

EMO 3. Akademik Kamp Etkinliği... BÜYÜK VERİ ve MADENCİLİĞİ

EMO tarafından İzmir Şirince’de Nesin Matematik Köyü’nde düzenlenen Akademik Kampın üçüncüsü kapsamında “Büyük Veri ve Madencilik” konusuna ilgili eğitimlere başlandı. Tüm dünyada giderek büyüyen veri tabanı ve bu verilerin değerlendirilmesi süreçlerinin ele alındığı kampta, Prof. Dr. Hamit Serbest’in eğitimi sorgulayan sunumu büyük ilgi gördü. Prof. Dr. Tayfun Akgül, “Bilim Etiği” sunumunda “kopyala-yapıştır” ve intihal sorunlarını masaya yatırdı.

Akademik Kamp, ilk gün Elektrik Mühendisleri Odası Müdürü Emre Metin ve EMO Yönetim Kurulu Saymanı İrfan Şenlik’in açılış konuşmasıyla başladı. Yrd. Doç. Dr. Şenlik, Akademik Kamp etkinliğini, EMO ve üniversiteler arasında işbirliğini geliştirmek, genç araştırmacılarla uzman akademisyenler arasında iletişimi sağlamak, bir araya getirmek amacıyla düzenlediklerini belirtti. İlk kampta “İşaret işleme” konusunun ele alındığını ve 50’nin üzerinde yüksek lisans öğrencisinin katıldığını; “Görüntü işleme” üzerine olan ikinci kampta da 60 civarında yüksek lisans ve doktora öğrencisinin katılımcı olarak yer aldığını kaydetti. EMO bünyesinde gerçekleştirilen bilimsel ve teknik sempozyumlar hakkında da bilgi veren Şenlik, EMO’nun 6 ayda bir yayımladığı Bilimsel Dergi’nin de 9. sayısı için hazırlıkların sürdürüldüğünü belirtirken, katılımcılara EMO Bilimsel Dergi’de makaleleriyle yer almaları için de çağrıda bulundu. “Tek amacımız, bilimsel dergiyi taranır bir dergi haline getirmek” diyen Şenlik, kampın verimli geçmesini temenni ederek sözlerini tamamladı.

Genç Akademisyenlere Öneriler

Akademik Kamp’ın yaratıcılarından EMO MİSEM Komisyonu Başkanı Orhan Örucü, “Burada, dünyada ve ülkemizde benzeri az bulunan bir etkinliği yapıyoruz. Bizim gençliğimizde izcilik vardı. Daha rahat, samimi ortamda bire bir ilişkiler yaratmak önemli. Normal kongreler, okul ilişkilerinde daha bürokratik ve soğuk oluyor. Biz onu ortadan kaldırarak böyle bir ortam yaratmaya çalıştık” sözleriyle konuşmasına başladı. 3. Akademik Kamp’a 37 doktora, 26 lisans öğrencisi ile 10 farklı üniversiteden 16 öğretim üyesinin katıldığını belirten Örucü, Kamp’ın gençler için önemli bir fırsat olduğuna dikkat çekerek, gençlere daha girişken olmaları konusunda tavsiyede bulundu. Gençlere Akademik Kamp gibi etkinliklerde yapılan anketleri önemsemeleri ve başvuru formlarında da çok dikkatli olmalarını öğütleyen Örucü, bu dokümanların kişilerin işe yaklaşımını ortaya koyduğunu ifade etti. Orhan Örucü, bundan sonra yapılacak olan kampın ana konusunu da “Elektrik Makinaları ve Güç Elektromanyetik” olarak açıkladı.

Türkiye’de her gün kötü haberle uyandığımızı, ama iyi şeyler de olduğunu belirten Örucü, Prof. Aziz Sançar’ın Nobel ödülü almasını kutlayarak, “8 çocuklu Mardin Savur’dan yoksul bir ailenin çocuğu, İstanbul Üniversitesi’nden mezun olduktan sonra ABD’ye gidiyor. 300’e yakın makale, ulaşılamayacak sayıda akademik bir atıf var arkasında” diye konuştu.



BÜYÜK VERİ VE MADENCİLİĞİ

Koordinatör: Prof.Dr. H. Altay GÜVENİR (Bilkent Üniversitesi)

