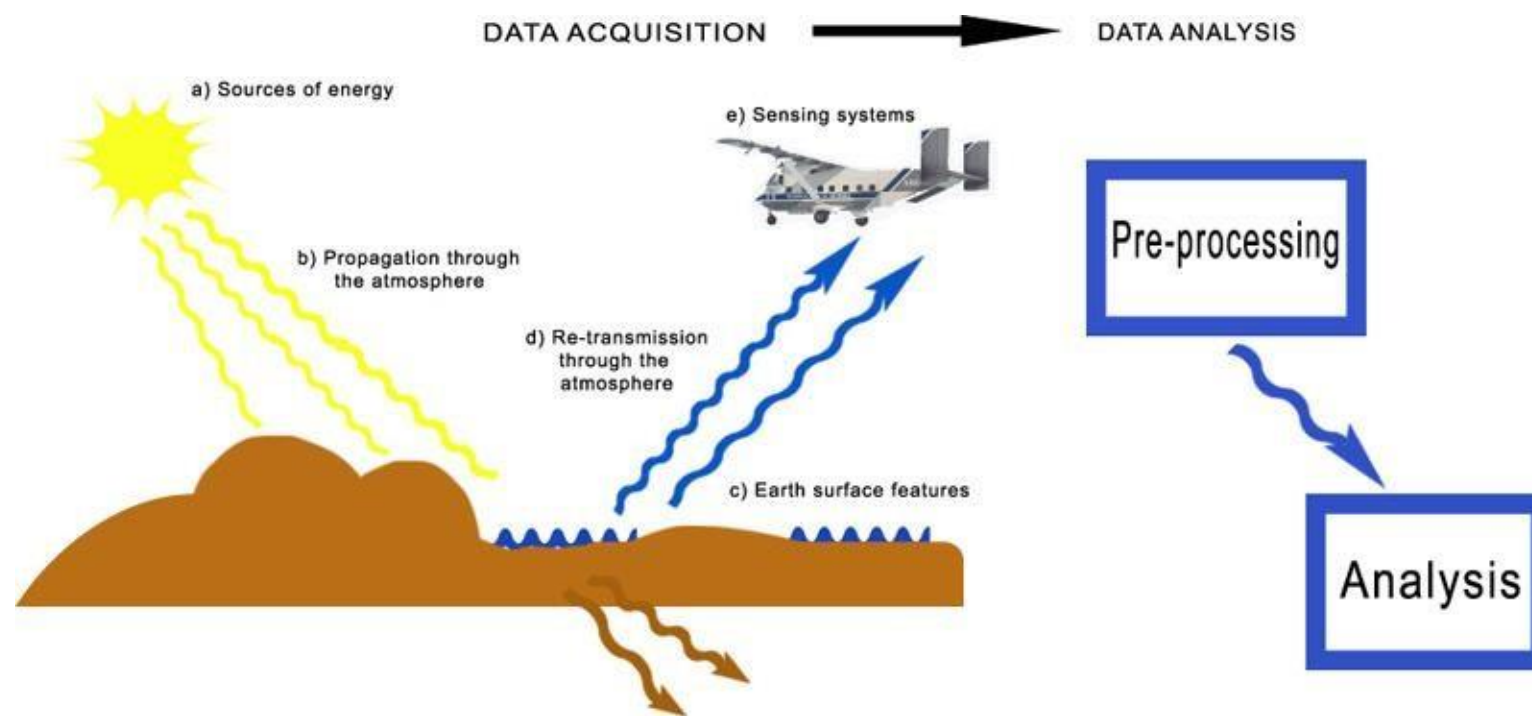
- Συστήματα που **δεν έρχονται σε άμεση επαφή** με τα αντικείμενα ενδιαφέροντος:
 - μέσω φωτογραφικών μηχανών που καταγράφουν σε ταινία, η οποία μπορεί έπειτα να σαρωθεί (κυρίως αεροφωτογραφίες)
 - μέσω αισθητήρων, που παράγουν άμεσα ψηφιακά προϊόντα (δορυφορικές εικόνες)

Αεροφωτογραφίες / Φωτογραμμετρία



- Οριζοντιογραφικά και υψομετρικά δεδομένα συλλέγονται με χρήση επικαλυπτομένων φωτογραφιών.
- Τύποι φωτογραφικών ταινιών:
 - μονοχρωματική (άσπρη & μαύρη)
 - φυσικού χρώματος
 - υπέρυθρη (πέρα του ορατού κόκκινου)
- Ψηφιοποίηση ή σάρωση για μετατροπή των αναλογικών φωτογραφιών σε ψηφιακές
- Διόρθωση / αναγωγή και DTM (ψηφιακό μοντέλο εδάφους) για τη δημιουργία ψηφιακών ορθοφωτογραφιών

