

**AtoZ: Novas Práticas em Informação e Conhecimento**  
www.atoz.ufpr.br

Universidade Federal do Paraná  
Setor de Ciências Sociais Aplicadas  
Curso de Gestão da Informação  
Av. Prefeito Lothário Meissner, 632 - Campus III  
Jardim Botânico  
80210-170 - Curitiba, PR - Brasil  
Fone: +55(41)3360-4389  
Fax: +55(41)3336-4471  
E-mail: revistaatoz@ufpr.br

#### **Comitê Editorial**

Dra. Patrícia Zeni Marchiori  
UFPR / DECIGI, Brasil  
Msc. Eduardo Michelotti Bettoni  
UFPR / DETUR, Brasil  
Dra. Helena Nunes Silva  
UFPR / DECIGI, Brasil  
Dra. Denise Fukumi Tsunoda  
UFPR / DECIGI, Brasil

#### **Editores de Seção – Artigos**

Dra. Patrícia Zeni Marchiori  
UFPR / DECIGI, Brasil  
Msc. Eduardo Michelotti Bettoni  
UFPR / DETUR, Brasil

#### **Editores de Seção – Entrevistas**

Andre Luiz Appel  
*Conselho Regional de Psicologia 8ª Região (CRP-08), Brasil*

#### **Editores de Leiaute**

Andre Luiz Appel  
*Conselho Regional de Psicologia 8ª Região (CRP-08), Brasil*

#### **Diagramação Web**

Msc. Eduardo Michelotti Bettoni  
UFPR / DETUR, Brasil

#### **Projeto Gráfico**

Andre Luiz Appel  
*Conselho Regional de Psicologia 8ª Região (CRP-08), Brasil*

#### **Colaboradores**

Amanda Cristina Santos Costa (Diagramação)  
*Curso de Gestão da Informação / UFPR, Brasil*  
Prof. Javier Canzani Cuello (Resumos em castelhano)  
*Escuela Universitaria de Bibliotecología y Ciencias Afines (EUBCA) / Universidad de la República, Uruguay*  
Msc. Lígia Leindorf Bartz Kraemer (Normalização)  
*UFPR / DECIGI, Brasil*  
Luis Antônio Borges Filho (Logomarca)  
*Curso de Gestão da Informação / UFPR, Brasil*  
Dr. Mauro José Belli (Apoio Técnico)  
*UFPR / DECIGI, Brasil*  
Msc. Viviane Helena Kuntz (Projeto Gráfico)  
*Laboratório de Mídias Digitais (LabMidi) / GPCIT / UFPR, Brasil*

---

AtoZ : Novas Práticas em Informação e Conhecimento. – Vol. 1, n. 1 (jan./jun. 2011)- . –  
Curitiba : Universidade Federal do Paraná, Curso de Gestão da Informação, 2011- .  
v. ; 18 cm.

Semestral.  
Publicação online: <<http://www.atoz.ufpr.br>>  
ISSN [em processo]

1. Comunicação científica – Periódico. 2. Informação – Periódico. 3. Conhecimento – Periódico.  
I. Curso de Gestão da Informação. II. Universidade Federal do Paraná.

CDD 001(8162)

## Bem vindos à AtoZ!

A Gestão da Informação, enquanto atividade acadêmica no Brasil, surgiu a partir do pioneirismo na criação do curso de Gestão da Informação da Universidade Federal do Paraná. Com uma perspectiva interdisciplinar, que agrupa conhecimentos da Ciência da Informação, Administração e Tecnologia da Informação, este Curso volta-se ao estudo da informação enquanto fenômeno social, econômico e tecnológico capaz de contribuir com as organizações em seus diferentes processos, produtos e serviços.

O universo da produção científica nas instituições de ensino superior é vasto, estimulante e complexo, devido às ramificações das matérias do conhecimento e aos modos de apreensão dessas matérias, uma vez que ao saber científico não se podem estabelecer fronteiras. O campo específico, para o qual se volta esta revista, a Gestão da Informação e do Conhecimento, apresenta especificidades devido à característica transdisciplinar com que atua nos campos teóricos e práticos ao romper os limites disciplinares para comprometer-se com a descoberta, a criação e aplicação do seu domínio como resposta as necessidades da sociedade.

Os canais para publicação da produção científica no Brasil, em diferentes áreas, voltam-se aos estudos desenvolvidos por pesquisadores experientes, enquanto aqueles menos experientes ou incipientes dificilmente conseguem publicar seus resultados. Neste sentido, a Revista **AtoZ: novas práticas em informação e conhecimento** tem como proposta possibilitar aos jovens pesquisadores a divulgação dos seus estudos (especialmente aqueles derivados de trabalhos de conclusão de curso e dissertações), cujas metodologias ou abordagens sejam inéditas. A produção da AtoZ tem caráter inovador ao propor, na plataforma OJS/SER, um espaço potencial para aplicação de conteúdos (e como laboratório de práticas) de diferentes disciplinas do curso de graduação em Gestão da Informação, em nível estratégico, gerencial e operacional da revista.

A entrevista sobre questões metodológicas em Ciências Sociais, com a professora doutora Leilah Santiago Bufrem abre o número inaugural da Revista AtoZ. Os artigos publicados espelham o caráter interdisciplinar da área, com contribuições nos temas indexação e recuperação; produção tecnológica da Universidade Estadual de Londrina; metodologias para compartilhamento de conhecimento; plano estratégico em ambiente acadêmico; e mineração de dados.

O artigo intitulado indexação e recuperação de teses e dissertações por meio de sintagmas nominais utiliza uma ferramenta para análise de textos (OGMA) na extração de sintagmas nominais para indexação automática e recuperação de informações de teses e dissertações no Banco de Teses e Dissertações da Universidade Federal de Pernambuco.

O mapeamento da produção tecnológica da Universidade Federal de Londrina é realizado a partir do Currículo Lattes de 51 docentes da área de Ciências Agrárias.

O compartilhamento de conhecimento é investigado sob a perspectiva das metodologias: Investigação Apreciativa e Facilitação Gráfica, utilizadas no Global Fórum América Latina.

A aplicação de um Plano estratégico ao Grupo de Pesquisa Aplicada em Ciência, Informação e Tecnologia (GP-CIT) apresenta como contribuição o valor agregado a processos e decisões dando visibilidade ao Grupo.

A mineração de dados é utilizada na busca de conhecimento, ou seja, de encontrar padrões que relacionem a evolução do canto, a biologia e o micro-habitat da família Thamnophilidae (Aves).

Pretende-se, ao se oferecer este primeiro número de uma nova revista, entre tantas já existentes no Brasil na área de Ciência da Informação, destacar seu enfoque diferenciado. Espera-se que seja um marco na divulgação de trabalhos de pesquisadores jovens e promissores.

Boa Leitura, e que esta o estimule a ser um colaborador na **AtoZ**!

**Professora Dr<sup>a</sup> Helena de Fátima Nunes Silva**  
Coordenadora do Curso de Gestão da Informação

# Questões de metodologia – parte 1

Leilah Santiago Bufrem

Leilah atualmente é Professora Titular do Curso de Gestão da Informação, do Programa de Pós-Graduação em Educação e do Mestrado em Ciência, Gestão e Tecnologia da Informação da Universidade Federal do Paraná. Membro de Conselhos Assessores de periódicos científicos. Coordenadora do Grupo de Pesquisa Educação, Pesquisa e Perfil profissional em Informação, pesquisadora do Grupo de Pesquisa Cultura, práticas escolares e educação histórica e do Núcleo de Produção Científica (NPC) da Escola de Comunicações e Artes/USP. Membro do Conselho Fiscal da International Society for Knowledge Organization (ISKO-Brasil). Tem experiência na área de Ciência da Informação e Educação, com ênfase em Metodologia e Produção Científica. Graduada e licenciada em Filosofia pela Pontifícia Universidade Católica do Paraná (1964), graduada em Biblioteconomia e Documentação pela Universidade Federal do Paraná (1963). Mestre em Educação pela Universidade Federal do Paraná (1981), doutora em Ciências da Comunicação pela Universidade de São Paulo (1991), pós-doutora pela Universidad Autónoma de Madrid (1995) e quatro especializações (Ação Cultural – 1987; Cultura Portuguesa – 1975; Métodos e Técnicas de Ensino – 1975; Teoria do Conhecimento – 1974). Tem livros e artigos publicados e orienta teses e dissertações de mestrado e doutorado.



## Palavras-chave

Metodologia da pesquisa; Ciências Sociais Aplicadas; Ciência da Informação; Pesquisas científicas.

## Keywords

*Scientific method; Social Sciences; Information Science; Scientific research.*

Antes de iniciar as reflexões originadas das perguntas, enfatizo o agradecimento aos estudantes que, com o entusiasmo apostolar de pesquisadores iniciantes, construíram um caminhar inteligente e instigante, com questões geradoras deste repensar da pesquisa. Elas mereceriam não modestas respostas como as que tentarei construir, mas ensaios ou teses, com recuperação de casos concretos, de teorias consolidadas e de transgressões desafiantes, não fossem as limitações de tempo e pertinência, em se tratando de uma entrevista para um número inaugural de periódico. E aos organizadores desse fascículo, responsáveis pela oportunidade, também agradeço a honra que me conferem, tentando vincular esforços para um resultado compatível com o valor das questões levantadas e o entusiasmo de realização de uma edição inaugural.

Leilah Santiago Bufrem

**1. Existe um roteiro a seguir para a delimitação do tema de uma pesquisa, de modo a torná-la bem estruturada e exequível? Qual seria esse roteiro? Escolha pessoal ou imposta?**

Estudos sobre escolhas e procedimentos de pesquisa científica reservam especial atenção à definição do tema e do problema, uma das fases mais significativas do processo, pelos cuidados e preciso que essa tarefa requer. Isso porque, somente com um sentido evidente do problema, após a realização de uma preliminar, mas cuidadosa análise da literatura, há condições de identificar aspectos, pouco ou nada explorados anteriormente, sobre o tema escolhido. Já nos assinalou Popper que a ciência ou o conhecimento não começaria de percepções ou observações, de uma coleção de fatos ou números, porém, mais propriamente, de problemas. Assim, não pode haver conhecimento sem problemas, mas tampouco há problema sem conhecimento, assim como não há problema sem ignorância (POPPER, 1978).

Pode-se mesmo afirmar que quanto maior é o conhecimento sobre um determinado tema, mais problemas são levantados pelo investigador que o domina, razão pela qual a análise da literatura por meio de leituras preliminares é especialmente importante, não apenas para aqueles cujo conhecimento sobre o tema é incipiente ou insatisfatório, mas também para os quais o aprofundamento nas questões permite identificar meandros, variáveis ou características capazes de provocar novas interrogações.

Percebe-se, portanto, em concordância com o argumento de Bachelard, que na vida científica os problemas não se formulam de modo espontâneo. Com efeito, a experiência nos mostra que eles podem ser reconhecidos nas experiências cotidianas de trabalho ou estudo, a partir das leituras teóricas, das necessidades pessoais ou profissionais, mas serão sempre uma forma de contradição a partir da contraposição ativa e consciente do sujeito com o meio circundante. Pode-se afirmar que, enquanto modo especial de contradição dialética, um problema representa uma dificuldade teórica ou prática, no conhecimento de algo cuja importância exige solução, em qualquer domínio do saber.

Mas somente a experiência não bastaria para responder a essa pergunta, pois, como diria Vandré (1968),

Quanto mais eu ando, mais vejo estrada  
Mas se eu não caminho, eu sou é nada.

Fazer pesquisa é um caminhar intencional que permite a visão de muitas estradas e rumos. Impõe-se o reconhecimento desses caminhos, intenção que justifica a recorrência ao testemunho daqueles com quem caminhamos e dos quais usufruímos saberes, procurando fundamentar teoricamente nossas experiências. Isso porque o caminho não é um andar às cegas e a trajetória da pesquisa não se faz por ensaio e erro.

Quando a escolha do tema e do problema é pessoal, a pesquisa adquire um caráter histórico, pois suas questões decorrem de estudos e experiências anteriores. Assim, a escolha nunca se verifica no vazio, pois fatores internos como a curiosidade, a imaginação e a criatividade, aliados ao referencial de conhecimentos pessoais e às concepções filosóficas do pesquisador estimulam a sua resolução. Por outro lado, há pressões do contexto, razões políticas, institucionais ou empresariais, que podem ainda ser locais, regionais, nacionais ou internacionais. Além desses fatores, também podem contribuir para o exercício da escolha as determinações legais. Entretanto, sempre se deve levar em conta critérios como a *relevância* do problema, isso é, o interesse intelectual ou prático que ele revele. A esse critério estariam relacionados fatores como a *oportunidade* de se tentar resolvê-lo, o *comprometimento*, o *reconhecimento dos modismos* para evitá-los, assim como as *implicações políticas, sociais, econômicas ou culturais* concorrentes para sua resolução.

A modo de roteiro, podemos nos pautar nos seguintes critérios, tanto quando a escolha é livre e pessoal, quanto em caso de imposição do tema: o problema deve traduzir uma *dificuldade, preocupação, ou necessidade* sentida; sua definição

implica o *reconhecimento* do objeto, sujeitos ou agentes para os quais se voltará a pesquisa, do meio ambiente ou circunstâncias que os envolvem, dos fatores estimulantes ou influências mais evidentes, assim como das modificações, respostas ou alterações advindas da solução do problema, portanto, ele deve ser *relevante* e

sua resolução, significativa e viável. O terceiro critério seria a *originalidade*, seja em relação ao seu objeto, seja em relação ao modo como será pesquisado esse objeto, pois um problema pode ser relevante, mas não necessariamente original e vice-versa. Conciliar esses três critérios tem sido tarefa decisiva para o pesquisador.

## **2. Em um relato de pesquisa, qual o papel da introdução, e que informações deve conter? Citações na introdução podem conflitar com aquelas do referencial teórico?**

A introdução situa o problema em seu contexto teórico e empírico, permitindo a compreensão do que vai ser apresentado ao longo da pesquisa, podendo conter um breve histórico sobre o problema definido, assim como as motivações que levaram o autor a realizar o estudo. Como parte inicial do texto, a introdução deve incluir os elementos necessários para contextualizar o tema e o problema, informar sobre os objetivos e o objeto da pesquisa, indicando sucinta e claramente o assunto, a área ou o campo de conhecimento e o período abrangidos pela pesquisa, enquanto elementos delimitadores, assim como a natureza e a importância do problema. Justificar o trabalho significa valorizar os aspectos que caracterizem sua relevância científica, social, institucional, legal

ou mesmo pessoal, esclarecendo o ponto de vista que o orienta. Parte do referencial teórico, com vistas a delimitar o objeto da pesquisa ou indicar a importância do problema, às vezes é incluída na introdução, o que não requer exaustividade, pois deve conter as referências necessárias para o embasamento dos pressupostos do trabalho. A introdução pode finalizar com uma questão de pesquisa ou com a formulação de hipóteses.

Também é na introdução que enunciamos os tópicos principais do texto, o roteiro ou ordem de exposição, sem, entretanto, mencionarmos os resultados, embora indicando que chegamos a eles.

## **3. O objetivo geral da pesquisa deve ser, obrigatoriamente, a resposta ao problema?**

Prefiro trocar o termo obrigatoriamente por preferencialmente. Parece evidente que quando se define um problema como forma de orientar o estudo, o objetivo geral pode ser considerado o alvo da pesquisa, portanto, relacionado ao conteúdo intrínseco do problema. Ele esclarece e direciona o foco da pesquisa de modo abrangente e normalmente é redigido em uma frase, iniciada com o verbo no infinitivo. O objetivo geral pode se constituir, portanto, em proposta de solução a um problema crítico a resolver ou que explore uma oportunidade de transformação ou aperfeiçoamento. Pode, também, descrever o comportamento de uma característica ou

conjunto de características. Deve, portanto, ser claro, de modo a explicar o que realmente queremos com o estudo.

Além de definir o campo no qual está inserido o objeto da pesquisa, a enunciação do objetivo geral auxilia a delimitação da abrangência do estudo, a partir do que se definem modos de raciocínios e procedimentos a adotar. Portanto, além de se relacionar ao conteúdo intrínseco da pesquisa, o objetivo geral contribui para a explicitação do problema e deve ser enunciado quando se introduz o trabalho, para expressar claramente o que se pretende com a investigação.