KONUŞMACILAR ve KONULAR

| | | | | |
|--|---|---|--|---|
| <p>Prof. Dr. Namik Serbest (Çukurova Üniversitesi)</p> <p>Bilim, Piyasendık ve Öğretim Kurumları</p> | <p>Prof. Dr. Nuran Ferhatosmanođlu (Bilkent Üniversitesi)</p> <p>Veri Ambarları ve Veri İnkü Analizleri</p> | <p>Prof. Dr. Tayfun Akçıl (İstanbul Teknik Üniversitesi)</p> <p>Bilim Etik</p> | <p>Prof. Dr. Adilna Bır (İstanbul Teknik Üniversitesi)</p> <p>Bilim Tarihinden Bir Kesit</p> | <p>Prof. Dr. Ali Nesin (Nevşehir Vakfı)</p> |
| <p>Prof. Dr. Zehra Çataltepe (İstanbul Teknik Üniversitesi, tezi.iz)</p> <p>Sosyal Ağlarda Sınıflandırma ve Özetlilik Seçimi</p> | <p>Doç. Dr. Buğra Gedik (Bilkent Üniversitesi)</p> <p>Büyük Veri Süreçlerinde Özetlenebilir Çıkarımlar</p> | <p>Doç. Dr. Ali Taylan Cemgil (Boğaziçi Üniversitesi)</p> <p>Matris ve Tensor Ayrıştırma Teknikleri</p> | <p>Doç. Dr. Yücel Saygın (Sakarya Üniversitesi)</p> <p>Matris Madenciliđi: Temel Yöntemler ve Diğer Analizler</p> | <p>Doç. Dr. Osman Abul (TOBB ETÜ)</p> <p>Kanun Madenciliđi</p> |
| <p>Y. Doç. Dr. İ. Sengül Altıngövede (ODTÜ)</p> <p>Arama Motorları: Sınırlanmış İhtiyaçlar ve Bulunabilirlik</p> | <p>Mustafa İspir (Google)</p> <p>Sihirli Gece: Büyük Veri Uygulamaları</p> | <p>Y. Doç. Dr. İsmail Arı (Özyeğin Üniversitesi)</p> <p>Bulut Bilgi ve Büyük Veri Uygulamaları</p> | <p>Y. Doç. Dr. Erman Aydın (Bilkent Üniversitesi)</p> <p>Büyük Veri Sistemlerinde Veri Kaliteliđi: Genel Veriye Durumu</p> | <p>Y. Doç. Dr. Kamer Kaya (Sakarya Üniversitesi)</p> <p>Tıbbi Performans ve Büyük Veri Sistemleri</p> |
| <p>Y. Doç. Dr. Ahmet Bulut (İstanbul Şehir Üniversitesi)</p> <p>Özdeşlenebilir, İnteraktif ve Bulunabilir Büyük Veri Analizi</p> | <p>Başvuru Koşulları</p> <p>1- EMO Üyesi olmak 2- Büyük veri ve madencilik, makine öğrenmesi ve benzer alanlardaki bir konuda Yüksek Lisans veya Doktora öğrencisi olmak</p> <p>Son Başvuru Tarihi: 04 Eylül 2015</p> | | | |

EMO Akademik Kamp'a dair tüm bilgileri www.akademikkamp.org adresinden ulaşabilirsiniz.

Kamp Sekreteryesi: TMMOB Elektrik Mühendisleri Odası İhlamur Sokak No: 10 Kat:3 Kızılay/Ankara Tel: 0312 425 32 72 Dahili: 122 E-Posta: akademikkamp@emo.org.tr

Temel bilimlerin hak ettikleri yere gelmeleri temennisinde bulunan Örücü, konuşmasını şöyle sürdürdü:

“Bulduğumuz bu mekan da temel bilimlerin yeri olarak nitelendirilebilir. Taşından toprağından imece usulüyle yapılmış, onlarca binlerce insanın emeğı olan bir yer. Böyle bir ülkede böyle bir iş yapıyor. İzmir Üniversitesi’nde Biyotıp ve GENOM Merkezi var. Hala anlayamadığım şekilde bu işi bilen bir insan başına getirilmiş. Ben buna da olumlu yönde şaşırdığımı söyleyeyim. Liderler Vadisi diye bir grupla tanıştık. Yapmak istediğiniz projelere uluslararası düzeyde yardımcı olan hoş bir girişim.”

Veri Madenciliğinin Sosyal Boyutu Unutulmamalı

Orhan Örücü, 3. Akademik Kamp’ın ana konusu olan veri madenciliğinin dünyanın her yerinde yaygınlaştığını, ancak konunun bilimsel ve sosyal yanının göz ardı edilmemesi gerektiğini belirtirken, şu uyarıda bulundu:

“Herkesin davranışını alıp bir şeyler yaparsak, Orwell’in 1984’ündeki hale geliriz. Biz bu etkinliğimizi; kuru kuruya bilimsel konular dışında, bilim ve etik nedir, bilimsel tarih konularında biraz daha farklılaştırmaya çalışıyoruz. Hipokrat’ın lafı var: ‘Önce başkasına zarar verme.’ Bu tüm alanlarda uygulamamız gereken bir söz.”

Orhan Örücü, EMO’nun mühendislik camiasının değerlerine yönelik anma etkinlikleri gerçekleştirdiğini, ancak bu değerlerin yaşarken onurlandırılmalarının öneminden yola çıkarak ilki Prof. Mithat İdemen için olmak üzere etkinlikler düzenlenmesinin planlandığını açıkladı. 2016 yılında da, bilgi kuramının babası olarak bilinen, 2001 yılında yaşamını yitiren ABD’li Matematikçi, Elektronik Mühendisi Claude Elwood Shannon anısına 100. yaş gününde etkinlik düzenleneceğini kaydeden Örücü, katılımcıları bu etkinliklere de davet etti. Shannon’un “İletişimin Matematik Kuramı” çalışmasının da haberleşmenin önünü açtığını belirten Örücü, bu kitabın da çevirisinin EMO tarafından yaptırıldığı bilgisini katılımcılarla paylaştı.

