Felsőoktatási Földmegfigyelés tananyag- Fejlesztés

HunEdu Team (Ronczyk Levente)

DATelite Kft, Óbudai Egyetem, Soproni Egyetem

Fény-Tér-Kép, 2026.10.01.

Célkitűzés

- Háttér
- Projekt státusza
- Most futó feladatok
  - Minta tananyagok
- Hátralévő feladatok
- Egyéb tervek, további lépések

HunEdu

- Europai Űrügynökség által finanszírozott felsőoktatási (200k EUR) tananyagfejlesztő projekt.
- DATelite fővállakozó,
- Soproni Egyetem, Óbudai Egyetem, Wageningeni Egyetem alvállalkozó
- 2025.03 - 2027.03

Teljesített feladatok

- Több lépcsőben történt az oktatási tematika fejlesztés,
  - Nemzetközi és hazai szakértők bevonásával,
- Megkezdődött a próba oktatás
- és ezzel párhuzamosan a képzés népszerűsítése
  - További arculati elemek kidolgozása

Tematika

sid	Subject Name	Semester	Credit (adju
1	Remote Sensing fundamentals	1	5
2	EO and GIS data management	1	
3	Geostatistics	1	5
4	EO Cloud platforms	1	5
5	Geospatial Analysis and Geovisualization	1	5
6	Elective course	1	5
1	Earth Observation in Practice 1.	2	
2	Change detection	2	
3	Machine Learning	2	5
4	Photogrammetry	2	5
5	Proposal writing, project management	2	
6	Elective course	2	5
1	Earth Observation in Practice 2.	3	
2	Deep Learning	3	5
3	Advanced EO : Hyperspectral	3	5
4	Advanced EO: SAR	3	
5	Advanced EO : Laser scanning	3	5
6	Thesis plan	3	5