Τηλεπισκόπηση - I



Τηλεπισκόπηση - II

- Οι δέκτες (sensors) τοποθετούνται σε πλατφόρμες (αεροπλάνα ή δορυφόρους) και καταγράφουν ανακλώμενη/ εκπεμπόμενη ακτινοβολία από τη ΦΓΕ σε μία ή περισσότερες ζώνες (κανάλια) του ηλεκτρομαγνητικού φάσματος
- Κατάλληλοι συνδυασμοί φασματικών καναλιών μπορούν να αποκαλύψουν ή να επισημάνουν σημαντικά φαινόμενα του περιβάλλοντος (π.χ. ύπαρξη νερού, μόλυνσης, ασθένειας βλάστησης)

Τηλεπισκόπηση - III

- Τύποι δεκτών:
 - **παθητικοί** (οι πιο κοινοί), καταγράφουν ηλεκτρομαγνητική ακτινοβολία προερχόμενη από μία φυσική πηγή (συνήθως ο ήλιος)
 - **ενεργοί** (radar), "φωτίζουν" (προσβάλλουν) οι ίδιοι το στόχο χρησιμοποιώντας την δική τους πηγή ακτινοβολίας (π.χ. RADARSAT του Καναδά, SIR-C/X-SAR της NASA)

Τηλεπισκόπηση - IV

- Άλλος διαχωρισμός:
 - **Μονοφασματικοί** καταγράφουν σε μία φασματική ζώνη και ανιχνεύουν την φασματική αντανάκλαση σε ένα μικρό τμήμα του ηλεκτρομαγνητικού φάσματος ή σε μια ευρύτερη περιοχή. Όταν καταγράφουν σε όλο το ορατό φάσμα και το εγγύς υπέρυθρο δίνουν **παγχρωματικές εικόνες**
 - **Πολυφασματικοί** (multispectral) καταγράφουν δεδομένα σε περισσότερες από μία φασματικές ζώνες
 - **Υπερφασματικοί** (hyperspectral) καταγράφουν την αντανάκλαση της ηλεκτρομαγνητικής ακτινοβολίας σε πολλές φασματικές ζώνες, στενού εύρους (10-20 νανομέτρων)



Landsat 7 (+ETM) Satellite Image
Acquired on: 06-AUG-1999
Resolution: 15 meter

SPOT 5 - 2.5 m - Berlin - Germany - August 2002

© Cnes 2002 - Distribution Spot Image - All rights reserved

SPOT
IMAGE





Copyright © 2007 GeoEye/EUSI. All Rights Reserved.
IKONOS satellite (0.8 meter resolution)
Faliro Coast, Greece
August 19, 2004



www.satimagingcorp.com

Τηλεπισκόπηση - V

- Η επεξεργασία των δεδομένων και η εξαγωγή χρήσιμης πληροφορίας μπορεί να γίνει με διάφορες μεθοδολογίες φωτοερμηνείας και επεξεργασίας εικόνας, λιγότερο ή περισσότερο αυτοματοποιημένες
- Για να επιτευχθεί μεγαλύτερο ποσοστό αξιοπιστίας, η διαδικασία αυτή σχεδόν πάντα εμπλουτίζεται με εξωγενή δεδομένα από χάρτες, αεροφωτογραφίες μεγαλύτερης κλίμακας, επιτόπιους ελέγχους, κ.λπ.

Τηλεπισκόπηση - VI

- Τα δεδομένα κωδικοποιούνται ως στοιχεία εικόνας (**picture elements-pixels**) σε ένα διδιάστατο πίνακα (raster)
- Η τιμή κάθε pixel αντιπροσωπεύει την ισχύ της ανακλώμενης ηλεκτρομαγνητικής ακτινοβολίας
- Η γεωγραφική αναφορά των κανονικοποιημένων (raster) εικόνων προηγείται της συλλογής των δεδομένων

Υφιστάμενες Βάσεις CAD Δεδομένων - I

- GIS (σύστημα βάσης δεδομένων) vs. CAD (σύστημα γραφικής απόδοσης)
 - GIS: τα γραφικά στοιχεία είναι απλώς οι αναπαραστάσεις των δεδομένων.
 - CAD: τα γραφικά στοιχεία είναι σημαντικά, δηλαδή το σχέδιο/ χάρτης είναι η πληροφορία.