Her Yerden Veri Akıyor

3. Akademik Kamp’ın Koordinatörlüğü’nü gerçekleştiren Prof. Dr. H. Altay Güvenir, “Büyük veri konusu neden gençlerin ve hocaların çok ilgisini çekiyor?” sorusuyla konuşmasına başlarken, şu yanıtı verdi:

“Teknolojinin ilerlemesi sayesinde, İnternet, sosyal medya devrimi sayesinde bilginin gücü öne çıkıyor. İş yapma şeklimizi kökten değiştirdi. Firmalar bir adım öne geçebilmek için fark yaratmak zorundalar. Bunun için de bu var olan teknolojiyi kullanabilirler. Yıllardır veri topluyorlar, bu verilerden nasıl bir değer yaratılabilir diye düşünüyorlar. Veri güdümlü yaklaşımla firmalar rakiplerine göre üstün duruma geçiyorlar.”

“Nerede bu büyük veri?” diyen Güvenir, konuşmasını şöyle sürdürdü:

“Her yerde var. Bilgisayar tuşlarına bastığımız zaman, İnternet’te arama yaparken, her yaptığımız iş bir veri olarak giriyor. Sensör her yerde yaygınlaşıyor. Bağlı buldukları ortamlarda sürekli veri üretiyorlar. Gelecekte daha da yaygınlaşacağı kesin olan, otomobillerde kullanılan teknolojilerde sürüşlerle ilgili veriler toplanıyor. Google’da sürücüsüz araçlara ilişkin araştırmalar yapılıyor. Çok enteresan yerlere gidiyoruz.”

“Bilgi ne zaman büyük oluyor?” sorusuyla konuşmasını sürdüren Güvenir, İnternet ortamında oluşturulan verilere ilişkin rakamsal bilgiler paylaştı. “2000 yılında 800 bin

petabyte gibi olan veri birikiminin 2020 yılında 35 Zettabyte'a çıkacağı tahmin ediliyor. Bu büyüklükteki veriyi işleyip, kabul edilebilir bir zamanda bilgi çıkarma işi; özel yeni teknolojiler gerektiriyor. Büyük veri, İngilizce 5 V harfiyle başlayan sıfatla tamamlanıyor. Birincisi variety (Çeşitlilik), ikincisi velocity (Hız). Hız çok yüksek. Sabit durmuyor veri sürekli geliyor. Volume (Veri büyüklüğü); 16 yılda 44 kat olacak veri miktarı. Verification (Doğrulama). Veri güvenliği de çok önemli bir konu. Beşincisi value (Değer). Akademik çalışma güzel, ama topluma insanlara bir faydası varsa o zaman bu işin değeri çok.”

SNS Research Report'tan bilgiler paylaşan Prof. Güvenir, 2014 yılında 30 milyar dolar olduğu belirtilen büyük veri pazarının her yıl yüzde 17 gibi büyüme oranı ile arttığını, 2020'de 76 milyar dolar civarına çıkmasının beklendiğini kaydetti. Pazar grafiğine bakıldığında da 3 önemli konunun ön plana çıktığını ifade eden Prof. Güvenir, ilkinin hizmet veren kurumlar ve danışmanlık olduğunu söyledi. “Diğer büyük kısım, donanım ile ilgili. Bellek ve işlemci altyapısıyla ilgili kısım veri büyüklüğüyle aynı oranda artacak 2020 yılına doğru” diyen Prof. Güvenir, üçüncü konuyu da Akademik Kamp sunumlarında da üzerinde çok durulacak bir konu olan etiketleme ve analiz etme platformları olarak aktardı.

Bu alandaki gelişime bakıldığında da her zaman olduğu gibi en büyük bütçenin savunma sanayide olduğunu, dikey pazarda finans, otomobil, uçak sektörü gibi alanların da bu alana büyük yatırım yapacağını düşünülüğünü kaydeden Prof. Güvenir, ardından medya ve eğlence, kamu hizmeti, perakende ve telekomünikasyon gibi alanların önemli gelişim perspektifi sunduğunu anlattı. Kıtalararasındaki beklenti dağılımına bakıldığında da bugün Asya-Pasifik ülkeleri ve Kuzey Amerika egemenliğinin 2020'ye doğru Asya'ya daha çok kayacağı tahminini de aktardı.