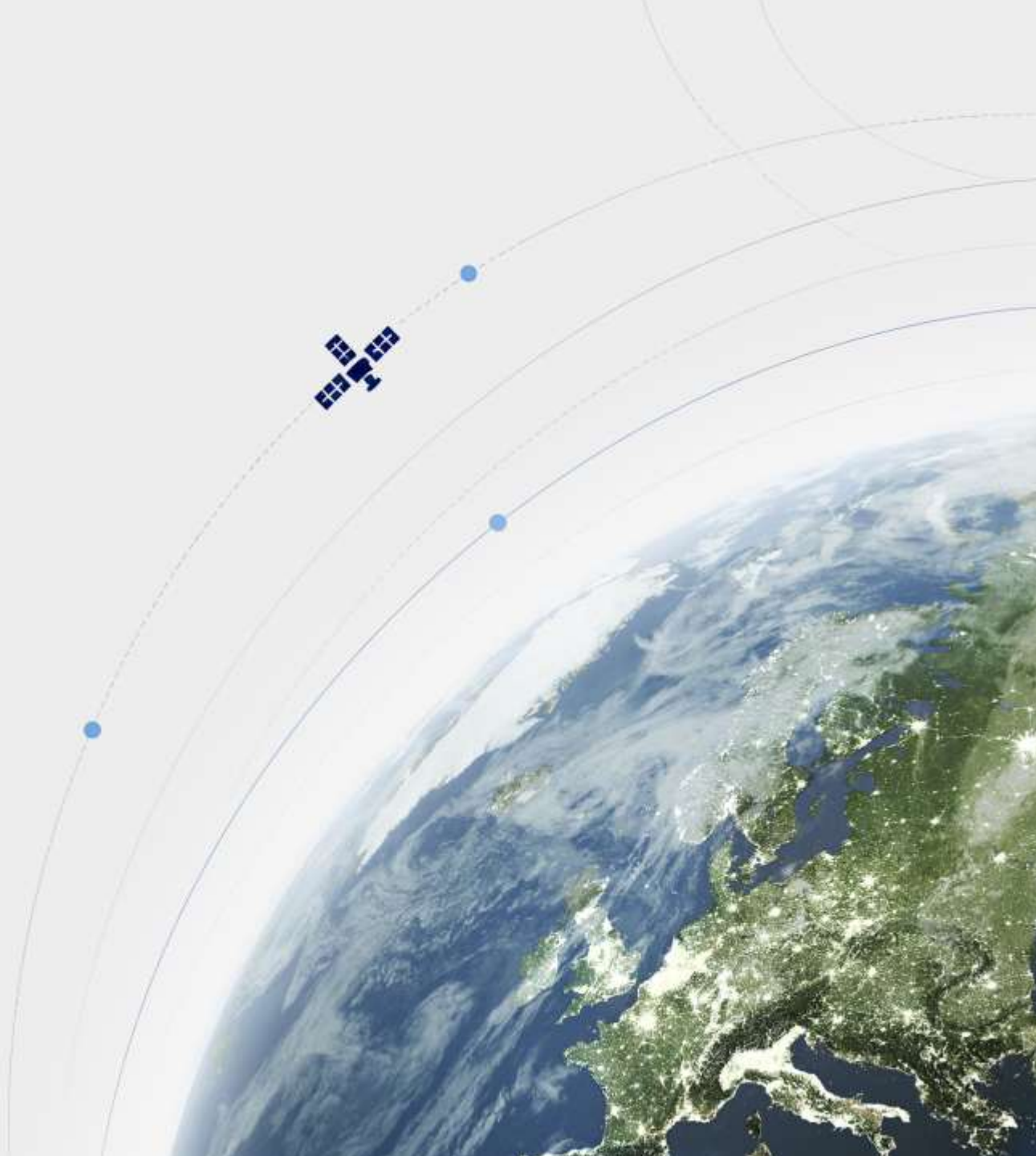
Gépi tanulás - tantárgyleírás

A tantárgyi program, 13 hétre osztott részletes tematika:

1. Nem felügyelt osztályozás: ISODATA
2. Nem felügyelt osztályozás: k-means, Maximális valószínűség
3. Gyakorlati vizsga
4. Felügyelt osztályozás alapjai
5. Tanító és tesztadatok kialakítása, osztálytávolságok
6. Jellemzők kiválasztása, optimalizálás, szóródási diagramok
7. Felügyelt osztályozási módszerek: Legközelebbi szomszéd, Maximális valószínűség
8. Gyakorlati vizsga
9. Felügyelt osztályozási módszerek: CART, Random Forest
10. Felügyelt osztályozási módszerek: SVM
11. Pontosságértékelés: Félreosztályozási mátrix, osztályozás pontossága
12. Pontosságértékelés: hibatérképek, osztályozás stabilitása
13. Gyakorlati vizsga

HUNEDU

Machine Learning



Kornél Czimber

University of Sopron, Hungary

14 July 2026

HUNEDU

02

Unsupervised Classification

Clusters, K-Means, ISODATA, Hierarchical

Unsupervised Classification

- Clustering: groups pixels with similar spectral, textural, or spatial properties.
- No labelling is required
- Iterative algorithms: k-means, ISODATA, DBSCAN
- Workflow:
  - Randomly Initialize K number of clusters
  - Assign every pixels to the nearest cluster
  - Recalculate cluster centres
  - Repeat until convergence

- **Unsupervised classification** is a type of machine learning in which a model groups or classifies data **without using labelled examples**. Instead of learning from known output classes, it identifies hidden patterns, similarities, or structures in the data.
- **Key characteristics of the unsupervised classification**
  - No labelled training data is required.
  - The algorithm discovers natural groupings (clusters) based on feature similarity.
  - Mainly used for exploratory data analysis and pattern discovery.
- Clustering is one of the oldest and most widely used techniques for **unsupervised image classification**, particularly in remote sensing, medical imaging, and computer vision. Instead of assigning pixels to predefined classes (e.g., forest, water, urban), clustering groups pixels or image objects with similar spectral, textural, or spatial properties.
- Common Clustering methods are K-Means, Hierarchical Clustering, DBSCAN.
- General workflow is extraction of image features (spectral bands, focal statistics, indices, texture, shape, elevation ...), feature normalization, clustering algorithm, cluster assignment and labelling, cluster analysis and merging, thematic mapping.
- K-mean algorithm input is the number of clusters. The algorithm randomly or linearly initialize K cluster centers, assigns every pixel to the nearest center, then recalculates cluster centers. It repeats until convergence.