**4. A delimitação da pesquisa depende de clareza e bom delineamento, também quanto à ortografia e à gramática. Nesse sentido, como identificar verbos ideais para a redação de objetivos geral e específicos?**

Na década de 1950, a *American Psychological Association* formou um grupo, liderado por Benjamim Bloom, para criar uma classificação de objetivos de processos educacionais, como forma de apoio à estruturação, organização, definição de objetivos e escolha de instrumentos de avaliação. Essa taxonomia partiu basicamente de três áreas não mutuamente exclusivas: a cognitiva, ligada ao saber, a afetiva, ligada a sentimentos e posturas e a psicomotora, ligada a ações físicas. A elas relacionam-se níveis de complexidade tais como: conhecimento; compreensão; aplicação; análise; síntese, avaliação; recepção; resposta; valorização; organização (de valores) e internalização. Esses níveis integram-se às ações dos sujeitos em seus contextos específicos e adquirem significados expressos em objetivos condizentes com as capacidades a adquirir. Como exemplo, quando se objetiva o conhecimento sobre critérios, regras e procedimentos, se esperam capacidades como adquirir, identificar, reconhecer, entre outras possibilidades expressas por verbos no infinitivo.

Como os processos de pesquisa e de ensino/aprendizagem são considerados indissociáveis, os objetivos de Bloom, originariamente destinados a facilitar a avaliação pedagógica têm sido um apoio mestre para a formulação de objetivos de pesquisa. Sua definição clara e estruturada, considerando-se a construção de conhecimento e de resultados adequados aos problemas que expressam as necessidades ou contradições a serem solucionadas, direcionará o processo de pesquisa para a escolha adequada de estratégias, métodos, delimitação do conteúdo específico, técnica e instrumentos de pesquisa e, consequentemente, para a avaliação dos resultados provisórios ou definitivos do processo de pesquisa.

A taxonomia de Bloom, portanto, possibilitou o ajustamento da linguagem no meio acadêmico e, com isso, o aperfeiçoamento do modo de definição, tanto de objetivos gerais, quanto específicos.

**5. Considerando o seguinte exemplo: Quando o ambiente para a pesquisa de campo é de alunos de graduação em Gestão da Informação, e precisamos delimitá-lo com base em critérios tais como: o grupo deverá ser composto de alunos cujos pais não têm curso superior; alunos que estudaram em escolas públicas; alunos que moram a mais de dez quilômetros da Universidade; alunos de ambos os gêneros; o que difere/delimita a composição de um universo e a composição de uma amostra?**

É importante distinguir inicialmente o que significa delimitação, recorte, universo e amostra. Ao delimitarmos um problema de pesquisa estamos definindo limites do que estará ou não dentro de nosso campo de estudo, ou seja, definiremos os aspectos, características ou variáveis concernentes ao objeto de pesquisa. Também fazem parte desse processo demarcatório os recortes espaciais e o período de tempo relativo à exposição do objeto ou à realização da pesquisa. Essa delimitação é necessária à adequada definição e futura resolução do problema pois, além de descrever a área em que se insere, define a abrangência do estudo. Deve tanto determinar

o que vai ser tratado no estudo, quanto o que não vai sê-lo.

Os termos universo e população em geral têm sido tratados como sinônimos, para significar o conjunto de todos os casos que se adequam a algum conjunto de especificações pré-definidas, ou o conjunto de todos os indivíduos aos quais se aplicam características ou propriedades definidas.

A amostra seria o subconjunto do universo ou população, a partir do qual os dados são recolhidos e pelo qual se estabelecem ou se estimam as características selecionadas do todo,

na impossibilidade de estudá-lo em sua extensão. Aproveitando o exemplo dado por vocês, quando temos um universo de estudantes de graduação em Gestão da Informação, podemos definir inicialmente se ele abrange os indivíduos de todo o Brasil ou se o foco recairá sobre estudantes de outros países. No primeiro caso, teremos um universo mais reduzido e ele será menor ainda se nosso foco se dirigir apenas aos estudantes da Universidade Federal do Paraná. Teremos então um novo universo, delimitado a uma instituição. Se continuarmos utilizando critérios, dessa vez em relação àqueles estudantes cujos pais não têm curso superior, teremos um grupo ainda mais reduzido e, assim por diante, em relação àqueles que estudaram em escolas públicas, ou àqueles que moram a mais de dez quilômetros da Universidade. Podemos ainda focalizar estudantes de ambos os gêneros ou especificamente do gênero feminino. Sempre que, intencionalmente, definimos critérios para a escolha de indivíduos que participarão de nosso estudo em prol da resolução do problema proposto, estaremos constituindo novos universos e, quanto mais focamos e atribuímos qualidades específicas a um grupo, estaremos, a cada qualidade atribuída, constituindo novo universo, ou seja, um conjunto de indivíduos aos quais se aplica a ideia, o que não se caracterizará, entretanto, como amostra. Caso seja necessário um procedimento amostral,

aproveitando o mesmo exemplo que serviu para ilustrar a questão, poderíamos optar por uma amostra aleatória, realizando um sorteio entre todos os estudantes, a partir de uma lista completa de seus nomes e a utilização da tabela dos números aleatórios ou de um aparato de sorteio. Poderíamos, ainda, retirar de cada uma das turmas do Curso um percentual idêntico de estudantes para formar uma amostra representativa do todo.

Outra possibilidade é recorrer a uma amostra não-aleatória também chamada de não-probabilística, ou seja, uma amostra em que nem todos os indivíduos tenham a mesma probabilidade estatística de entrar na sua composição. Neste caso, não é possível assegurar que seja representativa de toda a população, nem inferir o grau de confiança para os parâmetros da população que nos interessam. É o caso da amostragem de conveniência, ou por acessibilidade, que pode ser ilustrada com o mesmo exemplo. Se solicitarmos aos estudantes do Curso, presentes na cantina em determinado e oportuno horário, que respondam ao nosso questionário, estaremos utilizando este tipo de amostra, com menor valor científico, pois não é representativa do todo e da qual não se pode calcular a margem de erro, embora seja válida como ferramenta exploratória.

## 6. Há limites - e quais são os impactos - de um número reduzido de respondentes em pesquisas qualitativas (tendo como enfoque a técnica de entrevista)?

A realidade alvo de qualquer estudo científico apresenta múltiplas facetas, do que resultam possibilidades diferenciadas de entendê-la. Voltamo-nos a ela com intenções explicitadas pela conduta da investigação, diante dos elementos que a determinam e que permitem a aproximação intelectual, na interação entre sujeito e objeto de pesquisa. A construção do objeto empírico nem sempre é fácil de concretizar devido às peculiaridades da realidade quando a ênfase recai nos aspectos qualitativos.

Uma das características distintivas de uma pesquisa com enfoque qualitativo é a busca

de entendimento de um objeto ou fenômeno específico em profundidade. São priorizadas descrições, comparações e interpretações, em detrimento de estatísticas, regras e outras generalizações.

A pesquisa denominada qualitativa é mais participativa e, portanto, menos controlável. Os participantes da pesquisa podem direcionar o rumo da pesquisa em suas interações com o pesquisador. É a de ser mais adequada quando se deseja obter informações refratárias às medidas estatísticas, conservando-se a forma literal dos dados, o que enseja a percepção dos elos lógicos

e das relações entre fenômenos. Além de facilitar a percepção das relações de sentido entre as unidades ou as categorias que as unem, a pesquisa com enfoque qualitativo, pode penetrar em coisas concretas, mas ocultas, pois facilita a visão de elementos nem sempre notados ao primeiro olhar, o que na concepção de Ginzburg seriam os detritos ou refugos de nossa observação.

De um modo geral, pesquisas sob enfoque qualitativo realizam entrevistas, que exigem a definição de critérios para seleção dos sujeitos que comporão o universo ou amostra de investigação. A quantidade de componentes do grupo a entrevistar dificilmente pode ser determinada *a priori*, pois isso vai depender da qualidade das informações obtidas em cada entrevista, assim como da profundidade e do grau de recorrência e divergência das informações. A organização gradativa dos depoimentos relativos ao objeto da investigação vai ampliando as possibilidades de análise até que se atinja o chamado “ponto de saturação”, quando se dá por encerrado o trabalho.

Mas quando é muito amplo o universo ou população de estudo, coloca-se a questão da amostra e de sua representatividade. Há os defensores do procedimento de amostragem, para que os sujeitos participantes do processo de coleta de informações se constituam em grupo representativo do universo. Outro grupo defende a posição de busca exaustiva, de modo que todos os membros da população sejam consultados.

O que normalmente consideramos como amostra pode se converter em apenas um ou dois sujeitos numa pesquisa qualitativa, pois o critério de representatividade da amostra para esse tipo de pesquisa não é numérico. Essa amostra deve permitir que o pesquisador seja capaz de conhecer bem a realidade ou objeto em sua complexidade, intensificando sua busca nas características e qualidades perceptíveis, sem que a generalização dificulte seu propósito. Portanto, crucial para a definição da amostra é identificar os sujeitos cujas informações sejam mais significativas em relação ao problema investigado. Pode ser considerada uma amostra ideal aquela que reflete as múltiplas

dimensões do objeto de estudo.

É comum citar a entrevista como técnica exemplar de uma pesquisa com enfoque qualitativo, pois enquanto conversação pessoal entre um investigador que sabe o que deseja e um ou mais entrevistados que conhecem ou dispõem da fonte ou da informação desejada, ela favorece o aparecimento de sentimentos subjacentes a uma opinião expressa, assim como a obtenção de informações tanto complexas como emocionalmente carregadas.

A entrevista permite que se observem também as atitudes dos entrevistados, para que possam ser analisados profundamente, além das informações, a expressão de sentimentos, sensações, percepções, pensamentos, intenções, os comportamentos passados, entendimento de razões, significados e motivações dos sujeitos em relação ao objeto de pesquisa. A quantidade de pessoas entrevistadas deve, no entanto, permitir a reincidência de informações ou saturação dos dados, situação ocorrida quando nenhuma informação nova é acrescentada, portanto, a continuidade do processo de pesquisa torna-se inútil devido à redundância de informações. Portanto, os sujeitos sociais devem ser pesquisados em número suficiente para que se possa definir o momento de encerramento do processo, mas não exceder a dimensão de um estudo que não se volta à exaustividade, mas ao movimento deliberado de aprofundamento, numa superação dialética em prol de uma análise intensiva da realidade.

Resta acrescentar a essas reflexões, a partir das questões enunciadas, que a singularidade de cada situação de pesquisa impede fórmulas rígidas, padrões ou normas para a sua realização. A cada dúvida diante de decisões a tomar, especialmente do ponto de vista metodológico, desdobram-se modalidades alternativas e a cada realidade brotam sentidos diversos. Devemos tentar manter em nossa prática de pesquisa uma interação íntima entre o método como disciplina e o estudo em andamento, sempre conscientes do sentido do problema e determinados pela paixão para resolvê-lo.

## Referências

POPPER, K. **La lógica de las ciencias sociales.**

México: Grijalbo, 1978.

VANDRÉ, G.; ACCIOLY, H. **O Plantador.** Álbum

Canto Geral: 1968.

---

Questões elaboradas pelos alunos do Programa de Pós-Graduação (Mestrado) em Ciência, Gestão e Tecnologia da Informação (PPCGI/SA/UFPR) e alunos do Curso de Graduação em Gestão da Informação (CGI/UFPR).

---

# Indexação e recuperação de teses e dissertações por meio de sintagmas nominais

Renato Fernandes Corrêa<sup>1</sup>, Darliane Goes de Miranda<sup>2</sup>, Camila Oliveira de Almeida Lima<sup>3</sup>, Tiago José da Silva<sup>4</sup>

## Resumo

**Introdução:** Aborda a utilização dos sintagmas nominais no processo indexação automática das teses e dissertações depositadas na Biblioteca Digital de Teses e Dissertações da UFPE (BDTD-UFPE), considerando a hipótese de que os sintagmas nominais consistiriam numa melhor unidade de conhecimento para a indexação e recuperação de informação que as palavras isoladas, permitindo aumentar a satisfação da necessidade de informação do usuário durante a busca pela informação. Discute-se sobre o estado da arte dos sintagmas nominais e de sua extração automática, bem como sua aplicação na indexação automática e recuperação de informação. **Método:** Analisa, com base em ferramenta para análise de texto (OGMA), a aplicabilidade da extração de sintagmas nominais na indexação automática e recuperação de informação de teses e dissertações no contexto da BDTD-UFPE. Com base em resumos da área de Direito, Computação e Nutrição, definiram-se os valores totais para cada uma das variáveis observadas, o que permitiu avaliar a extração de sintagmas nominais através dos percentuais de precisão de sintagmas nominais relevantes; da taxa de erro ao extrair cadeias de caracteres que não constituem sintagmas nominais, e; do percentual de sintagmas nominais extraídos não relevantes. **Resultados:** O processo de extração de sintagmas nominais através do OGMA teve diferentes desempenhos para cada programa de pós-graduação, sendo obtido melhor desempenho (melhor índice de precisão) para resumos de Direito, seguidos dos de Computação e Nutrição. Esta diferença de desempenho pode em parte ser explicada pela diferente natureza dos termos técnicos presentes nos resumos. **Conclusões:** Embora existam limitações nas ferramentas disponíveis, a aplicação de métodos automatizados de extração e indexação por sintagmas nominais mostra-se promissora, pois os sintagmas nominais se configuram como melhores descritores e pontos de acesso aos documentos, eliminando os problemas causados pela sinonímia e a polissemia das palavras isoladas.

## Palavras-chave

Sintagmas nominais; Recuperação de informação; Indexação automática; Teses e dissertações.

## Introdução

Atualmente, vivencia-se uma transição das atividades relacionadas com a manipulação, a edição, o armazenamento, a distribuição das informações e a sua recuperação da forma impressa para a forma digital.

Durante os últimos anos, um volume crescente de informações tem sido registrado em várias bases de dados, nos mais diversos domínios do

conhecimento e sob diversas formas (numéricas, textuais, imagens etc.).

Muitas pesquisas vêm tentando contribuir para enfrentar alguns dos muitos desafios que surgem quando lidamos com massivas quantidades de dados, como nos grandes acervos de documentos digitais, notadamente quando estes precisam ser regularmente organizados, pesquisados, recuperando em tempo hábil informação relevante para algum objetivo específico.

Um dos grandes desafios encontrados na recuperação de informação é como atender às necessidades de informação do usuário de forma rápida e eficaz. Várias pesquisas foram e continuam sendo realizadas com o propósito de aumentar a precisão e revocação dos resultados de forma que o usuário possa encontrar todos os documentos que atendem às suas necessidades de informação.

O rápido desenvolvimento de novas tecnologias da informação criou um ambiente propício para o experimento de novos métodos de recuperação e de indexação da informação. A extração automática de sintagmas nominais faz parte da tentativa de indexar os documentos automaticamente.

O objetivo deste trabalho é discutir como os sintagmas nominais podem ser utilizados na indexação automática de teses e dissertações e prover uma melhor recuperação de informação na Biblioteca Digital de Teses e Dissertações da UFPE (BDTD-UFPE). Através de estudos de casos, analisa-se a extração automática de sintagmas nominais e a utilização dos mesmos como descritores e pontos de acesso a teses e dissertações de três programas de pós-graduação da UFPE.

## Metodologia

Para a realização deste artigo foram inicialmente definidas as formas como essa pesquisa seria estruturada e fundamentada, em termos de revisão de literatura, estudo de uma ferramenta de extração de sintagmas nominais e planejamento e realização dos estudos de caso.

Primeiramente, para coletar informações a cerca do objeto de estudo, foram realizadas buscas por palavra-chave em *sites* como o Google Acadêmico, SciELO e BRAPCI. Os dois últimos escolhidos por oferecerem um acervo de materiais científicos confiáveis e em maior quantidade e o primeiro por ser um *site* de busca de referência mundial.

As buscas utilizadas para tal pesquisa nos três *sites* foram: sintagmas nominais, “sintagmas nominais”, sintagma + nominal, recuperação + informação + sintagma + nominal, extração + sintagma + nominal, sintagma + nominal + indexação.

Após a coleta de artigos científicos, o material foi organizado de acordo com a relevância para este artigo, uma vez que o foco é discutir a forma como os sintagmas nominais podem ser utilizados em sistemas de recuperação de informação. Após o estudo destes artigos, foi elaborada a revisão de literatura, descrita na próxima seção, que busca abordar os seguintes pontos: esclarecer termos como sintagmas nominais, sistemas de recuperação de informação, extração de sintagmas nominais e descrever a estrutura e uso de *softwares* de extração de sintagmas nominais.

Após o levantamento do estado da arte dos sintagmas nominais e sua extração, foram planejados e realizados os estudos de caso. Os estudos de caso consistiram da extração de sintagmas nominais, através da ferramenta OGMA<sup>1</sup>, dos resumos de teses e dissertações de três programas de pós-graduação da UFPE: ciência da computação, direito e nutrição. Foram selecionados trinta resumos, procurando abranger as três grandes áreas do conhecimento (humanas, exatas e saúde), sendo desta forma escolhidos: dez resumos de Direito, dez de Ciências da Computação e dez de Nutrição. Os resultados dos estudos de caso são apresentados na seção Resultados.