Sınıflandırmadan Kümelendirmeye Veri Madenciliği

Akademik Kamp'ın ilk gününde katılımcıların da oldukça ilgisini çeken “Veri Ambarları ve Birliktelik Analizleri” başlıklı sunumunda Prof. Dr. Hakan Ferhatosmanoğlu, Facebook'un “veri bilimcisi” iş ilanıya yola çıktı:

“Bu ilan çok fazla veride analiz yapabilecek, modelleme yapabilecek kişiler arıyor. İlanda öncelikle istenen “sınıflandırma” bilgisiyle, etiketleri olan objeleri kullanarak yeni bir obje için etiketleme yapan modelleri bilme ve kullanma yetisi belirtiliyor. Sınıflandırma terimi daha çok kategorik etiketler için, ilanda geçen regresyon da numerik değerler için kullanılıyor. Örneğin, müşterileri risk skorlarına göre notlandırırken regresyon kullanabiliyoruz. İlanda sözü geçen bir başka sözcük olan “kümeleme”de de yine gruplandırma söz konusu ancak bu sefer etiketsiz veri üzerinde. Zaman serilerini modelleme ve tahmin modelleri de listelenmiş. Ayrıca SQL ve veri tabanı bilgisi de isteniyor. Yani veri bilimi alanında istatistik var, makine öğrenimi var, veritabanları var. Veri bilimcisi analizden anlasın, programlama yapabilsin ve veri tabanı tecrübesi olsun isteniyor. Hala değerli verinin önemli bir kısmı yapısal olarak tutuluyor. O yüzden veri tabanı bilgisi yine birçok firma tarafından istenen bir bilgi.”

Veri alanında 20 senelik dalgalar halinde gelişmeler yaşandığını aktaran Prof. Ferhatosmanoğlu, 60-70'li yıllarda ilk sorulan noktanın “Biri veriyi nasıl modellerim, nasıl yönetirim” olduğunu belirtirken, bugün de benzer şeylerin tartışıldığına dikkat çekti. 1970'li yıllarda veri tabanlarının, 1990'lı yıllarda veri ambarlarının meşhur olduğunu ve buradan nasıl bilgiyi

analiz ederim noktasına gelindiğini anlatan Ferhatosmanoğlu, 2010'lu yıllarda artık yapısal olmayan çok kaynaktan gelen verilerle araştırmalar yapıldığını söyledi. Ferhatosmanoğlu, aslında konuların benzer olduğunu, ancak farklı modellemelerin gündeme geldiğine dikkat çektiği konuşmasında, 60'lı yıllarda denenen ağaç modelinin başarılı olmadığını, o dönemde yine bir kısım insanın veriyi çizge olarak tutalım dediklerini, onun da başarılı olmadığını, 30 yıl sonra ise sosyal ağlarda çizge yapıların popüler olduğunu anlattı. 1969 yılında da bir matematikçinin “Veriyi ilkökulda öğrendiğimiz gibi kümeler halinde tutalım” dediğini, bunun bugün de uygulanan bir yöntem haline geldiğini aktaran Ferhatosmanoğlu, verinin yapısal olmadığı noktada da bu kez yeniden veriye yapısallık kazandırma arayışının başladığına işaret etti. Satır ve sütun tabanlı yapılardan, satır tabanlı veri ambarlarında sütun tabanlı analizler yapan yapılara varıncaya kadar, büyük verilerin analizine ilişkin yapısal mimariler ve süreçler hakkında bilgi veren Prof. Ferhatosmanoğlu, veri tabanından veri ambarlarına gelinen aşamada, veri madenciliği yapıldıktan sonra çıkan sonuçların değerlendirilmesinin de ayrı bir süreci kapsadığına vurgu yaptı.

Veri tabanı ile veri ambarı arasındaki farklılıkları da konuşmasında ele alan Ferhatosmanoğlu, veri tabanının sürekli, veri ambarının zaman zaman güncellendiğini; veri tabanlarında normalize edilmiş (az özellikli) çok sayıda dar tablolar kullanıldığını; öte yandan veri ambarlarında çok sayıda özellik içeren geniş tablolar kullanıldığını; veri tabanında satırların, veri ambarında sütunların önemli olduğunu; veri tabanını binlerce kişi, veri ambarını ise analizcilerin kullandığını anlattı. Veri ambarlarının oluşturulmasında, entegrasyon, yanlışlıkları düzeltme, analizden önce transformasyonları yapma konularına da değinen ve örnekler üzerinden öğrencilerin de katılımının sağlandığı Ferhatosmanoğlu'nun sunumunda, veri ambarlarındaki analiz süreçlerine ilişkin ana tablonun yanında oluşturulan ayrıntılı boyut tabloları, küp şeklindeki veri ambarı oluşturulması, yıldız ve kar tanesi şemaları anlatıldı. Bu süreçlerde karşılaşılan sorunlara da değinen Prof. Ferhatosmanoğlu, veriyi elde etme ve verinin güvenliği konusundaki sorunları da ele alınırken, yanlış veri girişinden, verinin boş bırakılmasına, verinin karşılıklı etkileşimine varıncaya kadar veri analizine ilişkin konunun çok boyutlu değerlendirmesini yaptı.