K-means Clustering

- Randomly or linearly initialize K number of clusters
  - Random: K random cluster centres
  - Linear: along mean ± standard deviation vector
- Assign every pixels to the nearest cluster
  - Spectral or normalized distance calculation
- Recalculate cluster centres
  - Averaging
- Repeat until convergence
  - Cluster centres move each step less than stop

ISODATA method

- Lab work Overview
  - Launch Google Colab
  - Two-dimensional data generation
  - Scattergram display
  - ISODATA classification
  - Displaying the result
  - Analysing the result

EO Cloud Platforms

Balázs Böröcz

Óbuda University, 202x-xx-xx

HUNEDU

11

AlphaEarth Foundations Model

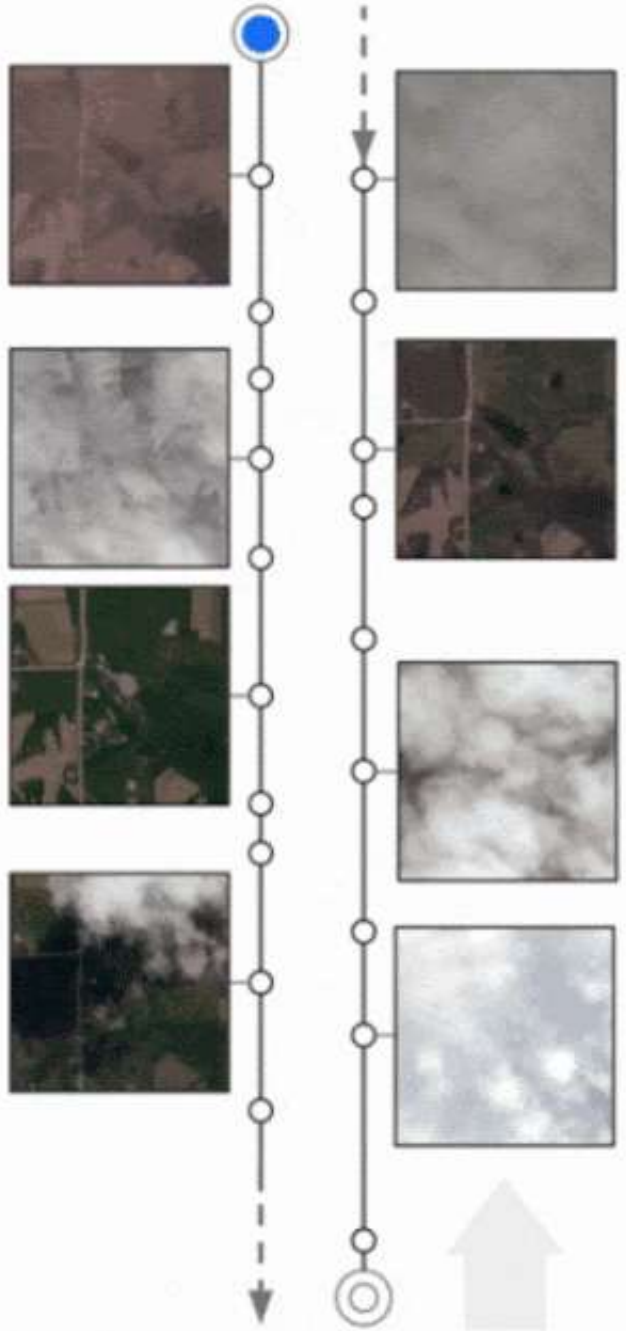
AEF model, embedding, Google DeepMind

From satellite data to geospatial embeddings

- Embedding models – Deep-learning models that transform complex, high-dimensional data into a compact numerical representation while preserving meaningful relationships between observations.
- From measurements to representations – Instead of working directly with individual spectral bands, an embedding captures information learned from many observations and data sources in a common feature space.
- AlphaEarth Foundations (AEF) – A geospatial foundation model developed by Google DeepMind that integrates large volumes of Earth-observation data into a unified representation of the Earth's surface.
- Multi-source input – The model learns from optical, radar, and other Earth-observation data, including temporal information, rather than relying on a single satellite image.
- 64-dimensional embedding – Each 10 × 10 m pixel is represented by 64 learned values, forming a single embedding vector. The individual dimensions do not correspond directly to physical quantities such as red or NIR reflectance.