Os resumos foram extraídos do sistema TEDE da BD TD-UFPE a partir dos metadados no formato MTD-BR das primeiras teses e dissertações inseridas no sistema para cada programa, ou seja, aquelas com menor valor do campo identificador. Definiu-se como resumo das teses e dissertações os valores dos seguintes seis campos presentes nos metadados das mesmas na seguinte ordem: título, titulação, resumo, assunto (contendo as palavras-chave), nome do programa, grande

<sup>1</sup> Ferramenta para análise de texto, cálculo da similaridade entre documentos e extração de sintagmas nominais (<http://www.luizmaia.com.br/ogma/>)

área do programa. Foi incluído o caractere ponto como separador de campos, bem como múltiplos valores de um campo como no caso das palavras-chave.

O seguinte roteiro foi utilizado para extrair os sintagmas nominais através do OGMA:

- a) abrir o texto de cada resumo na ferramenta;
- b) etiquetar os termos do resumo aberto;
- c) extrair os Sintagmas Nominais Pontuados do resumo etiquetado.

A partir deste roteiro foi possível verificar o resultado dos processos de etiquetagem dos termos e da extração dos sintagmas nominais pelo OGMA.

A avaliação do processo de extração automática de sintagmas nominais dos resumos de cada programa de pós-graduação foi realizada através do cálculo e análise dos percentuais de precisão em extraírem sintagmas nominais relevantes como descritores, a taxa de erro ao extraírem cadeias de caracteres que não constituem sintagmas nominais e o percentual de sintagmas nominais extraídos não relevantes como descritores.

Para cálculo dos percentuais foi necessária a classificação dos possíveis sintagmas nominais extraídos como verdadeiros sintagmas nominais ou falsos sintagmas nominais, sendo esta classificação realizada com base nas definições de sintagmas nominais descritas na revisão de literatura. Bem como a classificação dos verdadeiros sintagmas nominais extraídos em relevantes ou não como descritores para o resumo de onde foram extraídos, sendo esta última realizada com base na análise de assunto do resumo e as palavras-chaves constantes no resumo da respectiva tese ou dissertação.

## Estado da arte

Na presente seção apresenta-se o referencial teórico que norteou a pesquisa e deu

embasamento para a realização e conclusão deste trabalho.

Serão abordados nas subseções seguintes os conceitos de recuperação de informação e análise de texto, definição dos sintagmas nominais e a utilização dos mesmos pelos sistemas de recuperação de informação. Por fim, será apresentada uma ferramenta para extração de sintagmas nominais, descrevendo como se dá o processo de extração automática de sintagmas nominais.

## Recuperação de informação

A recuperação de informação (RI) é uma área que lida com o armazenamento de documentos e a recuperação automática de informação associada aos mesmos (BAEZA-YATES & RIBEIRO-NETO, 1999). Um objeto informacional é geralmente constituído de texto, tais como documentos diversos, páginas web e livros, embora possa conter outros tipos de conteúdo, tais como imagens, áudios, gráficos e figuras. A representação e organização desses objetos devem permitir às pessoas o acesso à informação relevante a partir da expressão de uma necessidade de informação, por exemplo, por meio de uma consulta. A essência da RI consiste na busca de documentos relevantes a uma dada consulta que expressa a necessidade de informação do usuário.

De acordo com Baeza-Yates & Ribeiro-Neto (1999) os sistemas de recuperação de informação (SRIs) são “sistemas que lidam com as tarefas de representação, armazenamento, organização e acesso aos itens de informação” e devem apresentar o conteúdo do documento ao usuário de uma maneira que lhe permita uma rápida seleção dos itens que satisfazem total ou parcialmente a sua necessidade de informação, formalizada através de uma expressão de busca.

Um sistema de recuperação de informação normalmente implementa uma ferramenta de busca passiva de informações que computa a relevância potencial dos documentos encontrados a partir da análise de similaridade,

podendo ser definido como um conjunto de dados padronizados, armazenados em meio eletrônico, utilizados para identificar informação e fornecer sua localização.

Atualmente os SRIs trabalham com diversos modelos para trazer a informação até o usuário. Os modelos utilizados na recuperação de informação têm como base o uso da palavra como unidade básica de acesso à informação. Vários modelos foram desenvolvidos com o objetivo de facilitar o acesso à informação, porém os modelos mais utilizados pelos SRIs são o modelo Espaço Vetorial e o modelo Booleano.

Para que um SRI possa responder as demandas dos usuários com tempos de resposta aceitáveis, primeiramente é preciso que os documentos que estão na base de dados sejam submetidos a um tratamento. Esse procedimento permite a extração de descritores e sua estruturação com o intuito de garantir um rápido acesso às informações (SOUZA et al, 2007). Um descritor consiste de uma palavra ou expressão que identifica, geralmente para fins de indexação, determinado conceito ou tema.

De acordo com Souza (2006), o processo de indexação produzindo um lista de descritores visa à representação dos conteúdos dos documentos com o objetivo de extrair as informações contidas nesses documentos organizando-as para permitir a recuperação destes últimos. Porém, na maioria dos SRIs atuais, os descritores não passam de palavras isoladas extraídas dos documentos que constituem o acervo.

Os primeiros estudos sobre avaliação de sistemas da informação relacionavam o nível de satisfação do usuário com o resultado apresentado pelo sistema de recuperação de informação, sendo quantificado pela porcentagem de resultados relevantes retornados por sua pesquisa. À medida que novos métodos, instrumentos e ferramentas com foco na relevância do resultado retornado dos SRIs foram utilizados, surgiu também a necessidade de identificar qual método - ou quais - seria capaz de retornar o maior número

de documentos relevantes para o usuário como resultado de uma consulta.

### Análise de texto

Segundo Maia (2008), a análise de texto corresponde a uma área que envolve outras subáreas como, por exemplo, a mineração de texto e a área de processamento de linguagem natural (PLN). A PLN também é uma subárea da inteligência artificial e da linguística que estuda os problemas da geração e tratamento automático de línguas naturais.

A mineração de textos refere-se ao processo de extração de informação útil (que possa gerar conhecimento) em documentos de textos não-estruturados. É basicamente um conjunto de métodos usados para organizar e descobrir informações em bases de textos. Com base no conhecimento a ser extraído, a mineração de textos define técnicas de extração de padrões ou tendências de grandes volumes de textos em linguagem natural, normalmente, para objetivos específicos.

Já o processamento de linguagem natural, de acordo com Maia (2008), trata-se de um conjunto de métodos formais para analisar textos e gerar frases escritas em um idioma humano. Um dos objetivos finais da PNL é fornecer aos computadores a capacidade de entender e compor textos, ressaltando que “entender” um texto significa reconhecer o contexto, fazer análise sintática, semântica, léxica e morfológica, criar resumos, extraír informação, interpretar os sentidos e até aprender conceitos com os textos processados.

É neste contexto de análise textual que são apresentados os sintagmas nominais, pois com o aparente esgotamento das estratégias atuais de indexação e representação de documentos, faz-se necessário investigar novas abordagens para sistemas de recuperação de informação. Dentre estas abordagens, há uma vertente que busca levar em conta a semântica intrínseca aos documentos textuais, e uma das formas de fazê-

lo é através da utilização de sintagmas nominais como descritores, ao invés de palavras isoladas.

## Sintagmas nominais

Segundo Dubois-Charlier (1977), chama-se sintagma uma sequência de palavras que constituem uma unidade. Um sintagma é uma associação de elementos compostos em um conjunto, organizados em um todo, funcionando conjuntamente. Sintagma significa, por definição, organização e relações de dependência e de ordem à volta de um elemento essencial.

O sintagma nominal “é a menor parte do discurso portadora de informação” (KURAMOTO, 1995), onde os signos linguísticos ligam-se uns aos outros formando grupos ao redor de substantivos. O sintagma nominal quando extraído do texto mantém o significado, o seu conceito. Por exemplo, na frase “O estudo dos sistemas de informação”, poderíamos extrair três sintagmas nominais, seriam eles:

- a) o estudo dos sistemas de informação;
- b) os sistemas de informação;
- c) informação.

De acordo com Miorelli (2001), os sintagmas nominais podem ser entendidos e tratados de forma sintática (privilegiando a forma) ou semântica, buscando os significados mais amplos, cada uma com suas especificidades e implicações. Segundo Perini (1998), os sintagmas nominais possuem duas estruturas: uma à esquerda do núcleo do sintagma que podem ser compostas por determinantes, possessivos, quantificadores e outras classes de palavras e, a estrutura à direita do núcleo que é composta por modificadores, que por sua vez podem ser classes abertas ou outros sintagmas.

A utilização dos sintagmas nominais como estrutura de acesso à informação contida em uma base de dados textual se apresenta como uma alternativa aos sistemas tradicionais de recuperação de informação. Os sintagmas nominais geralmente são extraídos do texto

e analisados a fim de facilitar o processo de indexação automática.

Os sintagmas nominais são compostos de grupos nominais constituídos de uma organização hierárquica em árvore. Diferentemente das palavras, o sintagma nominal quando extraído do texto mantém o significado. Com isso, muitos pesquisadores viram a possibilidade de utilizá-los em um processo de indexação e recuperação de informação.

Segundo Kuramoto (2002), a utilização dos sintagmas nominais na recuperação de informação oferece duas alternativas possíveis de implementação em termos de indexação automática e de interfaces de busca. Uma primeira alternativa seria implementar uma indexação automática nos moldes daquela tradicional baseada em palavras, apenas substituindo os índices contendo as palavras isoladas por índices contendo sintagmas nominais. Uma segunda alternativa seria o aproveitamento da organização hierárquica em árvore dos sintagmas nominais. O aproveitamento dessa organização não apenas cria um novo conceito em termos de indexação, como também introduz inovação em termos de uma interface de busca.

A grande maioria dos modelos de recuperação de informação utilizou e utiliza a palavra como forma de acesso à informação (KURAMOTO, 2002). Entretanto, a palavra possui algumas propriedades que geram alguns inconvenientes, como por exemplo, a sinonímia (concordância dos diversos nomes dados a uma mesma coisa) e a polissemia (ligada ao fato de uma palavra ter muitas significações). Isto gera problemas de ambiguidade não tratados pelos atuais sistemas de recuperação de informação, que podem retornar documentos que contêm palavras com significados diferentes da solicitada pelo usuário.

A utilização dos sintagmas nominais em substituição às palavras isoladas eliminaria estes inconvenientes.

Para que os sintagmas nominais possam ser utilizados nos sistemas de recuperação de informação faz-se necessário realizar o processo de extração dos sintagmas nominais.

### A utilização dos sintagmas nominais no processo de recuperação da informação

A indexação automática, segundo Le Guern (1991), consiste da mesma forma que a indexação realizada pelos indexadores seleciona, em cada documento, os elementos que permitirão ao usuário recuperá-lo, suprindo, assim, a sua necessidade de informação.

Os sistemas de recuperação de informação usualmente adotam termos índices para indexação de documentos, sendo que estes termos índice são usualmente palavras isoladas. Há uma ideia fundamental embutida de que a semântica dos documentos e as das necessidades de informação do usuário pode ser expressa através destes conjuntos de palavras. Porém, isto é claramente, uma grande simplificação do problema, porque grande parte da semântica do documento ou da requisição do usuário é perdida quando se substitui o texto completo por um conjunto de palavras (BAEZA-YATES & RIBEIRO-NETO, 1999, p. 19).

O processo de recuperação de informação baseado em sintagmas nominais, segundo Kuramoto (2002), aponta um potencial natural de organização que, se explorado convenientemente, poderia propiciar aos usuários maior facilidade no uso de um SRI e o retorno de resultados mais precisos em resposta a um processo de busca de informação. Ainda segundo o mesmo autor, a organização baseada em sintagmas nominais permite a navegação na estrutura hierárquica em árvore dos sintagmas nominais.

Considerando o exemplo de sintagma nominal citado anteriormente, ele pode ser organizado segundo os sintagmas que se encontram nele embutidos. Trata-se, portanto, de um sintagma nominal de terceiro nível, dado que ele contém dois outros sintagmas encadeados em seu interior.

A enumeração do nível dos sintagmas nominais poderá ser feita atribuindo-se ao sintagma mais simples (“informação”) o nível 1, ao sintagma que o contém (“os sistemas de informação”) seria atribuído o nível 2 e ao sintagma que contém os dois outros (“o estudo dos sistemas de informação”) seria enumerado como sendo o de nível 3, conforme abaixo:

- a) o estudo dos sistemas de informação (nível 3);
- b) os sistemas de informação (nível 2);
- c) informação (nível 1).

Uma interface de busca, baseada nessa organização, poderia funcionar da seguinte forma:

- a interface de busca aguarda que o usuário forneça um termo ou palavra que represente o centro do sintagma nominal de primeiro nível, por exemplo: informação;
- a partir desse termo a interface de busca recupera todos os sintagmas nominais de primeiro nível que tem “informação” como seu centro. No caso, seriam apresentados diversos sintagmas de primeiro nível que tem “informação” como centro do sintagma nominal, inclusive o sintagma “a informação”. A partir desse nível, o usuário seleciona o sintagma que possa vir atender a sua necessidade de informação. Nesse caso, ele escolheria “a informação” e solicita ao sistema que apresente os sintagmas de segundo nível que possua o sintagma nominal “a informação” em sua estrutura;
- em seguida a interface apresenta todos os sintagmas nominais do segundo nível que possua o sintagma “a informação” em sua estrutura. Essa navegação continua até o momento em que o usuário encontre o sintagma nominal que mais atenda a sua necessidade de informação. Nesse caso ele seleciona o referido sintagma e solicita que a

interface apresente todos os documentos de onde ele foi extraído.

## Ferramentas para extração de sintagmas nominais

Dentro do referencial teórico não foram encontrados muitos trabalhos, nem ferramentas que tinham como foco a extração de sintagmas nominais. A única ferramenta encontrada para

extração de sintagmas nominais em textos escritos em português foi o *software OGMA*.

O OGMA é uma ferramenta para análise de texto, cálculo da similaridade entre documentos e extração de sintagmas nominais. O aplicativo foi desenvolvido por Maia (2008) na ferramenta *Visual Studio.NET* em linguagem C#. O OGMA realiza também a identificação da classe do sintagma nominal, bem como o cálculo da

Quadro 1 – Regras de extração de Sintagmas Nominas do software OGMA

AR ← AD	de ← AR	AV ← AV ad	re ← SU	NS ← MD NS
AR ← AI	de ← PD	MD ← AV MD	de ← PP	NS ← NS pr NS
AJ ← VP	de ← PI	MD ← MD co MD	re ← NP	NS ← NS pr de NS
NU ← NR	qu ← AJ	NS ← NS MD	NS ← re	NS ← NS co NS
NU ← NC	qu ← NU	co ← CO	MD ← qu	NS ← NS co de NS
CO ← VG	qu ← PS	pr ← PR	SN ← NS	NS ← AV NS
CO ← CJ	ad ← AV		AV ← ad	SN ← de SN

Fonte: Maia (2008).

pontuação do mesmo como descritor de forma automática.

Para realizar a extração de sintagmas nominais o OGMA faz uso de um léxico da língua portuguesa construído a partir do vocabulário utilizado pelo dicionário BR.ISPELL e uma lista de 475 palavras irrelevantes criada tendo como base a gramática de Tufano (1990).

O léxico é utilizado para etiquetar cada palavra do texto com as possíveis classes gramaticais correspondentes.

Para resolver problemas de ambiguidade, o OGMA forma uma lista com todas as combinações de etiquetas encontradas para palavras de uma frase e submete cada combinação às regras para extração dos sintagmas nominais. Por exemplo, na frase: “O mato estava grande”, a ferramenta etiquetaria o texto da seguinte forma:

“O/AD mato/VBSU estava/VB grande/AJ”. Posteriormente, o OGMA submete às regras de extração as duas versões da frase etiquetada com diferentes combinações de etiquetas: “O/AD mato/VB estava/VB grande/AJ”; e “O/AD mato/SU estava/VB grande/AJ”.

Para extrair os sintagmas nominais, o OGMA faz uso do conjunto de regras presentes no Quadro 1, aplicando regra por regra na ordem de leitura até obter um sintagma nominal cujo símbolo é SN. Estas regras atuam sobre as etiquetas (que representam as classes gramaticais) atribuídas às palavras, visando marcar o início e fim do sintagma em cada sentença do texto. A correspondência entre as etiquetas presentes nas regras e as respectivas classes gramaticais podem ser obtidas em MAIA (2008).