Eğitime Sorgulayıcı Bakış

Prof. Dr. Hamit Serbest ise “Eğitim İyi Bir Şey mi?” başlıklı sunumunda, hangi inanç ve değerler sistemiyle eğitimin verildiğini sorgulayan, katılımcıların da sorularıyla yoğun ilgi gösterdikleri karşılıklı etkileşimli bir sunum yaptı. “Eğitim şart mı?”, “Hangi eğitim olmalı?” gibi şaşırtıcı sorular ve sorgulamaların yer aldığı sunumunda Hamit Serbest şöyle konuştu:

“Yanlış yapan eğitimli olabilir. O kişi büyük bir eğitim almış olabilir, ama onun aldığı doğrular kendi doğruları olabilir, başkanlarının doğruları olmayabilir? Kim kime hangi doğruları öğretecektir? Değer yargıları da zamanla değişiyor. Eğitim dediğimizde birilerinin doğru olarak belirlediği şeylerin diğer bireylere aktarılması mı, eğer bu doğruysa insanlar başkalarının doğrularını esas alarak yaşayacaklardır. Başkasının doğruları eğer hayatın esasını oluşturacaksa, bu kafatasının içindeki et parçası neye yarayacaktır?”

“Dindar ve muhafazakar bir gençlik yetiştirmek istiyoruz” sözü üzerinden eğitim konusunu tartışmaya devam eden Prof. Serbest, eğitimin tanımını şöyle yaptı:

“Eğitim, bireylere hayatta gerekli bilgi ve kabiliyetlerin sistematik bir şekilde verilmesidir. Okullar, üniversiteler, kurslar, uzaktan eğitim gibi örgün ve yaygın eğitimler vardır. Bu aslında bireyin doğumdan ölümüne kadar süren bir olgu. Bunun politik, sosyal, kültürel ve bireysel boyutları var.”

Prof. Hamit Serbest, farklı eğitim tanımlamalarıyla sorgulamasını sürdürürken, “Bireyin yaşantısı yoluyla davranışlarında kasıtlı olarak istenen değişimi meydana getiren bir süreç” tanımı üzerinde durdu. Eğitimle ilgili çeşitli alıntılar ve örnekler yoluyla irdeleme yapan Prof. Serbest’in Prof. Dr. Sami Zan’a atfen İstanbul Üniversitesi Tıp Fakültesi Yıllığı’ndan alıntılacağı şu sözler dikkat çekti:

“Meyvesi çamura düşüyor diye ağaca mı lanet edilir?
Bence en acınacak insan görevinde ücretten başka bir şey alamayandır.
Üniversiteye girip de çıkamayanlara profesör denir.
Hayatta bütün setler üzerinden geçilmek için yapılmıştır, önünde durmak için değil.
Yükselmek için kendi ayaklarınızı kullanınız. Başkalarının sırtını ve ellerini değil.
Söndüremeyeceğin ateşi yakma!
Yolun ilerisini göremiyorsanız, dönemece gelmişsiniz demektir.
Kader size bir limon veriyse ondan limonata yapacaksınız.”

Sunumda yer alan eğitim sistemindeki ödül, ceza ve yasaklamalar konuları da katılımcıların yoğun tartışmalarına yol açtı. Bunun üzerine Hamit Serbest, konuşmasında şu vurguya yer verdi:

“İnsan hayatının toplumsal yaşantının neyi referans alacağına kimse karışmamalı. İnsanın doğasını engelleyen birtakım dayatmaların kimseye faydası yok. Dünya görüşünüzün ne olduğu, dini inançlarınızın olup olmadığını, ne olduğunun hiç önemi yok, eğer insansanız sorgulamanız gerekiyor.”

Hamit Serbest, dünyada gelişimin nasıl sağlandığına ilişkin iki yolu şöyle anlattı:

“1. Yol: Üstün yetenekli birisi oturup bir şey icat eder. Bize teknoloji olarak döner. Bir gereksinim yaratır. Talep yaratmayı gerektirir. 2. Yol: Gereksinimi görür birisi yaratıcı çabaları tetikler. Gereksinimleri ve bunların sınırlarını tahmin etmek zordur. Her iki yolun da hangisinden giderseniz gidin ister icat eden, ister gereksinim karşılayan tarafında olun her iki ucunda da insan var. Teknolojik ürünler insan aklının eseridir diyoruz. İnsanın dünyası kurmaca mı? Doğaya dayalı, ama biz kendimize göre bir dünya kurmuşuz. Akıl, merak, hayal!”

Mühendisliğin; bilim, teknoloji ve ekonominin ara kesiti olduğunu ifade eden Hamit Serbest, mühendisliğin modellemeler yaparak çözüm ürettiğini, bu noktada bilime başvurduğunu ve bilimin de dogmaları içermediğinin altını çizdi. Bilimin de tutucu olsa bile temel bir kural olarak her konulan kuralın tartışmaya açık bırakıldığına dikkat çeken Hamit Serbest, bilgi kavramının İngilizce karşılığıyla ilgili tartışmaya ilişkin olarak da “Bilgi dediğimiz zaman bizim kastettiğimiz knowledge olması lazım. Üniversitelerin bilgi üreten yerler olması lazım” diye konuştu.