Υφιστάμενες Βάσεις CAD Δεδομένων - II

- Παραδοσιακά, τα CAD δεδομένα αναφέρονται σε **συστήματα αρχείων** και όχι σε **συστήματα βάσεων δεδομένων**
- Συνήθως αναφέρονται στην **γεωμετρία των αντικειμένων** (geometry-based), αποτελούνται δηλαδή, από **γεωμετρικά πρότυπα** (π.χ. σημεία, γραμμές, πολύγωνα)

Υφιστάμενα ΣΓΠ

- Πολύτιμη πηγή (ψηφιακών ήδη) δεδομένων!!!
- Προβλήματα:
 - Το σύννηθες: άλλο format, προσπάθειες προτυποποίησης
 - FGDC's, Spatial Data Transfer Standard (SDTS),
 - Open GIS Consortium, πρότυπα για δημιουργία GIS βάσεων δεδομένων
 - Έλλειψη metadata
 - Ανάγκη επικαιροποίησης

Κτηματολογικές Βάσεις Δεδομένων

- Πηγή δεδομένων που χρησιμοποιείται, τόσο στη Διαχείριση Πληροφοριών Γης, όσο και το σχεδιασμό. Περιλαμβάνει πληροφορίες για:
 - τα γεωτεμάχια,
 - την ιδιοκτησία και
 - την αξία
- Σε αυτές τα γεωμετρικά δεδομένα:
 - αναφέρονται σε **μεγάλες κλίμακες** (1:1000 έως 1:5000) και απαιτούν **μεγάλη ακρίβεια**

Βάσεις Περιγραφικών Δεδομένων - I

Μεθοδολογίες συλλογής

- Απογραφές πληθυσμού
 - Στην Ελλάδα υποχρεωτικές (ΕΛΣΤΑΤ), καταγραφή όλου του πληθυσμού
 - Τακτικές αλλά όχι τόσο συχνές (ανά 10 χρόνια) ώστε να θεωρούνται πάντα επίκαιρες και σημεία αναφοράς
- Έρευνες επικαιροποίησης
 - Μη υποχρεωτικές, επικαιροποιούν τις ήδη υπάρχουσες απογραφές
 - Καλύπτουν περιορισμένη γεωγραφική έκταση

Βάσεις Περιγραφικών Δεδομένων - II

- Ειδικές Έρευνες
 - Μη υποχρεωτικές, παρέχουν πληροφορίες για θέματα που δεν καλύπτονται από τις απογραφές
 - Συχνά πραγματοποιούνται από μεγάλους εθνικούς φορείς (εκτός ΕΛΣΤΑΤ)
- Διάφορες άλλες πηγές
 - Διοικητικά αρχεία
 - Προσθετικές πληροφορίες σε ήδη υπάρχουσες προερχόμενες από έρευνες μεγάλης κλίμακας

Κείμενα και Εγγραφές

Αναλογική και ψηφιακή πηγή δεδομένων

- Αναλογική μορφή
 - Εισαγωγή στο ΣΓΠ με πληκτρολόγηση: αποδελτίωση/ ψηφιοποίηση
- Ψηφιακή μορφή
 - Επεξεργάζονται όπως οι βάσεις περιγραφικών δεδομένων

Δεδομένα: Οι Τάσεις

- Ανάγκη/ Απαίτηση για δωρεάν ΨΗΦΙΑΚΑ Δεδομένα
 - Έκρηξη της χρήσης Google Maps
 - Web 2.0 → Web 3.0
 - World-wide Projects!!! (OpenStreetMap: τεχνολογία και δομή wikis)
 - OGC
 - INSPIRE

Συνοψίζοντας

Βήματα Εισαγωγής Δεδομένων σε ένα ΣΓΠ

- Εντοπισμός δεδομένων
- Επίλυση προβλημάτων/ ελλείψεων στην υφιστάμενη πηγή δεδομένων
- Μετατροπή σε ψηφιακή μορφή
- Διαδικασίες ποιοτικού ελέγχου
- Ορισμός κριτηρίων τελικής αποδοχής δεδομένων στο σύστημα

Η τελική επιλογή αφορά τελικά στη σχέση:
Ποιότητας-Κόστους-Οφέλους