Os Sintagmas Nominais encontrados entram em uma lista geral de Sintagmas Nominais da frase, e os duplicados são eliminados. Este tratamento possibilitou resolver o problema da ambiguidade de forma bem eficiente.

## Resultados

Com o objetivo de avaliar o processo de extração de sintagmas nominais, utilizou-se a ferramenta OGMA nos resumos de teses e dissertações de três programas de pós-graduação da UFPE: Ciência da Computação, Direito e Nutrição.

Após a extração de sintagmas nominais de cada resumo através do OGMA, contabilizaram-se o número de possíveis sintagmas nominais

**Tabela 1 – Resultado da extração de sintagmas nominais nos resumos de Teses e Dissertações em Direito/UFPE**

Resumo	Possíveis sintagmas extraídos	Não constituem sintagmas	Constituem sintagmas não relevantes	Constituem sintagmas relevantes
D1	46	17	10	19
D2	39	12	4	23
D3	54	18	4	32
D4	55	20	7	28
D5	44	13	6	25
D6	41	9	7	25
D7	55	16	10	29
D8	38	14	9	15
D9	29	5	9	15
D10	63	23	6	34
<b>Total</b>	<b>464</b>	<b>147</b>	<b>72</b>	<b>245</b>

Fonte: os autores.

**Tabela 2 – Avaliação da extração de sintagmas nominais nos resumo de Teses e Dissertações dos programas de Ciência da Computação, Direito e Nutrição da UFPE**

Programa	Precisão	Taxa de Erro	Percentual de não relevantes
Computação	0,47	0,33	0,20
Direito	0,53	0,32	0,16
Nutrição	0,47	0,33	0,20
<b>Média</b>	<b>0,49</b>	<b>0,32</b>	<b>0,19</b>
<b>Desvio</b>	<b>0,03</b>	<b>0,01</b>	<b>0,03</b>

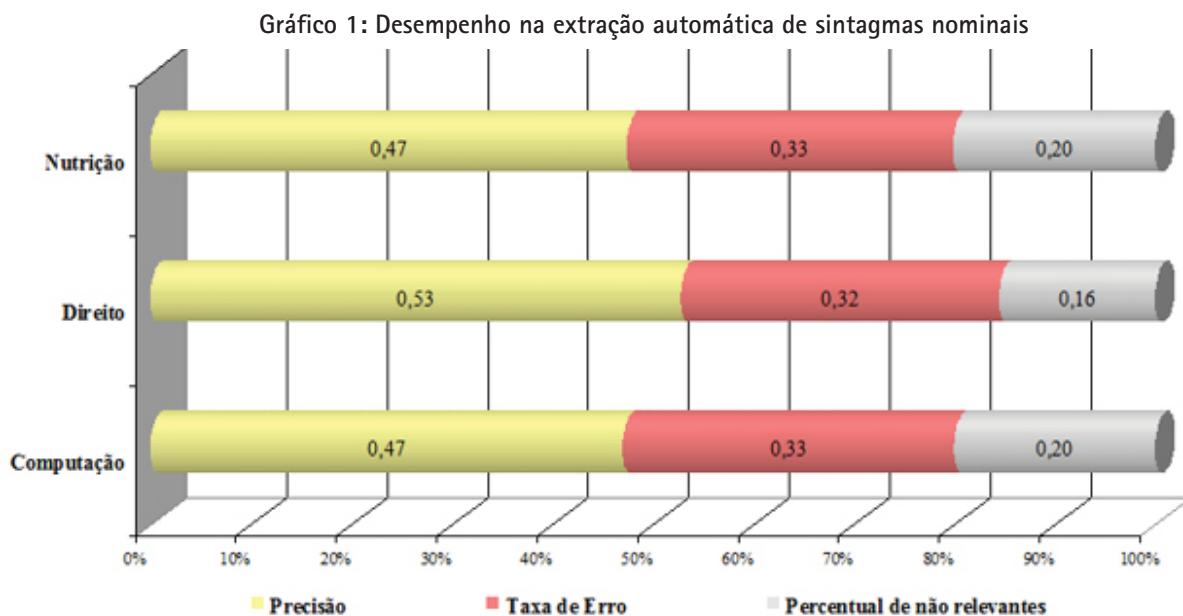
Fonte: os autores.

extraídos, quantos destes não constituem sintagmas nominais; quantos destes se constituem em sintagmas, mas não são relevantes como descritores para o respectivo resumo, e; quanto destes se constituem em sintagmas nominais relevantes como descritores para o respectivo resumo. Estes dados foram organizados em tabelas para cada programa de pós-graduação, como a mostrada na Tabela 1 para o programa de pós-graduação em direito.

A partir das três tabelas construídas (uma para cada programa de pós-graduação estudado) definiram-se os valores totais para cada uma das variáveis observadas, o que permitiu avaliar a extração de sintagmas nominais através dos percentuais de precisão em extrair sintagmas nominais relevantes; da taxa de erro ao extrair cadeias de caracteres que não constituem

sintagmas nominais, e; o percentual de sintagmas nominais extraídos não relevantes. Estes percentuais estão organizados na Tabela 2.

Na Tabela 2, pode-se observar que a taxa de erro do OGMA na identificação de sintagmas nominais foi em média de 42%, indicando que 42% dos possíveis sintagmas nominais extraídos pelo OGMA não são de fato sintagmas nominais. Os demais 58% dos possíveis sintagmas nominais extraídos são de fato sintagmas nominais. Entretanto, em média, a precisão na extração de sintagmas nominais relevantes dos resumos foi de 44%, significando que menos da metade dos possíveis sintagmas nominais extraídos pelo OGMA são de fato sintagmas nominais relevantes como descritores para os resumos, sendo 14% sintagmas nominais irrelevantes como descritores.



Fonte: os autores.

A partir da Tabela 2, foi construído do Gráfico 1, onde se pode observar com mais clareza que o processo de extração de sintagmas nominais através do OGMA teve diferentes desempenhos para cada programa de pós-graduação, sendo obtido melhor desempenho (melhor índice de precisão) para resumos de Direito, depois

Computação e Nutrição. Esta diferença de desempenho pode em parte ser explicada pela diferente natureza dos termos técnicos presentes nos resumos destes três programas: enquanto a terminologia da área do Direito faz uso de vocábulos mais comuns (provavelmente mais presentes no léxico do OGMA), em Computação há o

Quadro 2 – Sintagmas nominais relevantes extraídos do resumo D6 – do programa de pós-graduação de Direito da UFPE

A tutela dos direitos da personalidade por meio da aplicabilidade direta do princípio da dignidade da pessoa humana nas relações de direito privado.

Mestre em Direito.

O presente trabalho faz um estudo do sistema de proteção dos direitos da personalidade no ordenamento jurídico brasileiro com vistas a demonstrar a possibilidade de aplicação direta do princípio constitucional da dignidade da pessoa humana nas relações de direito privado. Esta possibilidade de apresenta a partir da constatação da mudança paradigmática da codificação experimentada com a promulgação do Código Civil de 2002 que adotou como uma das suas orientações metodológicas a técnica das cláusulas gerais. Por meio da efetividade das chamadas cláusulas gerais, o sistema normativo de direito privado deixa de ser um sistema fechado e passa a estabelecer conectividade com todo o ordenamento. A personalidade humana não se restringe a um requisito formal do sujeito de direito. Ela é mais do que isso, é um valor juridicamente tutelado. A tutela da personalidade confere unidade ao ordenamento jurídico, e torna evidente a superação da tradicional dicotomia entre o direito público e o direito privado. A aplicação direta do princípio da dignidade da pessoa humana nas relações de direito privado reflete e supera essa dicotomia e estabelece um verdadeiro sistema de proteção da pessoa humana no ordenamento jurídico brasileiro.

Cláusulas gerais. Dignidade da pessoa humana. Personalidade. DIREITO.

Pós-Graduação em Direito.

Ciências Humanas, Letras e Artes.

Fonte: os autores.

uso frequente de estrangeirismos e siglas, e em Nutrição há uso frequente de siglas, unidades de medidas e nomes científicos de animais, plantas e substâncias.

Entende-se que a precisão na extração de sintagmas nominais relevantes em torno de 50%, gera uma sobrecarga sobre profissional da informação atuando como indexador, que teria que avaliar os possíveis sintagmas nominais extraídos e praticamente descartar metade dos mesmos.

Entretanto, avaliando a qualidade dos sintagmas nominais relevantes extraídos como descritores dos resumos das teses e dissertações, percebe-se que este trabalho de indexação semi-automática, mesmo que enfadonho, é recompensador e que a aplicação dos sintagmas nominais na indexação de teses e dissertações é eficaz. Sustentando esta afirmativa, apresenta-se a taxa de revocação das palavras-chaves das teses e dissertações nesta lista de sintagmas nominais relevantes extraídos que é de 0,88 para Direito, 0,84 para Computação e 0,83 para Nutrição. Em outras palavras, cerca de 80% das palavras-chaves atribuídas às teses e dissertações são extraídas (recuperadas) como sintagmas nominais. Além disso, apresentam-se grifados no Quadro 2, para fins de exemplificação, os sintagmas nominais relevantes extraídos do resumo D6 do programa de pós-graduação de Direito, neste resumo a extração de sintagmas nominais obteve um dos melhores desempenhos. Observa-se, para o Quadro 2, que foi usado um código de cores para realçar quais sintagmas nominais estão relacionados a qual palavra-chave.

Analizando o Quadro 2, percebe-se que todas as palavras-chaves foram detectadas como sintagmas nominais (antepenúltima linha do quadro), e que os sintagmas nominais são mais específicos que as palavras-chaves que neles estão contidos. Este é o caso, por exemplo, da palavra-chave Personalidade, que fica menos ambígua e contextualizada através dos sintagmas: a personalidade da pessoa humana; os direitos da personalidade; a personalidade no ordenamento jurídico brasileiro; um estudo do sistema de

proteção dos direitos da personalidade no ordenamento jurídico brasileiro. O conceito ordenamento jurídico, que também se constitui como um importante descritor, também foi detectado como parte de sintagmas nominais, e complementaria bem a lista de palavras-chaves.

- alguns dos sintagmas nominais não representam relevância para o usuário no momento da busca, ou seja, embora sejam sintagmas, não constituem descritores e não corresponderiam à necessidade de informação do usuário no momento da busca por aquele documento. Tal fato mostra que a extração de sintagmas deve ser acompanhada de estratégias de ordenação por relevância dos sintagmas, levando em conta critérios de frequência e posicionamento, semelhantemente às propostas existentes para palavras isoladas. O OGMA gera uma pontuação de cada sintagma nominal que indicaria os mais relevantes como descritores, entretanto nem sempre a lista de sintagmas extraídos ordenada em ordem decrescente de pontuação aponta no topo os sintagmas mais relevantes;
- nem todos os sintagmas extraídos pelo OGMA são de fato sintagmas nominais, os motivos para a ferramenta retornar tais falsos positivos pode ser atribuída a um problema na etiquetagem do texto e uma possível falha nas regras de extração estabelecidas pelo OGMA;
- nem todos os sintagmas nominais são extraídos pelo OGMA como, por exemplo, os nomes próprios. Isto se deve a uma falha nas regras de extração estabelecidas pelo OGMA, que não contempla a possibilidade de dois ou mais substantivos adjacentes formarem um sintagma nominal;
- os sintagmas nominais extraídos classificados como relevantes constituem bons descritores para os resumos, constituindo em bons pontos de acesso as teses e dissertações no processo de recuperação de informação, de semântica muito mais precisa que

o conjunto de palavras isoladas que os compõem.

Baseando-se nestas observações, conclui-se que os processos de extração de sintagmas nominais e ordenação por relevância precisam ser mais profundamente investigados visando melhora na extração de sintagmas nominais relevantes como descritores.

## Conclusão

Um dos principais objetivos dos profissionais e cientistas da informação é garantir o acesso rápido e preciso à informação solicitada pelo usuário. A busca por novos métodos, novos caminhos para satisfazer essa necessidade são implementados e testados com o intuito de chegar mais próximo do ideal, ou seja, retornar o maior número de documentos relevantes à necessidade de informação do usuário.

A aplicação de métodos automatizados de extração e indexação por sintagmas nominais, embora se reconheçam as limitações nas ferramentas disponíveis, mostra-se bastante promissora. Os sintagmas nominais se apresentam como uma alternativa ao uso de palavras isoladas nos sistemas de recuperação de informação, pois se configuram como melhores descritores e pontos de acesso aos documentos, eliminando os problemas causados pela sinônima e a polissemia das palavras isoladas.

A literatura é carente de ferramentas para extração de sintagmas nominais em textos escritos em português, sendo o OGMA a única ferramenta disponível que se teve conhecimento durante a execução do estudo.

É possível perceber que os métodos para extração de sintagmas nominais implementados no OGMA precisam ser aperfeiçoados a fim de diminuir a taxa de falsos sintagmas nominais extraídos. Além disso, a extração dos sintagmas nominais não garante por si só a seleção de bons descritores, sendo necessário que a ferramenta de extração de sintagmas possa fazer a análise

dos textos e pontuar os sintagmas de acordo com a pontencialidade de serem bons descritores, este método de ordenação também precisa ser aprimorado no OGMA.

Assim, a extração automática de sintagmas nominais e utilização dos mesmos na indexação e recuperação de informação se constitui um amplo espaço de pesquisa e desenvolvimento de novos trabalhos. Para trabalhos futuros na temática, pretende-se avaliar a revocação na extração de sintagmas nominais; investigar com maior profundidade o processo de etiquetagem dos termos, extração e pontuação de sintagmas nominais com a finalidade de propor novos métodos; bem como investigar e aplicar outros métodos e ferramentas da área de processamento de linguagem natural nestes processos.

## Referências

- BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern information retrieval**. New York: ACM Press, 1999.
- KURAMOTO, H. Sintagmas nominais: uma nova proposta para a recuperação de informação. **DataGramZero: revista de Ciência da Informação**, v. 3, n. 1, 2002.
- \_\_\_\_\_. Uma abordagem alternativa para o tratamento e a recuperação de informação textual: os sintagmas nominais. **Ciência da Informação**, Brasília, v. 25, n. 2, 1995.
- LE GUERN, M. Un analyseur morpho-syntactique pour l'indexation automatique. **Le Français Moderne**, v. 59, n. 1, p. 22-35, juin 1991.
- MAIA, L. C. G. **Uso de sintagmas nominais na classificação automática de documentos eletrônicos**. 2008. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais – UFMG. Belo Horizonte, 2008.
- MIORELLI, S. T. **Extração do sintagma nominal em sentenças em português**. 2001. 98 f. Dissertação (Mestrado em Ciência da Computação) – Faculdade de Informática, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.
- PERINI, M. A. **Gramática descritiva do Português**. 3 ed. São Paulo: Ática, 1998.

SOUZA, R. R.; ALVARENGA NETO, R. C. D. de;  
MENDES, K. C. I. Mapeamento semântico através da análise de ocorrência de descritores sobre gestão do conhecimento. *Transinformação*, v. 19, n. 1, p. 19-30, 2007.

SOUZA, R. R. Uma proposta de metodologia para indexação automática utilizando sintagmas nominais. **Encontros Bibli:** revista eletrônica de Biblioteconomia e Ciência da Informação, v. 11, n. esp., p. 42-59, 2006.

TUFANO, D. **Estudos de língua e literatura.** 4. ed. São Paulo, Moderna, 1990.