From satellite data to geospatial embeddings

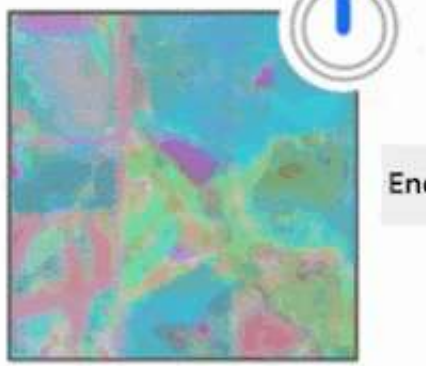
- Unit hypersphere – The 64-dimensional vectors are normalized to have a Euclidean length of 1, meaning that the embeddings can be interpreted as points on the surface of a 64-dimensional unit sphere.
- Similarity in embedding space – Locations with similar learned characteristics are represented by vectors pointing in similar directions, while different surface characteristics produce more distant directions.
- Beyond traditional indices – Instead of manually designing an index for a specific phenomenon, the embedding provides a general-purpose feature space that can be used for classification, regression, change detection, and similarity search.
- Analysis-ready data – Annual Satellite Embeddings are already generated and available in Google Earth Engine, so users can work with the learned representations without training the foundation model themselves.
- Key concept:
  - Multi-source Earth observation → Deep-learning model → 64-dimensional embedding → Similarity-based analysis