Bilginin üretiminde de artık bunun nasıl yapılacağına ilişkin ihtiyaçların net olarak ortaya konulup, planlamasının yapılmasının önemini anlatan Hamit Serbest, Cumhuriyet dönemindeki eğitim alanındaki gelişmeleri özetlerken, yaratılan gelişim ekosisteminin kesintiye uğradığı tarihleri şöyle aktardı:

“İhtira Beratı Kanunu 1879’da çıkmış, sonra unutmuz.1952’de yeniden hatırlamışız. 1994’de TPE kurulmuş, 2004 yılında yeniden hatırlamışız, şimdi yine unutmuz. Hepsinin Türkiye’yi bir noktada Batı için pazar yapacak tarihler olduğunu görüyoruz. Tanzimat, Marshall yardımları, Gümrük Birliği, AB uyum yasalarının başlangıcı.”

İlk Türk uçağının 1936’da yapılmış olması, Devrim Otomobili ve Teletaş örnekleri üzerinden Türkiye’nin gelişimi önündeki engelleri sorgulayan Hamit Serbest, çıkış yolunun sorgulama olduğunun altını çizdi.

“İntihal Cehalettir”

Prof. Dr. Tayfun Akgül ise “Bilim Etiği” sunumunda, bilim hırsızlarına ilişkin Türkiye’den ve dünyadan ilgi çekici örnekleri irdeledi. Gençlere bilimsel hırsızlıkla yayın yapmalarındansa hiç yayınlarının olmamasının daha iyi olacağını söyleyen Akgül, özellikle gençlerin bu konuda dikkatli olmaları gerektiğini, bugün yapılan bir hatanın mutlaka ortaya çıkacağı ve geleceğini karartacağını da aktardı. Dünyada ve Türkiye’de bilimsel etik kurallarına da değinen Akgül, İstanbul Teknik Üniversitesi’nde bu konuda verilen derslere, içeriklerine, kullanılan yayınlara değinirken, EMO’nun etik konulu kitaplarının da referans olarak öğrencilere okutulduğunu kaydetti. Akgül, EMO’nun ve üniversitenin bilimsel etik konusunda yaptığı etkinliklere de değindiği konuşmasında, IEEE’nin EMO’nun uluslararası niteliğinde olan bir örgüt olduğunu belirterek, IEEE’deki bilim ve etik ile intihal konusundaki duyarlılığı ve çalışmalarını aktardı. Yayın ilkeleri konusunda da bilgi aktaran Akgül, şöyle konuştu:

“Sayılarla oynamayın, yayınızı şişirmeyin. Çünkü intihal cehalettir. Bir kere yaparsan yanlışlıkla lekelendin bittin yani. Onun için çok dikkatli olmak gerekiyor. Hele işin başındaki sizin gibi arkadaşlar bu konuda olabildiğince dikkatli olun. Hocanızı bile dikkate davet edin. Bazı öğrencilerin başına geldi. Hocasının hatasından dolayı çocukların da kara listeye girdiği durumlar oluyor.”

Arama Motorları Aklımızı mı Okuyor?

Yrd. Doç. Dr. İ. Sengör Altıngöve ise “Arama Motorları: Samanlıkta İğne Aramak ve Bulmak” sunumuyla veri madenciliğini arama motorları bazında irdeledi. Arama motorlarının temel bileşenlerinin anlatıldığı, ağ tarama, indeksleme ve sorgu işleme alanındaki temel yöntemlerin anlatıldığı sunumda, ağ motorları üzerinde aramaların optimizasyonu tartışıldı. İndeks budama ve önbellekleme gibi ileri yöntemlerde arama motorlarına özgü problemler ve çözüm yöntemleri üzerinde duran Altıngöve, “Arama motoru sizin aklınızdan geçeni tabii ki bilmiyor, önceki sorgu loguna bakarak, artı pek çok ilave ipucunu kullanarak size sorgu önerisi yapabiliyor. Hem arama motorunun esas işlevini hem de pek çok analiz işlemini yapabileceğiniz bir veri kümesi oluşturuyor” dedi. Arama motorlarının sorgu işleminde ülkelerin kültürel farklılıklarından kaynaklı farklı algılamalardan başlayıp, kişilerin eğitim durumları, yaş gruplarına varıncaya kadar farklı analizler yaparak seçenekler sunmanın zorluğuna dikkat çekti. Bunları yaparken de çok önemli bir etken olarak “hız” sorununa işaret eden Altıngöve, “Tüm bunların hepsini anlamasını ve başarılı olmasını sağlayacaksınız ve kullanıcılar çok sabırsız” diye konuştu.

Google’dan İspir Sibyl’i Anlattı

Kampın ilk gün çalışmaları akşam yemeğinden sonra Google’dan Mustafa İspir’in “Sibyl: Google’da Büyük Veri Uygulamaları” sunumuyla tamamlandı. Google’daki yapay öğrenme

araçlarından biri olan Sibyl'in özelliklerini makine öğrenimiyle ilgili gelişmeleri ve sorunların çözüm yollarına ilişkin deneyimleri paylaşan İspir, büyük ölçekli verinin tanımını yaparken 100 milyardan daha büyük verinin öğrenme örneği ve yine 100 milyarlarca tekil özellikler, in bulunduğunu kaydetti. "Her sütun bir özellik, her satır bir öğrenme örneği. Dolayısıyla matrisin büyüklüğü 100 milyara 100 milyar. Küçük bir matris değil" dedi. Makine öğrenimiyle ilgili hangi yönlerin kullanılıyor olabileceği, Youtube örneği üzerinden kullanıcılara video önerisinin nelere göre belirlendiği gibi sorularıyla etkileşimli bir sunum yapan Mustafa İspir, Google'ın yeni bir ürünü olan kullanıcı ara yüzü ve makine öğrenimine dayalı "Google Photos" uygulamasını makine öğrenimine çok iyi bir örnek olarak gösterdi.

