

A principal motivação do projeto é a geo-espacialização de notícias agrícolas com base no pressuposto de que essas notícias possam indicar fatores de influência sobre modelos de previsão de safras. Para geo-espacializar e inferir quais são os fatores e qual sua influência na linha temporal, no projeto estão em desenvolvimento técnicas inovadoras em mineração de textos, que permitam identificar e extrair automaticamente das notícias agrícolas três categorias: **tópicos** (assuntos: palavras-chaves ou categorias agrícolas), **cobertura espacial** e **cobertura temporal**. A cobertura temporal visa, em primeiro plano, permitir um estudo da evolução daquele tópico ou assunto em relação aos demais, em intervalos de tempo; ou seja, verificar a tendência de dispersão desse tópico ao longo do tempo e futuramente como ela se relaciona aos modelos de previsão de safra agrícola. A cobertura espacial visa compreender qual é a cobertura geográfica daquela notícia, podendo ser de caráter geral (nacional) ou relativa a alguma micro ou macro-região geográfica. Como essas notícias são altamente dinâmicas, pois são, diárias ou, no mínimo, semanais, devem ser tratadas em tempo real; logo, os tópicos também devem ter uma evolução bastante dinâmica, que deve ser considerada em seu tratamento, embora também o projeto trabalhe com a identificação estática de tópicos. As técnicas em desenvolvimento são aplicáveis a quaisquer publicações textuais em mídia eletrônica.

## Objetivo

Esse *workshop* tem por objetivo relatar o estado atual de desenvolvimento, bem como identificar pontos que ainda devem ser atacados e estabelecer estratégias para atingir o alcance de suas metas até o final de julho. A presença de alguns convidados, juntamente com os membros do projeto e parceiros, nos permitirão discutir o futuro das aplicações dos produtos, bem como novas linhas de desenvolvimento do projeto. O encontro deve culminar com o relatório de acompanhamento e o planejamento de ações futuras.

## Programação

- 8h30 - 9h **Abertura** - Chefia Embrapa Informática Agropecuária
- 9h - 9h50 **Palestra 1** - O projeto TIENA - objetivos, inovações, intersecções e integrações de técnicas (Maria Fernanda Moura)
- 10h - 10h50 **Palestra 2** (Plano de Ação 2) - Extração de candidatas a termo e cobertura temporal - técnicas investigadas, ferramentas desenvolvidas (Leandro Henrique Mendonça de Oliveira)
- 11h - 11h50 **Palestra 3** (Plano de Ação 3) - Identificação de tópicos - técnicas investigadas e ferramentas desenvolvidas (Maria Fernanda Moura)
- 12h - 13h50 Almoço
- 14h - 14h50 **Palestra 4** (Plano de Ação 4) - Identificação da cobertura espacial de notícias - técnicas investigadas e ferramentas desenvolvidas (Eduardo Antonio Speranza)
- 15h - 15h50 **Palestra 5** (Plano de Ação 5) - Integração de dados e software - banco de dados e ferramental de integração das ferramentas (Fabiano Fernandes dos Santos)
- 16h - 18h **Mesa redonda** - Estratégias para completar o atingimento das metas propostas (Mediador: Solange Oliveira Rezende)

## 3 de maio

## Programação

- 9h-12h **Mesa redonda** - Consolidação das estratégias - elaboração do plano de ação (Mediador: Maria Fernanda Moura)
- 12h-13h50 Almoço
- 14h-14h50 **Palestra** (convidados) - O projeto TIENA - objetivos, inovações, aplicações (Maria Fernanda Moura)
- 15h-16h30 **Mesa redonda** (convidados) - Possibilidades futuras da tecnologia gerada (Mediador: Eduardo Antonio Speranza)
- 16h40-17h **Fechamento** - Chefia da Embrapa Informática Agropecuária

## Local

Sala de treinamento 3 da Embrapa Informática Agropecuária

## 4 de maio



TIENA  
projeto

Workshop



## Palestrantes

**Maria Fernanda Moura**, líder do projeto, doutora em Ciências, área de Computação e Matemática Computacional, pela Universidade de São Paulo (2009), mestre em Engenharia Elétrica pela Universidade Estadual de Campinas (1992) e bacharel em Estatística pela Universidade Estadual de Campinas (1987). Tem experiência na área de probabilidade e estatística, com ênfase em planejamento de experimentos e análise de dados, em engenharia de software, com ênfase em desenvolvimento e em mineração de textos para a gestão da informação e de competências, bem como em várias técnicas de aprendizado de máquina estatístico. É pesquisadora da Embrapa Informática Agropecuária desde 1989.

**Leandro Henrique Mendonça de Oliveira**, responsável por plano de ação, é graduado em Ciência da Computação pela Universidade Católica de Goiás - PUC-GO (1999) e doutor em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo - USP (2009). Atualmente é empregado efetivo da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) estando lotado na Embrapa Informática Agropecuária. Tem experiência na área de Ciência da Computação, com ênfase em projeto e desenvolvimento de Sistemas Internet/Web e Processamento de Língua Natural, atuando principalmente nos seguintes temas: terminologia computacional, lingüística de corpus, lingüística computacional, data mining, ontologias, thesaurus eletrônicos, web semântica, testes adaptativos informatizados e ambientes educacionais.

**Eduardo Antonio Speranza**, responsável por plano de ação e mediador, Bacharel em Ciências de Computação pelo Instituto de Ciências Matemáticas e de Computação - ICMC/USP. Mestre em Engenharia Elétrica pela Escola de Engenharia de São Carlos - EESC/USP. Atuação no desenvolvimento de software para Agricultura de Precisão. Atuação na Coordenação Geral de Tecnologia da Informação do Ministério da Agricultura (MAPA), na gestão de sistemas corporativos. Atualmente é Analista de Sistemas da Embrapa Informática Agropecuária, atuando no desenvolvimento de Sistemas de Informações Geográficas e Bancos de Dados Espaciais em projetos de pesquisa no Laboratório de Geotecnologias

**Fabiano Fernandes dos Santos**, responsável por atividade, graduado em Ciência da Computação pela Universidade Federal de Itajubá (2007) e mestrado em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo (2010). Atualmente é aluno de doutorado da Universidade de São Paulo, atuando principalmente nos seguintes temas: mineração de texto, descritores para agrupamentos hierárquicos, recuperação de informação contextual e regras de associação.

**Solange Oliveira Rezende**, responsável por atividade, mediadora, graduada em Licenciatura em Ciências Habilitação Matemática pela Universidade Federal de Uberlândia (1986), mestrado em Ciências de Computação e Matemática Computacional pela Universidade de São Paulo (1990) e doutorado em Engenharia Mecânica - São Carlos pela Universidade de São Paulo (1993). Atualmente é professor associado da Universidade de São Paulo. Tem experiência na área de Ciência da Computação, com ênfase em Metodologia e Técnicas da Computação - Inteligência Artificial, atuando principalmente nos temas relacionados com mineração de dados e textos.

Texto: Maria Fernanda Moura  
Criação e produção: Neide Makiko Furukawa  
Embrapa Informática Agropecuária - Núcleo de Comunicação Organizacional (NCO)  
Tiragem: 50 exemplares - 26 de abril de 2012

Mais informações: [sac@cnptia.embrapa.br](mailto:sac@cnptia.embrapa.br)



Ministério da  
Agricultura, Pecuária  
e Abastecimento



Workshop  
Projeto TIENA

# Tecnologias inovadoras em mineração de textos para a especialização de notícias agrícolas - piloto cana-de-açúcar

3 a 4 de maio de 2012

Embrapa Informática Agropecuária  
Campinas, SP