#### **Indexing and information retrieval of theses and dissertations through noun phrases**

##### **Abstract**

*Introduction: Discusses the use of noun phrases in the automatic indexing process of theses and dissertations deposited in the UFPE Digital Library of Theses and Dissertations (BDTD-UFPE), on the assumption that noun phrases consist of a better knowledge unit for indexing and information retrieval than individual words, allowing an adequate response to the users information need when searching for information. It presents the state of the art of noun phrases and their automatic extraction process, as well as its applicability in automatic indexing and information retrieval. Method: Based on text analysis tool (OGMA), analyses the applicability of the extraction of noun phrases in automatic indexing and information retrieval of thesis and dissertations in the context of BDTD-UFPE. Applied to abstracts from Law, Computer and Nutrition thesis and dissertations, the variables could be observed, allowing the research team assess the extraction of noun phrases using: the percentage of accuracy of relevant noun phrases; the error rate extract strings that are not noun phrases, and; the percentage of non relevant noun phrases extracted. Results: The process of extracting noun phrases by OGMA showed different performances for each graduate program, with better performance (better accuracy rate) for abstracts from Law Thesis and Dissertations, followed by Computer and Nutrition ones. This performance difference can be partly explained by the different nature of technical terms presented in the abstracts. Conclusions: It concludes that although there are limitations in the available tools, the application of automated methods of extraction and indexing by noun phrases appears to be quite promising, since the noun phrases are configured as best descriptors and access to documents, eliminating the problems caused by synonymy and polysemy of isolated words.*

##### **Keywords**

*Noun phrase; Information retrieval; Automatic indexing; Theses and dissertations.*

#### **Agradecimentos**

*Os autores gostariam de agradecer a Fundação de Amparo à Pesquisa do Estado de Pernambuco (FACEPE) pelo fomento dado ao projeto "Mapeador de teses e dissertações da UFPE (MTD-UFPE)" cujos equipamentos foram utilizados nesta pesquisa e ao laboratório Liber da UFPE pela estrutura física disponibilizada para realização deste trabalho.*

Recebido em 14 de fevereiro de 2011  
Aceito em 15 de fevereiro de 2011

<sup>1</sup> Bacharel em Gestão da Informação (UFPE), Doutor em Ciência da Computação (UFPE). Professor adjunto - Universidade Federal de Pernambuco (UFPE)/Departamento de Ciência da Informação.  
renato.correa@ufpe.br

<sup>2</sup> Bacharel em Gestão da Informação (UFPE). Estudante de graduação - Universidade Federal de Pernambuco (UFPE)/ Departamento de Ciência da Informação.  
darlygoes@gmail.com

<sup>3</sup> Bacharel em Gestão da Informação (UFPE). Estudante de graduação - Universidade Federal de Pernambuco (UFPE)/ Departamento de Ciência da Informação.  
camila.oalima@gmail.com

<sup>4</sup> Bacharel em Gestão da Informação (UFPE). Estudante de graduação - Universidade Federal de Pernambuco (UFPE)/ Departamento de Ciência da Informação.  
tiago.stallin@gmail.com

#### **Como citar este artigo:**

CORRÊA, R. F.; MIRANDA, D. G. de; LIMA, C. O. de A.; SILVA, T. J. da. Indexação e recuperação de teses e dissertações por meio de sintagmas nominais. **AtoZ**, Curitiba, v. 1, n. 1, p. 11-22, jan./jun. 2011. Disponível em: <<http://www.atoz.ufpr.br>>. Acesso em:

# Produção tecnológica da Universidade Estadual de Londrina (UEL): mapeamento da área de Ciências Agrárias pela Plataforma Lattes

Natan Tiago Batista Serzedello<sup>1</sup>, Maria Inês Tomaél<sup>2</sup>

## Resumo

**Introdução:** Os produtos e os processos tecnológicos empregados com o fim de solucionar os problemas práticos da sociedade compreendem a produção tecnológica. Esta se constitui como o meio essencial de mensuração do desenvolvimento tecnológico do ambiente em que a instituição está inserida. O objetivo desta pesquisa foi averiguar a produção tecnológica da área de Ciências Agrárias na Universidade Estadual de Londrina (UEL). **Método:** Identificação e análise da produção tecnológica, por meio dos currículos, disponíveis na Plataforma Lattes, de 51 professores que receberam bolsas do Programa de Iniciação Científica (PROIC) da UEL. Os dados coletados para a análise foram tabulados e categorizados por tipologias de produções tecnológicas. **Resultados:** Dos 65 itens cadastrados nos currículos Lattes dos professores referem-se à produção tecnológica, a qual foi mapeada, descrevendo-se as tipologias identificadas. Observou-se que as especialidades mais inovadoras são pertinentes a Agronomia; Recursos Florestais e Engenharia Florestal; Engenharia Agrícola; Zootecnia; Medicina Veterinária; Recursos Pesqueiros e Engenharia de Pesca; Ciência e Tecnologia de Alimentos. **Conclusões:** Verificou-se que a produção tecnológica encontrada, decorrente de atividades de pesquisas de desenvolvimento tecnológico, foi disseminada por meio da produção bibliográfica, principalmente por artigos de periódicos. Os pesquisadores participantes deste estudo fazem parcerias em um mesmo Departamento, ou em Departamentos variados, dentro e fora da Universidade, interagindo e compartilhando as experiências, o conhecimento e os equipamentos, o que potencializa uma produção tecnológica mais bem formulada e com maiores chances de êxito. Um percentual considerável de artigos foi publicado em periódicos adeptos da filosofia da *Open Archives Initiative* (OAI), o que possibilita o acesso livre e favorece a disseminação dos resultados das pesquisas.

## Palavras-chave

Produção tecnológica; Desenvolvimento tecnológico; Inovação tecnológica.

## Introdução

A produção tecnológica é caracterizada, por uma comunidade científica, pela geração de produtos e de processos tecnológicos, com o intuito de contribuir na solução de problemas práticos. A produção tecnológica geralmente tem a finalidade de atender as necessidades da sociedade, por meio da criação de invenções e, por consequência, impacta o desenvolvimento tecnológico, econômico e social.

O espaço tecnológico, ou ambiente de produção tecnológica, parte do conjunto da ciência, da tecnologia e da informação. Esse ambiente é assim denominado porque emprega a tecnologia

para a geração da produção tecnológica. Como exemplos dessa produção, podem-se citar as patentes, os manuais, as metodologias, os processos tecnológicos e de produção e os relatórios de pesquisas.

Os ambientes de produção tecnológica são configurados no Brasil, principalmente, pelas Universidades, parques tecnológicos, institutos tecnológicos e de pesquisa, os quais possuem a infra-estrutura espacial para desenvolver pesquisa aplicada.

As pesquisas aplicadas desenvolvidas nesses ambientes, normalmente, são fundamentadas pelo conhecimento científico, resultante de pesquisa

básica, que tem origem nas Universidades e estas resultam em produtos e processos denominados produção tecnológica.

O desenvolvimento social pode ser verificado por meio de indicadores estatísticos de produção tecnológica, que expõem o potencial qualitativo e quantitativo do progresso das Universidades, institutos tecnológicos e de pesquisa. Esse progresso está associado à criação de invenções, inovações de sucesso e quantidade de patentes. Quando essas instituições de pesquisa não apresentam alterações inovativas bem sucedidas (produções tecnológicas baseadas em inovações já existentes) são consideradas improdutivas e isso compromete a capacidade de obtenção de recursos.

Pela literatura, verifica-se que o indicador de produção tecnológica mais empregado é o número de patentes. Segundo Bernardes (2003, p. 152), “as patentes formam até hoje o principal indicador de produção tecnológica nos países centrais e o número delas é uma medida que auxilia a avaliação da capacidade de inovação”. Desde o final da década de 1990, vem-se aplicando e melhorando a eficiência e eficácia desse indicador focando somente nas patentes.

Os dados disponíveis sobre produção tecnológica são compilados por meio da elaboração de estatísticas dos produtos e processos tecnológicos, e estes auxiliam no desenvolvimento de pesquisas e no aprimoramento de artefatos tecnológicos, que refletem no setor produtivo e na qualidade de vida das pessoas.

Depois de tornarem-se públicas, as estatísticas de produção tecnológica servem de incentivo ao progresso, ao reconhecimento ou são utilizadas como base para novos estudos, formando o processo de inovação.

Além das patentes, outras formas de produção tecnológica também são de relevância para o desenvolvimento econômico, social, cultural, político e organizacional, como: manuais, metodologias, processos desenvolvidos e a participação em feiras e exposições. Essas formas

de produção foram foco desta pesquisa na área de Ciências Agrárias da Universidade Estadual de Londrina (UEL). As diferentes formas de produção tecnológica agregam conhecimentos para a área em que são produzidas; contudo, apenas as patentes aparecem na literatura como indicadores de produção tecnológica, para averiguação da evolução tecnológica.

O tema produção tecnológica é pouco explorado pela literatura. A identificação de artefatos tecnológicos na UEL, na área de Ciências Agrárias, pode garantir o acesso a informações com potencial de oportunidades mercadológicas, como também a visão do montante da variedade de invenções e a abertura a novos conhecimentos, por meio das inovações produzidas na Universidade.

A pesquisa foi realizada na área de Ciências Agrárias com o objetivo de prospectar a produção tecnológica da Universidade Estadual de Londrina (UEL). Os objetivos específicos da pesquisa foram: mapear as formas de produção tecnológica; descrever as tipologias de produção tecnológica; e identificar os canais de comunicação mais utilizados.

## Desenvolvimento tecnológico

Quando a sociedade sente a falta de algo necessário para o seu auxílio, progresso e melhoria surgem as invenções, que podem gerar produtos e processos com o propósito de sanarem as necessidades do meio em que estão inseridos. Schumpeter (1996, p. 11) define invenção como “processo de criação de novos produtos ou métodos técnicos direcionados para o meio econômico moldado por lógicas específicas.” Ou seja, a necessidade de satisfação serve de incentivo para o surgimento das invenções.

Pensando em maior conforto, comodidade e melhores condições de vida, os integrantes da sociedade sempre geram produtos, processos e instrumentos técnicos, vinculados em sua maioria à tecnologia.

Os artefatos de base tecnológica são denominados produtos ou processos tecnológicos, isto por possuírem informações tecnológicas embutidas, que geralmente são elaboradas para sanar os problemas práticos da sociedade.

## Tecnologia

Seguindo uma tendência evolucionista, vários autores mencionam que a tecnologia teve seu início quando o homem começou a usar as mãos para capturar seu alimento e a construir itens para melhorias no seu cotidiano, tais como as técnicas e as ferramentas por eles utilizadas.

A evolução do homem, até alcançar sua forma atual, permitiu seu progresso gradual, causando uma evolução também nas técnicas, à medida que o cérebro evoluía. Com o uso do cérebro, sobreveio a organização social e o homem se distanciou do mundo animal.

De acordo com Goldemberg (1978, p. 157), tecnologia é “o conjunto de conhecimentos de que uma sociedade dispõe sobre ciências e artes industriais, incluindo os fenômenos sociais e físicos, e a aplicação destes princípios à produção de bens e serviços”. A tecnologia agrupa valor também aos fatores técnicos, políticos, econômicos e sociais.

A concepção convencional da tecnologia é a de geração de produtos e processos de ciência aplicada, ou seja, resulta em itens tecnológicos. Nada mais é que a junção da teoria e prática, o saber e o fazer. A seleção das tecnologias é compreendida pela busca em conjuntos de tecnologias ou a procura por produtos e processos com tecnologias, que são denominados produção tecnológica.

## Produção tecnológica

O conceito de produção tecnológica foi inicialmente disseminado por sociólogos que criaram teorias que mencionavam produtos e processos tecnológicos gerados por inspirações,

sem embasamento ou acúmulo de conhecimento. A literatura (OGBUM, THOMAS, 1922; GILFILAN, 1935 *apud* CONCEIÇÃO, 2003), ressalta que sociólogos da Universidade de Chicago acreditavam que as invenções eram quase um milagre, nascidas por meio da intuição.

É necessário ter uma noção preliminar, ou uma base inicial – adquirida a partir de um conhecimento já existente – para se criar uma invenção. Marx confirma esta assertiva em suas teorias, quando menciona a necessidade de conhecimento prévio e de acúmulo de conhecimento para desenvolver os produtos e processos tecnológicos. Os sociólogos (OGBUM, THOMAS, 1922; GILFILAN, 1935 *apud* CONCEIÇÃO, 2003), fundamentados pela ótica marxista, ressaltam o fato de que os países menos desprovidos de desenvolvimento necessitam seguir os passos dos países industrializados para atingirem a sua evolução (MACHADO, 2006).

A identificação da produção e o crescimento dos produtos e processos tecnológicos nos países ocidentais surgiram no século XVIII, e com isso, também a demanda de produção tecnológica. O conjunto de produtos e processos tecnológicos compõe a produção tecnológica, a qual visa à

Quadro 1 – Tipos de produção tecnológica

Tipos de produção tecnológica	
Modalidades	Especificação
Softwares	Abrange os sistemas computacionais.
Produtos	Projetos, protótipos, pilotos entre outros.
Processos	Processo ou técnica com fins tecnológicos aplicáveis.
Cartas mapas ou similares	Produtos cartográficos desenvolvidos para tecnológicos aplicáveis.
Relatório de pesquisa	Relatórios periódicos ou finais que descrevem as pesquisas que contenham fins tecnológicos.
Patente	Fonte de informação tecnológica. Criações, invenções no âmbito da tecnologia.
Manuais	Manuais que documentem e descrevam o produto.

Fonte: Baseado na Plataforma Lattes<sup>1</sup> - CNPQ.

<sup>1</sup> <http://plsql1.cnpq.br/images/ajuda/index.html>.

solução dos problemas e o desenvolvimento tecnológico da sociedade.

A vasta produção de invenções e inovações constantes, de acordo com Schumpeter (1996), é classificada como idéia original que traz reflexos de ordem econômica, social e cultural. Há uma grande variabilidade das formas de produção tecnológica e as suas diversas classificações, como as exemplificadas no Quadro 1.

Essas tipologias (Quadro 1) são decorrentes de pesquisas de desenvolvimento tecnológico e representam ou descrevem inovações. Os tipos de produção apresentados são consolidados, visto que figuram na Plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## Inovação

A inovação acontece quando o conhecimento gerado pela pesquisa básica se torna aplicado, possibilitando a criação de produtos e processos, seu incremento ou aperfeiçoamento, ou seja, quando esta agrupa benefícios a uma tecnologia existente, gerando uma nova. Joseph Schumpeter foi o primeiro a relacionar a inovação ao desenvolvimento econômico, pois acreditava no processo de criação fundamentado no conhecimento acumulado, ligado ao processo de desenvolvimento econômico.

Para Lemos (1999, p.126) inovação é um “processo pelo qual produtores dominam e implementam o projeto e produção de bens e serviços que são novos para os mesmos, a despeito de serem ou não novos para seus concorrentes – domésticos ou estrangeiros”. Esse conceito é complementado por Caldas e colaboradores (2001, p. 2) quando dizem que o conceito de inovação “ainda não está completamente internalizado nas nossas instituições, e serão necessários muitos esforços até que se tenha no Brasil um sistema nacional de inovação forte e consolidado”.

## Inovação tecnológica nas universidades

Por meio da pesquisa de caráter inovador, ocorre o desenvolvimento que abre oportunidades a serem apropriadas pela sociedade.

No Brasil, por meio da pesquisa nas Universidades, surgem produtos e processos tecnológicos inovadores, que podem contribuir para sanar as necessidades da sociedade. De acordo com Fava-de-Moraes (2000, p. 9), “a pesquisa básica executada ‘espontaneamente’ pela Universidade ainda é comprovadamente a maior fonte de resultados aplicáveis do que a pesquisa dita ‘encomendada’ por empresa.” E corrobora com a pesquisa aplicada e com processos de inovação tecnológica nas Universidades, o que possibilita a formação de alunos empreendedores, que são capacitados a inovar fora do meio acadêmico.

De Acordo com Fava-de-Moraes (2000, p. 1):

O papel da universidade é inquestionável para a formação de pessoas qualificadas e para o desenvolvimento socioeconômico de um país, principalmente quando fundamentado em uma política científica-tecnológica articulada no espaço das relações entre universidade, Estado, empresariado e outros setores sociais.

O processo de produção do conhecimento na inovação tecnológica é formado pelo ciclo de idéias no âmbito da ciência, por meio de pesquisa, criatividade e experimentação, obtendo um produto ou processo de fator aplicável. Este processo de desenvolvimento científico e tecnológico aprimora o bem-estar da sociedade e a tecnologia acarreta a evolução da população, devendo estes ser conduzidos em conjunto.

O desenvolvimento de inovação tecnológica em Universidades, no Brasil, ainda é muito baixo e necessita expansão para atender as necessidades da sociedade com tecnologias adequadas, produtos, processos e serviços que promovam o crescimento econômico e social.

Para isso é necessária a interação da comunidade científica acadêmica com o setor privado, visto

que muitas vezes a inovação tecnológica inicia-se nas Universidades, mas não se concretiza.

O que acontece nas Universidades é a dominação do formal, a forma de como se executa, não havendo o questionamento sobre as finalidades dos produtos ou processos elaborados. Enfim, não se disseminam informações que poderiam, muitas vezes, ser úteis à sociedade ou também servir de apoio a outras invenções. Caldas e colaboradores (2001, p. 2) consideram que:

O momento é o mais propício para que o País perceba que o único caminho a ser trilhado por países em desenvolvimento em busca de suas chances de competir e sobreviver no mundo do futuro deve apropriar-se dos resultados da ciência e utilizá-los em benefício da sociedade, ou seja, promover e estabelecer o processo de inovação.

