

# TÜBİTAK-ARDEB

## Veri Madenciliği ve Veri Depolama Çağrı Programı

### “1003-BIT-VERI-2016-1 Yenilikçi Bilgi Keşif Yöntemleri”

#### Çağrı Metni

## 1. Genel Çerçeve

Büyük veri, geleneksel veri saklama ve işleme teknolojileri ile saklanması ve işlenmesi mümkün olmayan ölçüde büyük, çeşitli ve karmaşık veri olarak tanımlanabilir. Büyük verinin temel özellikleri hacim, çeşitlilik, hız, değişkenlik, güvenilirlik (veracity), izlenebilirlik (provenance) ve karmaşıklık olarak sıralanabilir. Büyük veri ile ilgili temel sorunlar büyük verinin edinilmesi, filtrelenmesi, serileştirilmesi, sıkıştırılması, yönetimi, saklanması, modellenmesi, üzerinde arama yapılması, paylaşılması, aktarılması, çözülmesi ve görüntülenmesini içermektedir. Bütün bunlarla ilgili sistemlerin (algoritma, donanım ve/veya entegrasyonları) gerçek zamanlı ve koştur sistemlerde tasarlanıp uygulanması gün geçtikçe daha fazla önem arz etmektedir.

Sürekli gelişen ve değişen teknoloji ile birlikte büyük verinin kapsamı ve ortaya çıkardığı sorunlar da değişmektedir. Büyük veriden faydalı bilgilerin çıkartılması için gelişmiş çözümleme yöntemlerinin kullanılmasına ihtiyaç vardır. Bu da mevcut teknolojilerin büyük ölçekli ve karmaşık veri kümelerinde kullanımı için geliştirilmesi ve büyük verinin özelliklerine uygun yenilikçi çözümleme yöntemlerinin geliştirilmesini gerektirmektedir. Çağrı kapsamında Büyük Veri alanına yönelik Ar-Ge niteliği olan yenilikçi araştırmalar ve uygulamalar hedeflenmektedir. Araştırma alanları mühendislik, istatistik ve matematik gibi disiplinlerdeki gelişmeleri hedeflerken, uygulama alanları da birçok disiplini (sosyal bilimler, mühendislik, sağlık, fen bilimleri, finans vb.) hedeflemektedir.

## 2. Amaç ve Hedefler

Veri tabanlarında Bilgi Keşif Yöntemlerine Yönelik Projeler aşağıdaki ve benzeri konulara odaklanacaktır:

### \* Veri yapısını inceleme yöntemleri

- ✓ Önişlemede serileştirme ve sıkıştırma algoritmaları,
- ✓ Büyük Veri ayıklama algoritmaları,
- ✓ Akan veri önişleme başarımı – benchmark oluşturma,
- ✓ Akan veri önişlemede ölçelenebilirlik,
- ✓ Yazılım tanımlı veri saklama yöntemleri

### \* Yenilikçi, seçme, ayıklama, önişleme yöntemleri

- ✓ Önişleme araçları (Flink, Storm Spark gibi) geliştirme,
- ✓ Nesnelerin interneti (IoT, M2M) verilerinin elde edilmesi,
- ✓ Yüksek başarımli hesaplama ve/veya gerçek zamanlı sistemler için yenilikçi yaklaşımlar,
- ✓ Akıllı veri toplama yöntemleri

### \* Yenilikçi veri madenciliği yöntem ve algoritmaları

- ✓ Veri madenciliğinde veri güvenliği ve kişisel verilerin korunmasına yönelik uygulamalar,
- ✓ Yeni veri madenciliği modellerinin ve uygulamaların geliştirilmesi

\*Yenilikçi veri dönüşümü yöntemleri

\* Özel kaynaklardan veri madenciliği

- ✓ Tıp ve eczacılıkta veri madenciliği,
- ✓ Biyoteknolojide veri madenciliği,
- ✓ Meteorolojide veri madenciliği,
- ✓ Kablosuz sistemler için veri madenciliği,
- ✓ İnternet ve sosyal ağlardan veri madenciliği,
- ✓ Tarımda veri madenciliği,
- ✓ Modelleme, görselleştirme, kişiselleştirme ve öneri sistemleri için veri madenciliği,
- ✓ Nesnelerin internetinde (IoT), M2M veri madenciliği

\* Otomatik geri getirme (retrieval) ve veri madenciliği

- ✓ Otomatik bilgi çıkarımı,
- ✓ Çoklu ortam arşivlerinden otomatik geri getirme,
- ✓ Otomatik telif hakkı ihlali tespiti,
- ✓ Otomatik konsept tespiti, sınıflandırma ve özet çıkarımı

### 3. İlgili Destek Programı

Bu çağrı konusu kapsamında önerilecek projelere "1003-Öncelikli Alanlar Ar-Ge Projeleri Destekleme Programı" kapsamında destek verilecektir.

### 4. Çağrıya Özel Hususlar

- Önerilecek projeler küçük, orta veya büyük ölçekli projeler olarak hazırlanabilir.
- Bu çağrı kapsamında altyapı oluşturmaya yönelik olan projeler desteklenmez ve proje bütçe kalemleri arasında dengeli bir dağılım olması beklenir.
- Farklı disiplinlerden araştırmacıların proje ekibinde görev alması ve konunun disiplinlerarası bir yaklaşımla ele alınması önerilmektedir.
- Orta ve büyük ölçekli projelerin farklı kurum/kuruluşlarda yürütülen ve birden fazla kurumun yer aldığı alt projelerden oluşması (bir proje en fazla 1 ana ve 3 alt projeden oluşabilir) ve üniversite ile özel sektörün katılımı önerilmektedir.
- Orta ve büyük ölçekli projelerin büyük veri üretme ve tüketmeye yönelik uygulama içermesi zorunludur.
- Bu çağrı programına önerilecek projelere, yeni kurulan üniversitelerden (2006 yılından itibaren kurulmuş üniversiteler) proje yürütücüsü ve/veya araştırmacıların katılımının sağlanması teşvik edilmektedir (\*).

(\*) Bilimsel değerlendirme sırasında aynı/yaklaşık puan alan proje önerilerinden belirtilen koşulu sağlayanlara bütçe imkânları da gözetilerek öncelik sağlanacaktır.

### 5. Çağrı Takvimi

Aşama	Çevrimiçi Başvuru Sistemi Kapanış Tarihi	Elektronik Başvuru Çıktısının Gönderilmesi İçin Son Tarih*
Birinci Aşama	01.04.2016 17:30	08.04.2016 17:30
İkinci Aşama	17.06.2016 17:30	01.07.2016 17:30

\* Elektronik başvuru çıktısının ıslak imzalı nüshasının belirtilen tarih ve saate kadar Kurumumuza ulaştırılması gerekmektedir.

## 6. Ek Belgelere Referanslar

- 1003 Destek Programı Web Sayfası
- 1003 Destek Programı Bilgi Notu
- 1003 Öncelikli Alanlar Ar-Ge Projeleri Destekleme Programı Usul ve Esasları
- Ulusal Bilim, Teknoloji ve Yenilik Stratejisi (UBTYS) 2011-2016
- 1003 Proje Önerisi Değerlendirme Formu
- Yasal/Özel İzin Belgesi Bilgi Notu
- Etik Kurul Onay Belgesi Bilgi Notu

## 7. İrtibat Bilgileri

### Altuğ Çil

<b>Tel</b>	0312 4685300-1227
<b>e-posta</b>	altug.cil@tubitak.gov.tr

### Betül Öztürk

<b>Tel</b>	0312 4685300-1226
<b>e-posta</b>	ozturk.betul@tubitak.gov.tr