Youtube'da video önerilerinin nasıl yapıldığına ilişkin katılımcıların da görüşlerini aldıktan sonra İspir, konuşmasını şöyle sürdürdü:

"Diyelim ki biz tıklamaya göre seçeceğiz. Teoride bu söylenenlerin hepsi bir sinyal. En son bu sinyallerin hepsini toplayıp bir sayı çıkarıyor ve o sayıya göre sıralama yapıyorsunuz. Sinyallerin en büyüğü şu anda izlediğimiz video ve bu verileri kullanarak öyle bir sıralama yapacağım ki bunların tıklanma sayısı en yüksek olacak. Böyle bir sistem kurduğumuzu varsayalım. Makine öğreniminde en önemli faktör, veri ve algoritmanın etkinliği. Ben bir video sunabilirim. Bu gösterdiğim veri tıklanabilir. Ama ben gösterirsem tıklanabilir. Gerçekten böyle bir döngü var. Eğer seçeneklerde ben ilgili olmayan tıklanmayacak bir video gösterirsem tıklanmaz. Benim algoritmam bana bunu öğretir. Kötü videolar bu döngüyle elenir. Ama neyi kaybedebiliriz? Gerçekten ilgili bir video vardır. Onu kaybedebiliriz. Rastgele bir tanesini buraya eklemem gerçekten çok büyük bir maliyet. Bir değil, iki yer ayıralım desek bunun normal Youtube izlenme zamanına direk etkisi var."

Öncelikle sistemde her şeyin elle başladığını, eldeki bir ürünü geliştirmek ya da yeni bir ürün için elle yapılan girişin ardından makine öğrenime geçilince en büyük etkinin orada ortaya çıktığını anlatan İspir, "Burada önemli olan şey veri. Düzgün hale getirilmiş bir veriye basit bir makine öğrenimi ile birçok katkıda bulunuyoruz."

İspir, makine öğreniminde "linear ve logistic regression" temel olarak kullanıldığını kaydetti. Eldeki büyük veriyi kullanırken bölümlenmeler yapıldığını bunun da hem elle hem de makine öğrenimi aracılığıyla yapıldığını belirten İspir, büyük verinin nasıl işleneceği konusunda yapılan hesaplamaları da örneklendirdi. Büyük veriyi okumak ve yazmanın önemli diğer bir zorluk olduğunu ifade eden İspir, makinenin yok olması durumuna karşı da veriyi korumak için olabildiğince sürecin ortalarında kaydetme çözümüne başvurduklarını anlattı. Eldeki veri ile makine öğrenimi arasında veri adlandırma gibi alanlarda yaşanan uyumsuzluk sorunun en büyük problem olduğunu, buna yönelik çözümlerin de karlılık oranına doğrudan etkide bulunduğunu söyledi. Yapılan uygulamalarda başarı ölçümü olarak da A/B testinin kullanıldığını belirten İspir, online ve offline yapılan uygulamalar arasında ciddi farklılıklar bulunduğunu bildirdi. Mustafa İspir, Google Araştırma'nın en önemli işinin aslında danışmanlık olduğunu; ürünü geliştirmenin yetmediğini, o ürünü kullanacak kişinin yanında olup, problemleri görerek çözümlerin sağlandığını; Sibyl'in de bu nedenle çok başarılı ve yaygın olduğunu sözlerine ekledi.

EMO 3. AKADEMİK KAMP 8-11 EKİM 2015 DEĞERLENDİRME VE SONUÇ RAPORU

Elektrik Mühendisleri Odası'nın ilkinin 2012 yılında gerçekleştirdiği Akademik Kampın üçüncüsü 8-11 Ekim 2015 tarihlerinde "Büyük Veri ve Madencilik" başlığı altında gerçekleştirildi. Nesin Vakfı'nın İzmir Şirince'deki Matematik Köyü'nde yapılan kampta görüntü işleme alanında çalışan EMO ve diğer Oda üyesi yüksek lisans ve doktora öğrencileri ile alanda çalışan uzman akademisyenler bir araya geldi.

Belli bir konu altında bir araya gelen araştırmacılara; işyeri, eğitim, seminer ve kongrelerden farklı olarak daha yoğun ve doğrudan etkileşim sağlanabilecek bilimsel bir tartışma ortamı sağlamak amacıyla düzenlenen EMO 3. Akademik Kampla, Odamızın bilim dünyasıyla olan ilişkilerinin güçlendirilmesi hedeflenmiştir.