Para tanto, o ato de explorar toda forma de inovação tecnológica nas comunidades científico-acadêmicas é fundamental para que aconteça o desenvolvimento econômico, social e político da sociedade, gerando bem-estar e avanço tecnológico. Fava-de-Moraes (2000, p. 11) afirma que:

O Brasil não deve se iludir com importantes conquistas eventualmente exploradas com personalismo e ufania na mídia. O Brasil precisa conhecer o quadro real para ousar com humildade e exigir que decisões de política científico-tecnológica sejam tomadas conjuntamente pelas Universidades-Governo-Setor privado (triângulo de Sábatto ou tríplice hélice).

É muito importante disseminar resultados de estudos científicos, mesmo daquelas que não obtiveram sucesso, para que ocorram mudanças futuras. Esta ação possibilitará que as informações resultantes de pesquisas possam fazer a diferença. Embora seus efeitos não sejam perceptíveis no curto prazo, tais informações poderão ser utilizadas em longo prazo, se estiverem disponíveis.

## Procedimentos metodológicos

A pesquisa foi realizada na área de Ciências Agrárias da UEL, por apresentar um grande

índice de professores bolsistas de produtividade em pesquisa do CNPq – comparada a outros setores da UEL.

Situada na cidade de Londrina (Paraná), a UEL é a maior universidade do Estado do Paraná e apresenta índices anuais de produtividade crescente. O Centro de Ciências Agrárias (CCA) da UEL possui os cursos de graduação em: Agronomia, Zootecnia e Medicina Veterinária. Contempla também os programas de pós-graduação em: Agronomia (Mestrado e Doutorado), Ciência Animal (Mestrado e Doutorado), Ciência de Alimentos (Mestrado e Doutorado) e diversos cursos de especialização (*lato sensu*).

O CCA possui 118 professores. Para a definição da amostra, utilizou-se o edital que discrimina os nomes dos professores que receberam bolsas, na UEL, do Programa de Iniciação Científica (PROIC) que distribui bolsas de Iniciação Científica da própria UEL, da Fundação Araucária e do Programa Institucional de Bolsas de Iniciação Científica (PIBIC) do CNPq. Este critério foi empregado por eleger para o recebimento de bolsas os professores/pesquisadores mais produtivos. Cinquenta e um professores orientadores listados no Edital PROIC na área de Ciências Agrárias da UEL foram os que fizeram parte da nossa pesquisa.

Para a coleta de dados foram desenvolvidos os seguintes procedimentos: a) Identificação do currículo dos professores na Plataforma Lattes; b) Levantamento e análise da produção na seção de Produção em CT&A do currículo Lattes. Foi feito um exame das produções bibliográficas, produções técnicas e demais trabalhos.

Para a produção bibliográfica, considerou-se:

- Título de cada produção;
- Resumo e leitura técnica do artigo em questão (sempre que disponível na Web).

Para a produção técnica, considerou-se:

- Título de cada produção;
- Resumo (quando este estava disponível na Web).

Para a categoria de “demais trabalhos” cadastrados no currículo lattes, analisou-se:

- Título de cada produção;
- Resumo (quando este estava disponível na Web).

O instrumento utilizado para a coleta foi um formulário, que possibilitou levantar as informações de identificação dos docentes e dados para o levantamento de sua produção tecnológica. Para cada produção tecnológica encontrada, preencheu-se um formulário.

## Produção tecnológica: resultados da pesquisa

Dentre os 51 currículos de professores pesquisados, 25 não possuem produção tecnológica, porém apresentam produção bibliográfica. Isto corrobora a afirmação de Fujino, Stal e Plonski (1999) de que há a necessidade de aumento progressivo de conhecimento científico para a obtenção do progresso técnico.

Para os outros 26 participantes desta pesquisa, foi identificada produção tecnológica disseminada por meio da literatura científica – cadastrada no currículo lattes no item Produção Bibliográfica, na Produção Técnica, ou nos Demais Trabalhos.

A produção dos professores do CCA da UEL não possui indicador de mensuração relevante que a evidencie, mesmo que esta traga benefícios, modificações, e melhorias à sociedade. Isto se verifica em virtude de nenhum dos currículos analisados incluírem dados sobre a solicitação ou concessão de patentes. Sabe-se que a patente é o principal indicador de produção tecnológica, conforme destaca Bernardes (2003, p. 152), isso confirma que as produções tecnológicas encontradas neste estudo não foram

mensuradas e incluídas como parâmetro para o desenvolvimento econômico e tecnológico.

## Tipologia e formas de produção tecnológica

A produção tecnológica encontrada inclui: softwares, produtos, processos e manuais, os quais são classificados como idéias originais, e podem ter reflexos de ordem cultural, social e econômica conforme evidenciado por Schumpeter (1996). O Quadro 2 sintetiza os dados encontrados na pesquisa.

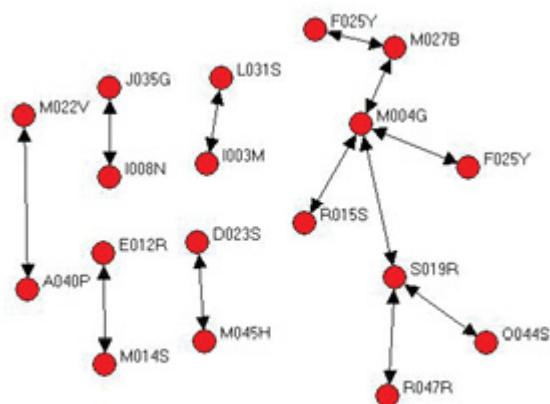
**Quadro 2 – Especialidade e produção**

Especialidade	Tipo de Produção Tecnológica
Entomologia Agrícola, Defesa Fitossanitária	03 processos
Microbiologia Agrícola	05 produtos
Manejo e conservação do solo	03 processos
Manejo e Tratos Culturais	02 produtos
Fitopatologia	06 softwares 04 processos
Tecnologia de Aplicação	02 produtos
Instrumentação em máquinas e Implementos Agrícolas	02 produtos
Agricultura; Produção e Beneficiamento de Sementes	01 produto
Controle populacional de Animais	01 processo
Engenharia de Alimentos	01 processo 01 software
Avaliação e controle da qualidade de alimentos; Química, física, Físico-química e Bioquímica dos Alimentos e das Matérias-primas	01 manual 01 processo 02 produtos
Embalagens de produtos alimentares	02 produtos
Clinica Veterinária; Obstetrícia Animal	01 processo
Doenças parasitárias de animais	06 produtos
Zoonoses e Saúde Pública	03 manuais 02 produtos 01 processo
Medicina Veterinária Preventiva	02 manuais
Valor nutritivo de Alimentos	03 processos
Tecnologia de produtos de origem vegetal	04 processos 02 produtos
Produção animal em ruminantes	01 produto
Conservação de Alimentos para animais	01 software
Química, física, Físico-química e Bioquímica dos Alimentos e das Matérias-primas; Tecnologia de alimentos dietéticos e Nutricionais	01 software 01 produto

Fonte: os autores.

Observa-se que as especialidades mais inovadoras são pertinentes a: Agronomia; Recursos Florestais e Engenharia Florestal; Engenharia Agrícola; Zootecnia; Medicina Veterinária; Recursos Pesqueiros e Engenharia de Pesca; Ciência e Tecnologia de Alimentos.

Figura 1 – Mapeamento da produção tecnológica na área de Ciências Agrárias da UEL



Fonte: os autores.

Na Figura 1, cada círculo contém um código que corresponde a um pesquisador. Pode-se observar que cinco das produções tecnológicas encontradas em coautoria foram produzidas por pelo menos dois atores<sup>2</sup> participantes desta pesquisa. Os atores com maior índice de trabalhos em coautoria são os que receberam os códigos M004G e S019R. O M004G possui quatro ligações de coautoria, e o S019R sustenta três ligações (coautoria), ambos com diferentes professores.

Dos pesquisadores que tiveram os currículos analisados, os pertencentes ao Departamento de Agronomia foram os que mais produziram, com 43% do total dos produtos e com 52% do total de processos encontrados. Esses produtos e processos pertencentes ao Departamento de Agronomia são focados nas temáticas: valores nutritivos de alimentos, engenharia de alimentos, entomologia agrícola, tecnologia de aplicação de máquinas e implementos agrícola, manejo e conservação do solo, e defesa fitossanitária.

Os pesquisadores desse departamento também se destacaram com 67% do total da produção de softwares desenvolvidos. Estes softwares podem ser classificados nas especialidades de epidemias agrícolas, riscos climáticos e fitopatologia.

O Departamento de Agronomia, devido ao alto número de professores contemplados com bolsa de iniciação científica, apresentou o maior número de produções tecnológicas, e também abrangeu temáticas variadas e diversificados tipos de produções.

Do Departamento de Medicina Veterinária Preventiva, os pesquisadores detiveram 29% do total de produtos identificados. Os produtos oriundos desse departamento têm o foco nas temáticas: zoonoses e saúde pública; medicina veterinária preventiva; e doenças parasitárias.

Em relação à produção total de manuais, os pesquisadores do Departamento de Medicina Veterinária Preventiva atingiram 83% da produção, deste total, seis manuais são na temática de Zoonoses e Saúde Pública.

Observa-se que os pesquisadores do Departamento de Ciência e Tecnologia de Alimentos foram os únicos que produziram quatro diferentes tipos de produção tecnológica. Porém, em relação à quantidade, os pesquisadores desse departamento estão com baixa produtividade tecnológica (quando comparados a pesquisadores do Departamento de Agronomia), somando oito produções, o total de produções tecnológicas identificadas nesta pesquisa foi de 65.

Os pesquisadores dos Departamentos de Zootecnia e Ciência e Tecnologia de Alimentos abordaram as temáticas: valor nutritivo de alimento; tecnologia de produtos de origem vegetal; química; física; físico-química e bioquímica de alimentos e das matérias-primas alimentares; avaliação e controle de qualidade de alimentos; e embalagens de produtos alimentares.

<sup>2</sup> Professores/pesquisadores.

As produções tecnológicas encontradas tiveram como principal temática a avaliação e controle de qualidade de alimentos e embalagens de produtos alimentares, formando parcerias, por meio de coautoria, entre os Departamentos de Agronomia, Zootecnia, Ciência e Tecnologia de Alimentos.

A titulação dos professores não influenciou na importância ou na forma da produção tecnológica, não havendo diferença entre a produção do livre docente, dos pós-doutores e dos doutores. Um fator que pode ter contribuído para isso são as coautorias, pois com a produção em conjunto há o compartilhamento de experiências e informações, acarretando benefícios para ambas as partes.

Todas as produções estão em sintonia com a área de especialidade dos professores, inclusive as em coautoria; 51% dos professores apresentaram pelo menos uma produção tecnológica em sua produção cadastrada no currículo Lattes, nos itens Produção Bibliográfica, Produção Técnica, ou nos Demais Trabalhos.

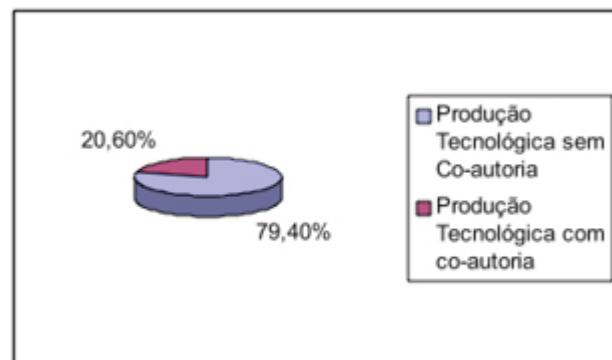
A área de Ciências Agrárias da UEL possui um estímulo maior para a produção bibliográfica, visto que todos a possuem, e que em vários casos a produção bibliográfica dissemina uma produção tecnológica.

Fujino, Stal e Plonski (1999) mencionam que a questão da pesquisa acadêmica deve ser caracterizada pela liberdade de investigação, e pela obrigatoriedade de incentivar o livre fluxo das informações por meio de publicações, e outros meios de divulgação dos conhecimentos gerados para a sociedade. Logo, os professores da área de Ciências Agrárias da UEL com esse alto índice em produção bibliográfica condicionam o livre fluxo das informações por meio das publicações para que seja disseminado a todos.

Independente da coautoria, a produção de cada professor foi considerada individualmente, situação que potencializa o aumento da produção tecnológica identificada. Outro fator que influenciou na quantidade da produção é o fato

de publicarem a mesma produção em diferentes línguas, ou quando o professor desenvolve uma produção tecnológica a dissemina em publicações e também a cadastrá como produção técnica.

**Gráfico 1 – Total da produção tecnológica da área de Ciências Agrárias**

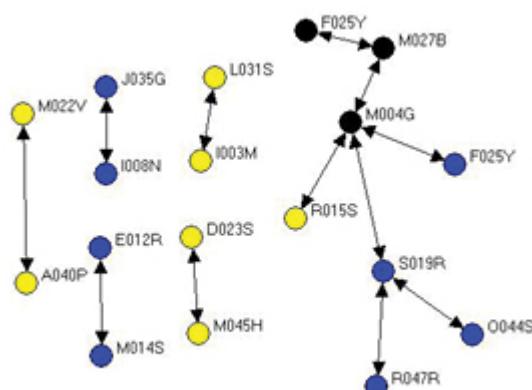


Fonte: os autores.

Logo essa produção pode ser cadastrada mais de uma vez nos itens de: Produção Bibliográfica, Produção Técnica e Demais trabalhos do currículo Lattes, o que também potencializou o aumento da produção. Esse impacto quantitativo ilusório na produção só foi percebido no final da análise dos dados levantados.

O total de 65 itens de produção tecnológica identificados nesta pesquisa teve um aumento de 20,6%, que equivale a 14 produções tecnológicas, devido às coautorias entre os professores contemplados com a bolsa de iniciação científica (Gráfico 1).

**Figura 2 – Mapeamento por produção tecnológica**



Fonte: os autores.

As produções em coautoria atingiram 20% do total de coautores, a temática centrou-se, principalmente, na especialidade embalagens para alimentos. Os Departamentos de Agronomia, Ciência e Tecnologia de Alimentos e o de Zootecnia estão no grupo que corresponde aos 20% do total de coautorias na temática de embalagens para alimentos, identificadas na Figura 2 pelas ligações do pesquisador M004G.

Aproximadamente Setenta e nove por cento (79,4%) da produção tecnológica encontrada não foi desenvolvida em parceria, percentual que totalizou 51 produções tecnológicas. Em alguns casos, as publicações foram compiladas com outros autores (não participantes desta pesquisa).

A Figura 2 possibilita a visualização por coautorias dos atores e por tipo de produção. Os círculos com a cor amarela se referem às produções tecnológicas classificadas como processos tecnológicos. A cor azul se refere aos produtos tecnológicos que foram produzidos aos pares, bem como em grupos de três professores, com destaque para o ator S019R que liga três pesquisadores. E os pontos com a cor preta se referem aos pesquisadores que têm produção tecnológica de produtos e também processos tecnológicos, o maior número de ligações mantidas foi pelo ator M004G, com quatro ligações.

### Produção Bibliográfica, Técnica e Demais Trabalhos

A análise, aqui apresentada, envolve (no geral) os três tipos de produção: a bibliográfica, a técnica e os demais trabalhos. A análise específica, demonstrada a seguir, evidencia as produções, cada tipo com sua singularidade.

O Departamento de Agronomia apresenta maior número de publicações bibliográficas, que disseminam produções tecnológicas, e produção técnica, obtendo 43% do total dessas produções mencionadas, e o de Medicina Veterinária Preventiva apresenta maior variedade nas produções.

Os Departamentos de Ciência e Tecnologia de Alimentos e Zootecnia mostram um bom nível de produção bibliográfica em relação ao Departamento de Agronomia que é o que possui o maior índice (39%), visto que os dois departamentos citados possuem menor número de professores e apresentam respectivamente 27% e 23% da quantidade total de publicação bibliográfica.

O Departamento de Clínicas Veterinárias possui poucos professores contemplados com bolsa do PROIC, o que justifica o baixo percentual de 1,4% da produção bibliográfica com produto tecnológico.

A produção técnica é liderada pelo Departamento de Agronomia com 56% da produção técnica. O Departamento de Agronomia demonstra, pelos altos percentuais em suas produções bibliográficas e técnicas, o interesse pela busca, descoberta, experimentações, dessa forma ocorre o desenvolvimento e adoção de novos produtos, novos softwares, e novos processos.

