Education and Culture DG

Lifelong Learning Programme

Database Technology Professional Extended
DBTech EXT

Education and Culture DG

Lifelong Learning Programme

This presentation is the result of collaborative work undertaken along the lines of the DBTechNet Consortium. The speakers participate in DBTech EXT, a project partially funded by the EU LLP Transversal Programme (Project Number: 143371-LLP-1-2008-1-FI-KA3-KA3MP

Education and Culture DG

Lifelong Learning Programme

DBTech EXT

Virtual Laboratory Workshop

on

Business Intelligence & Knowledge Discovery from Databases (BI & KDD VLW)

Dimitris A. Dervos
Information Technology Department
Alexander Technology Educational Institute
Thessaloniki, Greece

BI&KDD VLW Overview (1)

A) The tutorial part

Video presentation and ppt slides on:

- Data vs. Information
- Shallow/Multidimensional/Hidden/Deep Information
- The KDD Process
- Data Mining Strategies
- Information Representation and Interpretation

BI&KDD VLW Overview (2)

B) Self-practicing part

Exercises with model answers on:

- Affinity Analysis (Apriori algorithm)
- Decision Trees (J48/C4.5)
- Clustering (K-Means, DBSCAN, Agglomerative)

BI&KDD VLW Overview (3)

C) Hands-On Self-Study and Practicing

- Association rules mining (WEKA, IBM I.M.)
- K-MEANS clustering (WEKA)
- Classification (J48/C4.5 WEKA)
- Using the K-MEANS output to improve the quality of the J48 output (WEKA)
- Recommender system processing with association rules mining, improved with demographic clustering*

* IBM DB2 DWE IM Tutorial: *Managing your business in retail*

http://www.ibm.com/developerworks/edu/dm-dw-dm-retail_tutorial-i.html

BI&KDD VLW Overview (4)

D) Laboratory Report

Association Rules Mining:

- Software used: WEKA and IBM Intelligent Miner
- Datasets: (a) HEAL-Link e-resources/journals dataset
  (b) Titanic dataset

E) Quiz-Assessment

MCQ type

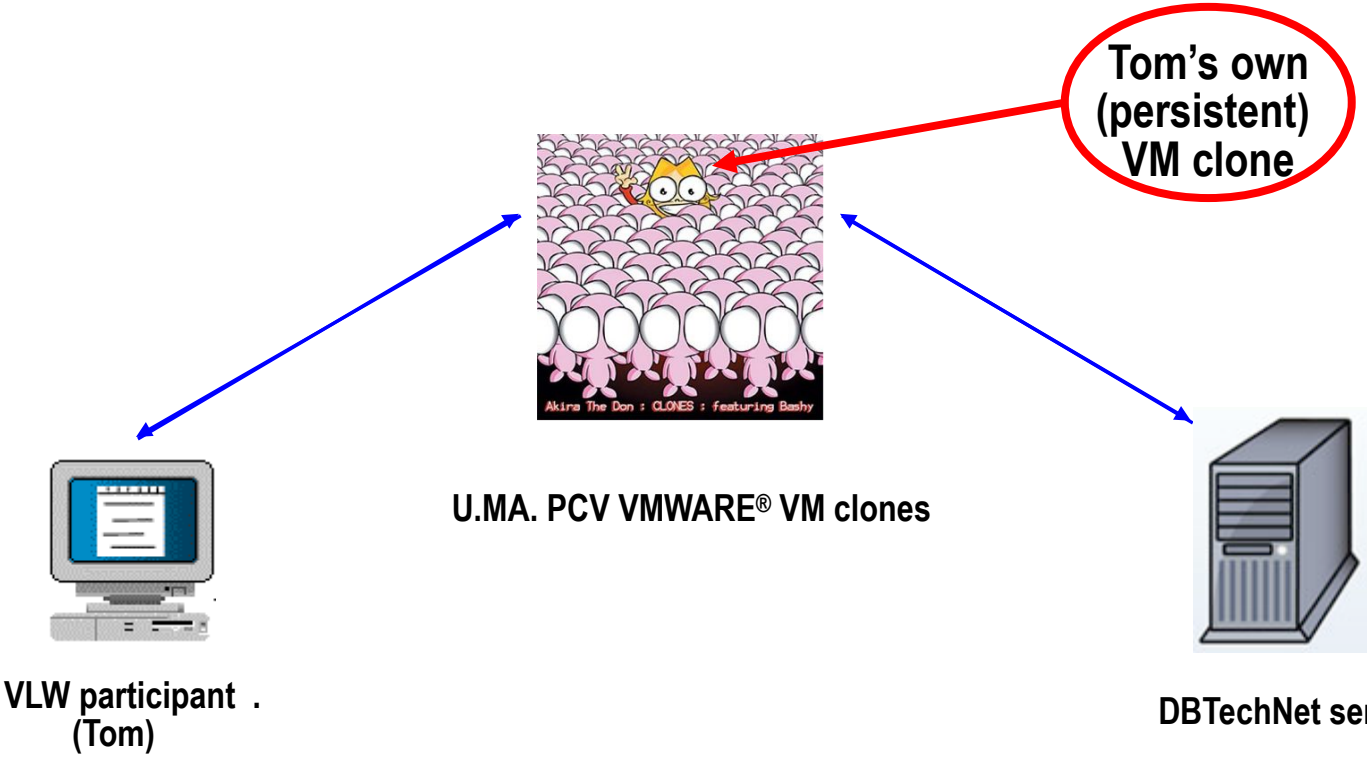
F) Supporting Material

'How To" Guides, Software installation and user guides (with videos), Trainer's/Trainee Guides, bibliography, etc.

Content Organization (http://dbtech.uom.gr)

- Three-level access authorization scheme:
  public access, portal user, BI&KDD user
- One-point of entry to all VLW content (BI&KDD "RoadMap")

The UMA PC Virtual Environment



...resolves the commercial software licensing issue

Trainer(s)-to-Trainee(s) Communication

A) Asynchronous

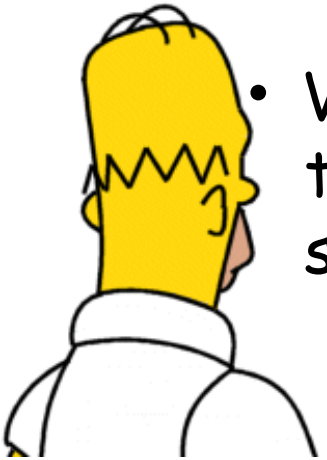
- email
- Discussion fora
- Google Docs forms

B) Synchronous

- 1-1: Skype (with screen sharing, if needed)
- 1-M: Adobe Connect Pro (UMA, HAAGA-HELIA)

Conclusion

- DBTech EXT BI&KDD VLW overview
  - Content outline and organization
- Self-study, self-practicing (model answers, step -by-step instructions for the 'hands-on' part
- Open source and commercial software utilized)



- Virtualization technology utilized, both to distribute the sw used, as well as to overcome the commercial software user licensing issue

Thank You
for
Your Attention!