Büyük Veri ve Madencilik alanında çalışan, 35 farklı üniversiteden 64 doktora veya yüksek lisans öğrencisinin katıldığı kampta, alanında uzmanlıkları ile tanınan 14 farklı üniversiteden 17 akademisyen ders verdi. Katılımcıların 35'i doktora öğrencisi, 26'sı yüksek lisans öğrencisi ve 3'ü misafirlerden oluştu. Ayrıca katılımcıların 20'sini kadınlar, 40'ını erkek öğrenciler oluşturdu. Kampa EMO Yönetim Kurulu Başkanı Hüseyin Yeşil, Yönetim Kurulu Saymanı İrfan Şenlik ve MİSEM Komisyonu Başkanı Orhan Örucü katıldı.

Birinci Gün

Açılış: EMO 3. Akademik Kamp 8 Ekim 2015 Perşembe günü 09.00'da açılış töreni ile başladı. Açılıшта dört konuşmacı yer aldı:

- Oda Müdürü Emre Metin kamp çalışmalarına ilişkin temel bilgileri katılımcılara aktardı.
- EMO Yönetim Kurulu Saymanı İrfan Şenlik katılımcıları selamlayarak kampın amacı ve EMO'nun akademi dünyası ile ilişkileri hakkında bilgi verdi.



- EMO MİSEM Komisyonu Başkanı Orhan Örucü daha önce düzenlenen akademik kampa ve bundan sonra yapılması planlanan kamplara ilişkin bilgi sundu.
 - EMO 3. Akademik Kampın akademik içeriğinin organizatörü Prof. Dr. H. Altay Güvenir, kampın öğrencilere yapacağı katkılar ve kampın içeriği hakkında genel bilgi verdi.
- Açılış konuşmalarının ardından katılımcılar kendilerini tanıtarak yaptıkları çalışmalara ilişkin bilgi aktardılar. Sonrasında birinci gün derslerine geçildi, ilk gün beş ders verildi:
- Prof. Dr. Hamit Serbest, “Eğitim İyi Bir Şey Mi?”
 - Prof. Dr. Hakan Ferharosmanoğlu, “Veri Ambarları ve Birliktelik Analizleri”
 - Prof. Dr. Tayfun Akgül, “Bilim Etiği”
 - Yrd. Doç. Dr. İ. Sengör Altıngöve, “Arama Motorları: Samanlıkta İğne Aramak, ve Bulmak!”
 - Mustafa İspir “Sibyl: Google’da Büyük Veri Uygulamaları”

İkinci Gün

9 Ekim Cuma günü altı ders verilmiştir:

- Prof. Dr. H. Altay Güvenir, “EMO Bilimsel Dergi”
- Prof. Dr. Atilla Bir, “Öklid’ten Nasreddin Tusi’ye, Tusi’den Uluğ Bey’e Bilim”
- Yrd. Doç. Dr. İsmail Arı, “Bulut Bilişim ve Büyük Veri Uygulamaları”
- Doç. Dr. Buğra Gedik, “Büyük Veri Dünyasında Ölçeklenebilir Çizge İşleme”
- Doç. Dr. Ali Taylan Cemgil, “Matris ve Tensor Ayrıştırma Teknikleri”
- Prof. Dr. Ali Nesin

Üçüncü Gün

10 Ekim Cumartesi günü dört ders verilmiştir:

- Doç. Dr. Yücel Saygın, “Metin Madenciliği: Temel Yöntemler ve Duygu Analizi”
- Prof. Dr. Zehra Çataltepe, “Sosyal Ağlarda Sınıflandırma ve Öznitelik Seçimi”
- Doç. Dr. Osman Abul, “Konum Mahremiyeti”
- Yrd. Doç. Dr. Erman Ayday, “Büyük Veri Sistemlerinde Veri Gizliliği: Genetik Verinin Durumu”

Dördüncü Gün

11 Ekim Pazar günü 2 ders verilmiştir:

- Yrd. Doç. Dr. Kamer Kaya, “Yüksek Performanslı Büyük Veri Sistemleri”
- Yrd. Doç. Dr. Ahmet Bulut, “Ölçeklenebilir, İnteraktif, ve Hataya Dayanıklı Büyük Veri Analizi”

Dördüncü gün derslerin ardından Efes Harabeleri ve Şirince Köyü’ne yapılan gezi ile kamp programı tamamlandı. Kamp sırasında katılımcılara anket formu dağıtılarak kampı değerlendirmeleri istendi.

Sayılarla Akademik Kamp

- 26 yüksek lisans, 35 doktora öğrencisi, 3 misafir öğrenci olmak üzere toplam 64 katılımcı
- Katılımcıların 20’si kadın 44’ü erkek
- 35 ayrı üniversite/kurumdan katılım
- 16 öğretim üyesi, 1 firma çalışanı (Google) ders verdi.
- Kamp esnasında 17 ayrı konuda toplam 31 saat ders yapıldı.
- Dersler saat 09.00’da başlayıp, 22.00’de sona erdi.