#### a) Produção bibliográfica

A Tabela 1 apresenta as publicações bibliográficas que disseminam produções tecnológicas. A análise desta tabela evidencia que:

- a) o Departamento de Agronomia tem a maior quantidade de publicações em coautoria, em relação aos outros Departamentos,
  - 40% do total das publicações são em coautorias;
  - 46% são publicações de autores vinculados apenas da UEL – autorias institucionais;
  - 51% estão publicadas em língua nacional;
  - 65% das publicações estão disponíveis em suporte on-line;
  - metade da produção encontrada está indexada no *ISI Web of Science*.
- b) os Departamentos de Zootecnia detêm 23% da produção bibliográfica total, e

**Tabela 1 – Produção Bibliográfica**

Departamento	Produção bibliográfica											
	Total	Ind.	Coaut.	Inst.	Interinst.	Pub. em Port.	Pub. em Inglês	Impresso	Eletr. On-line	CD-ROM	Qualis	ISI
Agronomia	17	01	16	15	02	20	06	07	13	00	13	05
Clínicas Veterinárias	01	00	01	01	00	01	00	01	01	00	01	00
Medicina Veterinária Preventiva	04	03	01	03	01	02	02	01	02	01	02	02
Ciência e Tecnologia de Alimentos	12	00	12	08	04	09	04	06	12	00	12	02
Zootecnia	10	00	10	06	04	07	06	02	09	00	04	02

Fonte: os autores.

- o de Ciência e Tecnologia de Alimentos 27%;
- c) o Departamento de Ciência e Tecnologia de Alimentos participa em 36% das produções bibliográficas de autoria institucional e, 72% de produção bibliográfica de autoria Interinstitucional, e o Departamento de Zootecnia detêm 30% de produção bibliográfica de autoria institucional e 36% produção bibliográfica de autoria interinstitucional;
- d) há a prevalência de publicação em periódicos eletrônicos com a média de 28% do total das produções bibliográficas;
- e) no Departamento de Medicina Veterinária Preventiva constatou-se que 9% da produção bibliográfica dissemina uma produção tecnológica, porém, apresenta 50% dessa produção no Sistema Qualis<sup>3</sup> da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e indexada no *ISI Web Of Science*. Este Departamento demonstra que produz tanto em veículos impressos (25%), como em eletrônicos *on-line* (50%) e *off-line* (CD-ROM) – 25% da sua produção bibliográfica;
- f) o Departamento de Clínicas Veterinárias publicou somente em língua nacional,

em periódico de caráter impresso e eletrônico, certificado pelo sistema Qualis.

A modalidade de publicação eletrônica tem se tornado um meio de disseminação importante, e gerou o aumento da quantidade de publicação, por possibilitar diversas vantagens e facilidades (CAFÉ *et al.*, 2002).

O fato de haver autorias interinstitucionais indica que as parcerias podem possibilitar a troca e compartilhamento de conhecimento e equipamentos, auxiliando o desenvolvimento das produções. A interação entre a Universidade e outros órgãos externos garante vantagens de desenvolvimento e caracteriza o espaço ou ambiente de produção tecnológica.

Conforme discorrido por Santos (1992), sobre o espaço de produção tecnológica ser uma estrutura demográfica, com uma produção, com estruturação de arranjos e modelos produtivos que garantam o suporte adequado para o desenvolvimento da produção tecnológica. Com isso, nota-se que as autorias interinstitucionais são responsáveis pela estruturação do espaço tecnológico, pois por meio dessas publicações há a oportunidade de ampliar os recursos, o que possibilita desenvolver as produções tecnológicas com maior qualidade.

<sup>3</sup> <http://qualis.capes.gov.br/webqualis/>.

Tabela 2 – Produção Técnica dos professores contemplados com bolsa do PROIC

Departamento	Produção bibliográfica				
	Total	Individual	Co-autoria	Institucional	Interinstitucional
Agronomia	13	02	11	13	00
Clínicas Veterinárias	00	00	00	00	00
Medicina Veterinária Preventiva	07	01	06	07	00
Ciência e Tecnologia de Alimentos	02	01	01	02	00
Zootecnia	02	00	02	02	00

Fonte: os autores.

### b) Produção técnica

Pela análise da Tabela 2, que aborda a produção técnica dos professores contemplados com bolsa do PROIC, pode-se perceber que:

- o Departamento de Agronomia possui 55% do total das produções técnicas, 84% são em coautoria;
- o Departamento de Medicina Veterinária Preventiva detém 29% das produções técnicas, há também a prevalência de 85% de coautoria;
- o Departamento de Ciência e Tecnologia de Alimentos atinge 8% do total das produções técnicas, 50% de autoria individual e 50% em coautoria;
- o Departamento de Zootecnia também possui 8% do total das produções técnicas, com 100% em coautoria e,
- o Departamento de Clínicas Veterinárias não tem produção classificada como produção tecnológica;
- 100% da produção são de autores vinculados a uma mesma instituição – a UEL.

Na Tabela 2 observa-se a prevalência da produção em coautoria. Além disso, nem todas as produções técnicas encontradas se configuram como produção tecnológica, como por exemplo: as palestras e as consultorias não apresentam características tecnológicas.

Identificou-se, também, que a Instituição, pelo menos, no que se refere à Produção Técnica,

não faz parcerias com outras instituições, o que deixa a UEL como a única detentora da produção identificada. Percebeu-se que é possível produzir individualmente, ou em conjunto dentro da própria Universidade, mostrando certa autossuficiência.

Quadro 3 – Descrição de Manuais

Manual por Departamento	Finalidades/ Características
Medicina Veterinária Preventiva	Manual Educativo sobre Zoonoses, Manual Técnico de Leishmaniose Tegumentar Canina produzido por I008N – (Especialidade: Zoonoses e Saúde Pública).
Medicina Veterinária Preventiva	Manual com o Programa de controle as Zoonoses também produzido por I008N dentro de sua especialidade.
Manual por Departamento	Finalidades/ Características.
Medicina Veterinária Preventiva	Manual para vigilância de Epizootias em Primatas Não Humanos do Estado do Paraná (SESA-PR) também produzido por I008N.
Medicina Veterinária Preventiva	Manual em CD-ROM que contém o treinamento em métodos de diagnóstico e controle da brucelose e tuberculose e noções em EET. Elaborado por E033M inserido na sua Subárea de Medicina Veterinária Preventiva.
Medicina Veterinária Preventiva	Manual que consta o programa nacional de controle e Erradicação da brucelose e tuberculose animal, se caracterizando manual técnico, também elaborado por E033M dentro da sua perspectiva.
Ciência e Tecnologia de Alimentos	Manual desenvolvido com finalidade didática ou instrucional (apostila). Contendo o Perfil Livre e avanços em análise sensorial descritiva. Elaborado por M027B (Ciência e Tecnologia de Alimentos) dentro de sua especialidade em Avaliação e Controle de Qualidade de Alimentos; Química, Física, físico-químico e Bioquímica dos Alimentos. e das Matérias-Primas Alimentares.

Fonte: os autores.

### c) Demais trabalhos

No sub-tópico “Demais trabalhos” do currículo Lattes, o único Departamento que apresenta produção tecnológica é o de Medicina Veterinária Preventiva, com o índice de 67% de coautoria e 33% de autoria individual.

Foi identificado no Departamento de Medicina Veterinária Preventiva o percentual de 67% de autoria institucional e 33% de autoria interinstitucional do total de três produções tecnológicas encontradas nesse Departamento.

Na produção tecnológica identificada como manual (Quadro 3), há a prevalência de 83% da produção total pelo Departamento de Medicina Veterinária Preventiva, desenvolvida por dois doutores na especialidade de “zoonoses e saúde pública”.

Percebe-se pelo Quadro 3 que, além do Departamento de Medicina Veterinária, o Departamento de Ciência e Tecnologia de Alimentos também produziu manuais em temas diversificados. Os outros professores, contemplados com a bolsa de iniciação científica, vinculados aos demais departamentos não produziram manuais.

Os processos foram encontrados em maior número, em relação às demais tipologias de produção tecnológicas relatadas neste estudo, o que indica a preferência no desenvolvimento de processos. Todos os Departamentos produziram pelo menos um processo. Os processos desenvolvidos, segundo suas características, auxiliam no bom rendimento da agricultura, diminuindo os danos, gastos, epidemias, tempo do agricultor, e agilizando o processo de desempenho das atividades a serem executadas.

E ainda, os processos apresentam soluções que podem resultar em benefícios para o desenvolvimento econômico da sociedade. Encontramos 23 processos com maior ênfase na área de alimentos, especificamente no melhoramento da qualidade.

Alguns dos produtos originaram-se dos processos encontrados, como por exemplo, o método isolado para controle da broca do café, que gerou um protocolo (produto tecnológico) para teste *in vitro*. Outro exemplo é o processo para obtenção de fécula de inhame, que acarretou na criação do produto “filme biodegradável”. Constatou-se então que a partir do conhecimento prévio, ressaltado por Machado (2006) como parte do conhecimento acumulado, os professores investigados desenvolveram produtos tecnológicos.

Na tipologia categorizada como produto, identificou-se 25 itens com propósitos diversos, desde a fabricação de bandejas para alimentos, ou adaptações em produtos agrícolas, à criação de vacinas e micro-organismos. A contribuição dos produtos está relacionada a três finalidades distintas: a) aprimoramento, e/ou avanço da área de Ciências Agrárias; b) geração de desenvolvimento econômico e social; c) auxiliar a Universidade a cumprir parte de sua missão, que é produzir conhecimento para o desenvolvimento social.

Os *softwares* foram criados, em maior quantidade, pelo Departamento de Agronomia, seguido pelo Departamento de Zootecnia, e pelo Departamento de Ciência e Tecnologia de Alimentos.

Os *softwares* identificados (nove) têm como propósito auxiliar nos cálculos, na avaliação de riscos climáticos, simulações para aprendizado, e detecção de contaminação de alimentos.

Esses *softwares* respaldam a tomada de decisão do profissional que o utilizará, proporcionando a diminuição da margem de erro. Não foi encontrado nenhum *software* com registro no Instituto Nacional de Propriedade Industrial (INPI), sistema de propriedade industrial que garante por meio de leis e direitos, restrições e explicações em termos de proteção ao software (FRANÇA, 1997).

O INPI concede o registro de direito autoral para a modalidade “softwares” e neste estudo não foi identificado nenhum *software* que esteja

protegido pela legislação vigente de direito autoral.

Alguns dos *softwares* encontrados estão disseminados na Web para comercialização e podem ser recuperados pelo sistema de busca do mecanismo de busca Google.

### Canais de comunicação

A comunidade científica, com destaque neste estudo para os Departamentos de Agronomia e Zootecnia, usa amplamente os canais de informação, disseminando em grande escala suas publicações e divulgando suas pesquisas em periódicos variados, fatores que contribuem para que a informação alcance um número maior de pessoas.

Somente quatro dos periódicos que disseminam suas produções não são certificados pelo Sistema Qualis e metade das publicações está indexada no *ISI web of Science*. Também vale ressaltar que publicam em periódicos de acesso aberto (*Open Archives Initiative - OAI*) e restrito, sendo que a publicação neste último segmento dificultou um melhor detalhamento da produção tecnológica. Apenas 28% dos periódicos identificados são adeptos da filosofia de OAI. Isto indica que parte das pesquisas está sendo disseminada por acesso livre e fica evidenciado o interesse dos pesquisadores em divulgar os resultados ampla e irrestritamente.

Outro ponto evidente é a publicação endógena na revista “*Semina*” da Universidade Estadual de Londrina, que possui 32% do número de publicações.

### Considerações finais

A produção tecnológica da área de Ciências Agrárias da UEL alcançou o total de 65 itens, sendo a maior parte disseminada por meio da produção bibliográfica.

Percebeu-se que além da atividade de pesquisa básica, na área de Ciências Agrárias, há também

a atividade de pesquisa aplicada junto com a experimentação e com potencial para lançamento no mercado. Situação esta que pode acarretar em novos produtos, novos softwares, novos manuais e novos processos, como os identificados neste estudo.

O mapeamento das formas de produção tecnológica, além de ter proporcionado a identificação dos produtos, processos tecnológicos, softwares e manuais na área pesquisada, possibilitou a visualização dos trabalhos em parcerias. Os pesquisadores participantes deste estudo fazem parcerias em um mesmo Departamento, ou em Departamentos variados, dentro e fora da Universidade, interagindo e compartilhando as experiências, o conhecimento e os equipamentos, o que potencializa uma produção tecnológica mais bem formulada e com maiores chances de êxito.

Os canais de comunicação mais utilizados foram os periódicos, só havendo um livro e um capítulo de livro em que a produção foi divulgada.

Um percentual considerável de artigos foi publicado em periódicos adeptos da filosofia da *Open Archives Initiative (OAI)*, o que possibilita o acesso livre e alta disseminação dos resultados das pesquisas.

A produção dos pesquisadores demonstra que estão contribuindo para que a universidade atinja um de seus objetivos – contribuir no atendimento das necessidades da sociedade.

A interação entre os atores na universidade indica que a pesquisa acadêmica e a pesquisa que resulta em inovação tecnológica, além de fomentar o capital intelectual, auxiliam no bom desenvolvimento de atividades nas quais o compartilhamento de recursos e conhecimento é o principal ativo da produção tecnológica.

## Referências

- BERNARDES, R. Produção de estatísticas e inovação tecnológica PAEP 1996-2001. **São Paulo em Perspectiva**, São Paulo, v. 17, p. 3-4, 2003.
- CAFÉ, L. et al. Arquivos abertos: inovação tecnológica para comunicação científica na rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 25., 2002, Salvador. Comunicação para cidadania. **Anais...** Salvador: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2002.
- CALDAS, R. de A. et al. Gestão estratégica em ciência, tecnologia e inovação. **Parcerias Estratégicas**, Brasília, DF, v. 11, p. 48-73, 2001. Disponível em: <<http://www.cgee.org.br/parcerias/p11.php>>. Acesso em: 10 maio 2007.
- CONCEIÇÃO, C. P. Protagonistas e contextos da produção tecnológica em Portugal: o caso da invenção independente. **Sociologia: problemas e práticas**, n. 41, p. 119-138, 2003.
- FAVA-DE-MORAES, F. Universidade, inovação e impacto socioeconômico. **São Paulo em Perspectiva**, São Paulo, v. 14, n. 3, 2000. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-8839200000300003&lng=pt&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-8839200000300003&lng=pt&nrm=iso)>. Acesso em: 10 maio 2007.
- FRANÇA, R. O. Patente como fonte de informação tecnológica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 2, n. 2, p. 235-264, jul./dez. 1997. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/viewFile/636/425>>. Acesso em: 18 abr. 2007.
- FUJINO, A.; STAL, E.; PLONSKI, G. A. A proteção do conhecimento na universidade. **Revista de Administração**, São Paulo, v. 34, n. 4, out./dez. 1999. Disponível em: <<http://www.rausp.usp.br/download.asp?file=3404046.pdf>>. Acesso em: 13 jun. 2007.
- GOLDEMBERG, J. **Tecnologia apropriada**: encontros com a civilização brasileira. Rio de Janeiro, n. 3, 1978.
- LEMOS, C. Inovação na era do conhecimento. In: LASTRES, H. M. M.; ALBAGLI, S. (Org.). **Informação e globalização na era do conhecimento**. Rio de Janeiro: Campus, 1999. Cap. 5, p. 122-144.
- MACHADO, C. J. S. As relações entre tecnologia, inovação e sociedade. **Datagramazero**: revista de Ciência da Informação, Rio de Janeiro, v. 7, n. 1, fev. 2006. Disponível em: <[http://www.dgz.org.br/fev06/F\\_I\\_art.htm](http://www.dgz.org.br/fev06/F_I_art.htm)>. Acesso em: 19 abr. 2007.

SANTOS, M. **Espaço e método**. 3. ed. São Paulo: Nobel, 1992. 88 p.

SCHUMPETER, J. **Ensaios**: empresários, inovação, ciclos de negócio e evolução do capitalismo. Oeiras: Celta, 1996.