Multi-source Earth observations

AlphaEarth Foundations

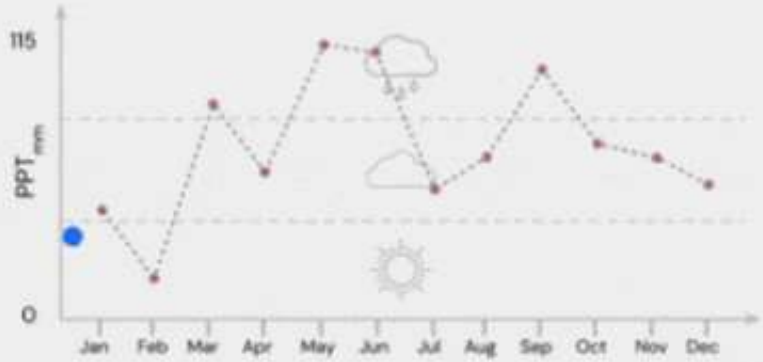
Date/Time



Annual embedding
-93.748°, 41.408°

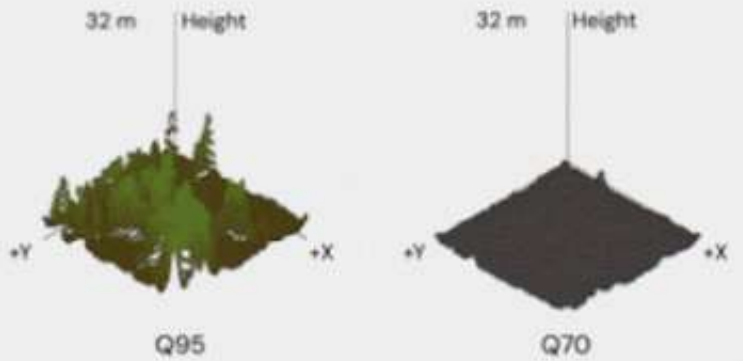
Encodes

Weather

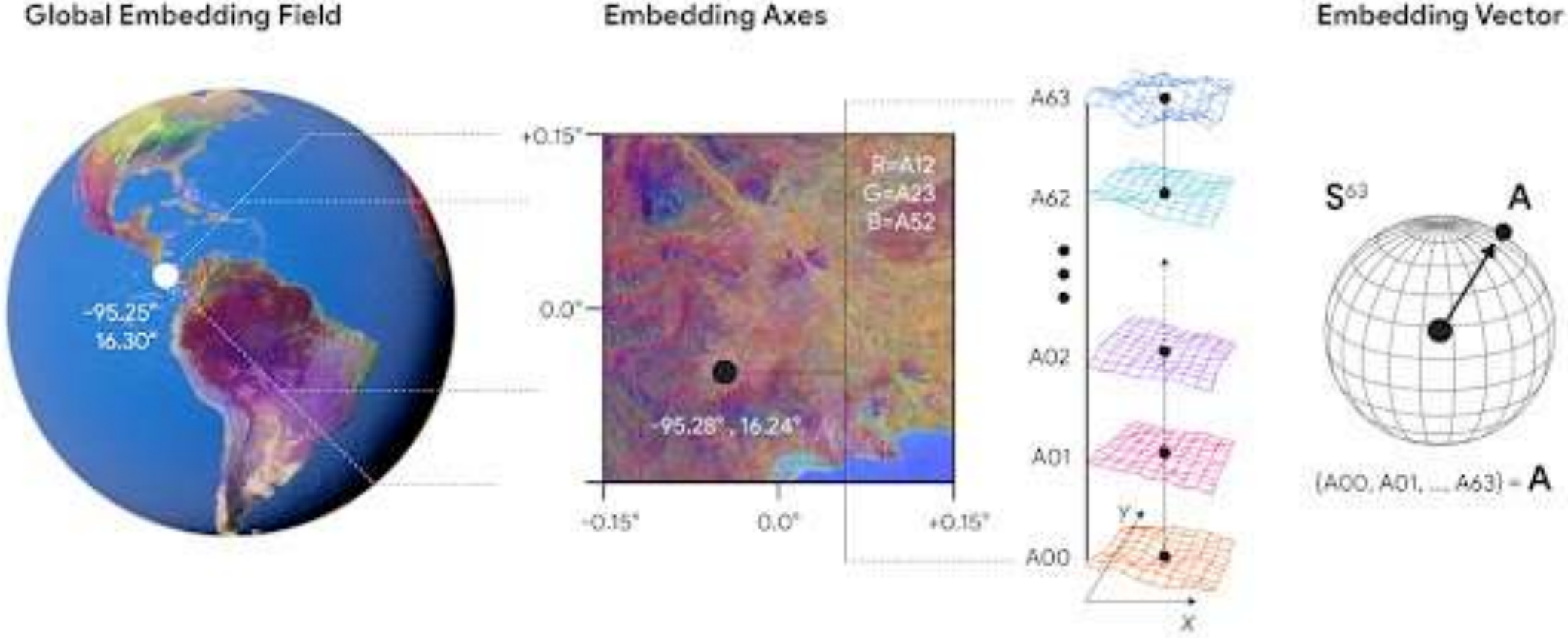


Optical

LiDAR

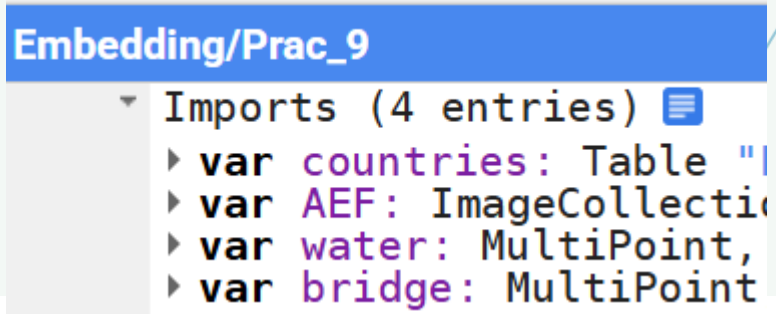


64-dimensional vector space



First steps

- **Independent task:**
  - **Create a new folder called Embedding.**
  - **Create a new file called Prac_9 inside the Embedding folder.**
  - **Import the ”LSIB 2017: Large Scale International Boundary Polygons, Simplified” database as countries**
  - **Import ”Satellite Embedding V1” as AEF**
  - **Create two geometries, one called bridge and the other called water.**



Hátralevő lépések

- Második körös próbaoktatás lezárása,
- Népszerűsítő események, rendezvények tartása,
- Tananyagfejlesztés lezárása,

További célok

- Hazai szakmai, intézményi konszenzus kialakítása az oktatási program bevezetésével kapcsolatban.
  - Mesterképzés akkreditálása és elindítása
    - Jelenleg formálódó megbeszélés tárgya
- Minisztérium tájékoztatás
- Nemzetköziesítés, Közép-Európai egyetemek (Ljubljana, Prága, Eszék, Temesvár, stb) bekapcsolása,

Köszönjük a figyelmet!