## Study of the technological production at the State University of Londrina (UEL): mapping the area of Agricultural Sciences by Platform Lattes

### Abstract

*Introduction: Technological products and processes employed to solve society's practical problems comprise technological production, which is the essential means to measure the technological development of the environment in which the institution is inserted. This research aimed to verify the State University of Londrina's (UEL) technological production in the field of Agro-Sciences. Method: The methodological approach was to identify and analyze technological production through the curricula of the 51 professors granted with scholarships by UEL's Scientific Initiation Program (PROIC) available at the Lattes Platform. The data collected for analysis was charted and categorized according to technological production typology. Results: Of the 65 registered items in the researchers' curricula lattes refer to the technological production. Those items were mapped in order to describe the types identified. It was observed that the most innovative specialties are Agronomy, Forestry and Forest Engineering, Agricultural Engineering, Animal Science, Veterinary Medicine, Fisheries Resources and Fishing Engineering, Science and Food Technology. Conclusions: The production resulting from research activities of technological development, has spread through the production literature, primarily journal articles. Researchers build partnerships in the same Department, or with others within and outside the University, interacting and sharing experiences, knowledge and equipment. This practice improves a technological production that is more likely to be successful. A considerable percentage of articles are published in Open Archives Initiative(OAI) journals, allowing free access and encourages the dissemination of research results.*

### Keywords

*Technological Production; Technological Development; Technological innovation.*

---

Recebido em 27 de fevereiro de 2011  
Aceito em 19 de março de 2011

---

<sup>1</sup> Graduado em Biblioteconomia - UEL, Pós-Graduando em Gestão Pública - Universidade de Mogi das Cruzes/SP. Especialista em Informações Técnicas, Culturais e Desportivas - Secretaria Municipal de Cultura / Biblioteca Mário de Andrade (São Paulo).

nserzedello@prefeitura.sp.gov.br

<sup>2</sup> Graduada em Biblioteconomia, Doutora em Ciência da Informação - UFMG. Professora - Universidade Estadual de Londrina/Departamento de Ciência da Informação.  
mitomael@uel.br

---

Como citar este artigo:

SERZEDELLO, N. T. B.; TOMAÉL, M. I. Produção tecnológica da Universidade Estadual de Londrina (UEL): mapeamento da área de Ciências Agrárias pela Plataforma Lattes. **AtoZ**, Curitiba, v. 1, n. 1, p. 23-37, jan./jun. 2011. Disponível em: <<http://www.atoz.ufpr.br>>. Acesso em:

# Metodologias que estimulam o compartilhamento de conhecimentos: a experiência do Global Forum América Latina – GFAL

Isabela Drago<sup>1</sup>, Karoline Aparecida Scroch Sato<sup>2</sup>, Maxiliano Ribeiro<sup>3</sup>, Helena de Fátima Nunes Silva<sup>4</sup>

## Resumo

**Introdução:** Investiga a contribuição das metodologias utilizadas no Global Forum America Latina (GFAL) como forma de compartilhamento de conhecimento, empregando a Investigação Apreciativa (IA) e a Facilitação Gráfica. Identifica a importância do ambiente para o compartilhamento, a importância da externalização do conhecimento e os fatores que facilitam o compartilhamento.

**Método:** Realiza uma pesquisa descritiva com abordagem qualitativa, a partir de questionários aplicados aos coordenadores da área de Desenvolvimento Organizacional da Federação das Indústrias do Estado do Paraná, promotora do evento no Brasil, e da análise da documentação relativa às metodologias utilizadas. **Resultados:** Caracterizou-se o ambiente de pesquisa e, com base na tabulação dos dados obtidos pelo questionário, procedeu-se a análise sob as seguintes categorias: importância do ambiente, externalização do conhecimento, fatores facilitadores e as metodologias como facilitadoras do compartilhamento de experiências. **Conclusões:** A Investigação Apreciativa (IA) pode ser considerada uma metodologia que facilita a construção de ambiente favorável (ba) para a criação de conhecimento assim como a Facilitação Gráfica contribui diretamente no processo de compartilhamento, pois facilita a explicitação dos conhecimentos.

## Palavras-chave

Compartilhamento de conhecimento; Socialização do conhecimento; Investigação apreciativa; Facilitação gráfica; Gestão do conhecimento; Global Forum América Latina.

## Introdução

Na sociedade do conhecimento, a constante busca pela inovação e pelos resultados tem levado às organizações a se estruturarem como organismos de aprendizagem, como produtoras de conhecimento. No entanto, para que haja criação de conhecimento, o ambiente organizacional deve ser propício, ou seja, é necessário ter um espaço compartilhado, denominado de contexto capacitante ou *ba*, em que as pessoas possam sentirem-se seguras e mobilizadas para compartilhar informações e conhecimentos com o grupo e com a organização.

Corroborando com esta premissa, de compartilhamento do conhecimento, o Global Forum America Latina – GFAL que, no Brasil, é promovido pela Federação das Indústrias do Estado do Paraná – FIEP utilizou duas metodologias, a Investigação Apreciativa e a Facilitação

Gráfica, com a finalidade de estimular o diálogo entre os especialistas e o compartilhamento de conhecimento. Os diálogos são norteados pelos passos da Investigação Apreciativa (IA) e a Facilitação Gráfica na composição do registro gráfico do conhecimento compartilhado entre todos os participantes dos eventos.

A Investigação Apreciativa utilizada como metodologia contribui para a gestão de mudanças, adaptando-se aos valores, crenças e desafios de negócios. Com a aplicação dessa metodologia propõe-se a ruptura dos paradigmas tradicionais de planejamento em uma organização, pois trabalha com alterações radicais em práticas de mudanças gerenciais. Sua abordagem é inovadora, desafia o modelo habitual de gestão, voltado para a resolução dos problemas ou dificuldades. Propõe o gerenciamento de mudanças pela busca do que há de melhor, do que dá vida a um sistema quando ele está em seu estado mais eficaz e capaz,

em termos econômicos, sociais, ecológicos e humanos.

A Facilitação Gráfica é utilizada no processo da Investigação Apreciativa e durante todo o encontro do GFAL, para registrar o conhecimento que está sendo compartilhado pelos participantes, ajudando a retratar sobre o que está sendo explicitado pelos grupos e para que as pessoas se vejam dentro do processo e o que podem, efetivamente, contribuir. Esta metodologia ajuda de maneira simples e clara no gerenciamento das informações e conhecimentos compartilhados durante o Global Forum América Latina, usando linguagem gráfica consistente em formato de desenhos por meio de documentos.

Este artigo tem uma abordagem qualitativa e descriptiva, a qual se justifica pelo objetivo de investigar a experiência do GFAL e a contribuição das metodologias utilizadas, Investigação Apreciativa e Facilitação Gráfica, para tornar efetivo o processo de compartilhamento de conhecimento. A coleta das informações deu-se na documentação da organização promotora do evento e por um questionário (com questões abertas, relacionado às metodologias utilizadas aos conceitos de compartilhamento de conhecimento) aplicado aos coordenadores da Área de Desenvolvimento Organizacional do Sistema FIEP, os quais foram selecionados pelo fato de terem visão sistêmica desse evento e pela vivência na aplicação das metodologias.

Para análise dos resultados levou-se em consideração por um lado, os conceitos referentes à Gestão do Conhecimento, em especial o compartilhamento. Por outro lado, as observações sobre o contexto do ambiente pesquisado e a aplicação das metodologias.

## Informação x conhecimento

Muito se diz que se vive na sociedade da informação e do conhecimento e que apenas as empresas que souberem gerenciar suas informações e conhecimentos vencerão no acirrado mercado competitivo. Por este motivo

torna-se importante iniciar a abordagem teórica diferenciando informação e conhecimento para a análise dos demais temas e para a compreensão da dificuldade do compartilhamento do conhecimento, objeto de estudo do presente trabalho. Informações e conhecimento passam a ser recursos de desenvolvimento, de criação de valor, ligados diretamente ao sucesso das organizações.

Informação é, de acordo com Davenport (1998, p. 18-19), um conjunto de dados dotados de relevância e propósito que requerem análise, exigem consenso em relação ao significado e exige mediação humana. Somente as pessoas transformam dados em informação. Informação é a transformação do dado, quando é interpretado e está contextualizado. Para Barreto (1999) a informação sintoniza o mundo. De acordo com McGee e Prusak (1994, p. 5), "a informação que cria valor significativo para as organizações possibilita que novos produtos e serviços sejam criados e aperfeiçoa a qualidade do processo decisório na organização". As informações se transformam em insumos ao processo decisório (BARBOSA, 1997). Obter vantagem competitiva é a principal razão para o uso estratégico da informação, por este motivo o recurso informação deve ser tratado da mesma maneira que qualquer outro recurso (STAREC; GOMES; BEZERRA, 2005, p. 61). Ter a informação certa, no momento certo é fator crítico de sucesso e sobrevivência para as organizações (JANNUZZI e MONTALLI, 1999).

O conhecimento é entendido como as informações que estão na mente das pessoas. Informações estas valiosas, que exigem reflexão, síntese e contexto. "Conhecimento é uma abstenção interior, pessoal, de alguma coisa que foi experimentada por alguém (SETZER, 1999). O conhecimento é de difícil estruturação, não possibilita transferência e é de difícil disseminação. O conhecimento, normalmente é tácito, só existe na mente humana e é de difícil explicitação (DAVENPORT, 1998, p. 18-19). No contexto das organizações, o conhecimento é entendido como a experiência acumulada de seus colaboradores.

Considerando a definição de conhecimento, é importante diferenciar o conhecimento explícito e o conhecimento tácito. O conhecimento explícito é codificado, é objetivo, está representado em linguagem formal e permite armazenamento (CARVALHO, 2006). Este tipo de conhecimento é fácil de ser compartilhado. O conhecimento tácito, em contrapartida, é difícil de formalização, por ser um conhecimento pessoal, e, por este motivo também é de difícil comunicação e subjetivo, constituído de insights e intuições. “Enquanto permanece guardado como *know how* pessoal, o conhecimento tácito tem pouco valor para a organização. Porém, apesar da dificuldade de se codificar o conhecimento tácito, seu valor substancial compensa o esforço” (DAVENPORT e PRUSAK, 1998, p. 99). O conhecimento tácito está, fortemente, associado ao processo de inovação nas organizações, para identificar e solucionar problemas, e antecipar ações (TERRA, 2001, p. 69).

Com a intensificação do uso da informação como fim estratégico começou-se a identificar os aspectos intangíveis que faziam parte dos processos organizacionais. Identificou-se que as informações por si só não poderiam solucionar todos os problemas da organização. De acordo com Terra (2001, p. 180) a transformação de informações para conhecimentos depende da mente humana e de interpretação, que são baseadas em experiências e conhecimento prévio acumulado. Com isto o conhecimento passa a ser valorizado cada vez mais e surge a expressão gestão do conhecimento, que aponta a integração e uso dos conhecimentos nas organizações com o objetivo de auxiliar na busca por bons resultados e aprendizagem organizacional.

## Criação de conhecimento

Retomando as colocações anteriores, de que para vencerem e se manterem competitivas as organizações precisam pensar em maneiras e práticas para criar conhecimento organizacional e gerar aprendizagem, deve-se saber, primeiramente, que a criação do conhecimento envolve a capacidade de “adquirir, criar, acumular

e explorar o conhecimento. Esta criação acontece quando o conhecimento dos indivíduos passa a ser cristalizado como parte da organização” (NONAKA e TAKEUCHI, 1997, p. 83). Porém, a criação não é uma “simples compilação de fatos, mas um processo humano singular e irredutível, que não se produz com facilidade” (VON KROGH; ICHIJO; NONAKA, 2001, p. 15).

Para que a construção de conhecimento exista é necessário um relacionamento sinérgico entre o conhecimento tácito e o explícito e a elaboração de processos sociais que possibilitam a criação de novos conhecimentos por meio da conversão do conhecimento tácito em conhecimento explícito (NONAKA e TAKEUCHI, 1997, p. 82-83). “Empresas criadoras de conhecimento, seriam, pois, aquelas que criam, sistematicamente, novos conhecimentos, disseminam-nos pela organização inteira e, rapidamente, os incorporam as novas tecnologias e produtos” (TERRA, 2001, p. 78).

Segundo a abordagem de Nonaka e Takeuchi (1997, p. 65) o conhecimento pode ser criado com base em duas dimensões: a ontológica, na qual o conhecimento só é criado pelos indivíduos; e a epistemológica, que considera que os homens adquirem conhecimentos criando e tendo suas próprias experiências. A conversão do conhecimento explícito para o tácito seria a dimensão epistemológica da criação do conhecimento. A dimensão ontológica que seria o nível organizacional ocorre quando vários processos de conversão do conhecimento tácito e explícito acontecem em um ciclo ascendente do indivíduo até pontos da organização com o ambiente (NONAKA e TAKEUCHI, 1997, p. 65). “E, nesse processo, o indivíduo assumiria o papel de criador, do grupo, de sintetizador e a organização, de amplificadora do conhecimento” (TERRA, 2001, p. 80-81).

A criação do conhecimento ocorre “pelo modelo da espiral do conhecimento baseada no comprometimento pessoal e em vários processos de conversão entre o conhecimento implícito e explícito, envolvendo desde o indivíduo até o grupo, a organização e o ambiente” (TERRA,

2001, p. 78). A partir desta perspectiva surge a criação do conhecimento a qual ocorre em forma de espiral, começando no nível individual e aumentando, ampliando comunidades de interação até chegar ao nível interorganizacional. Para os autores, a criação do conhecimento está, portanto, na mobilização e na conversão do conhecimento tácito em explícito. A partir de vários processos de conversão o conhecimento é criado (NONAKA e TAKEUCHI, 1997, p. 62).

O processo de conversão envolve: a socialização, interação entre conhecimentos tácito e tácito; a externalização, interação entre conhecimentos tácito e explícito; a combinação interação entre conhecimentos explícito e explícito, e a internalização, interação entre conhecimentos explícito e tácito. Entretanto, para que as organizações prosperem e gerem conhecimento é necessário que o processo de criação seja contínuo, pois é necessário que o conhecimento tácito seja convertido em explícito para alavancar e gerar vantagens competitivas para a organização como um todo.

De acordo com Nonaka e Takeuchi (1997, p. 83-95), são cinco as condições capacitadoras, em nível organizacional, que promovem a espiral do conhecimento: intenção, autonomia, flutuação e caos criativo, redundância e variedade de requisitos.

Cada vez mais as organizações estão considerando a criação de conhecimento como fator de competitividade. Porém, isto não basta para o sucesso. É necessário que as organizações realmente dissemelhem e promovam a criação de conhecimento. É importante que exista um processo sistematizado e consciente para que a criação de conhecimento possa ser efetiva, gerando assim resultados positivos.

## Compartilhamento de conhecimento

Reconhecida a importância da criação do conhecimento organizacional é necessário que este conhecimento seja compartilhado para que se gerem novos produtos, processos

e serviços. As experiências vivenciadas pelos colaboradores precisam ser transmitidas, pois o compartilhamento é vital para o sucesso das organizações e deve ser feito com rapidez e eficiência a toda a organização. “As idéias causam maior impacto quando são amplamente compartilhadas, e não quando mantidas em poucas mentes” (GARVIN, 2000, p. 68).

Vários autores têm ressaltado a dificuldade de compartilhar o conhecimento tácito, pois este reside na mente de cada um dos colaboradores. Repassar a alguém detalhes de habilidades é algo extremamente difícil e subjetivo. O conhecimento tácito é considerado ambíguo e de difícil transferência para as outras partes da organização. Para Stewart (1998, p. 65), o conhecimento tácito, além de ser de difícil explicação e comunicação, é ainda de difícil identificação e articulação: “as pessoas sabem mais do que percebem – ao longo dos anos, elas desenvolvem enormes repertórios de habilidades, informações e formas de trabalhar que internalizaram a ponto de esquecer”.

O desafio que se percebe é, portanto, conseguir converter o conhecimento individual de maneira que a organização constantemente aprenda e incremente seus conhecimentos.

A dúvida é: como conseguir compartilhar conhecimento? Outra dúvida é: quem precisa saber e quanto, para facilitar o compartilhamento? De acordo com Davenport e Prusak (1998, p. 107) o compartilhamento acontece quando a organização contrata pessoas perspicazes e deixa que estas pessoas conversem entre si. A questão já conhecida é que o conhecimento tácito é de difícil transferência e de que o compartilhamento é uma tarefa difícil. Por este motivo, as organizações devem facilitar o compartilhamento.

É preciso criar processos e orientação e/ou aprendizado para que os conhecimentos tácitos que são importantes na organização não fiquem concentrados em uma única pessoa (DAVENPORT e PRUSAK, 1998, p. 99). “Para que informações e conhecimentos possam ser compartilhados de modo fácil, confortável e transparente, é preciso que haja propício

ambiente para compartilhar o conhecimento a ser implantado" (MCINERNEY, 2006, p. 57). É necessário, em primeiro lugar, que a organização adote métodos de compartilhamento que estejam de acordo com a cultura organizacional.

Von Krogh, Ichijo e Nonaka (2001, p. 17) apontam que "a capacitação do conhecimento deve ser vista de maneira circular; objetivar sempre a ampliação do potencial de criação do conhecimento". Para isto, apresentam cinco capacitores do conhecimento que são atividades que afetam positivamente a criação do conhecimento e preparam o ambiente para o compartilhamento, são eles: Capacitor 1- Instalar a visão do conhecimento; Capacitor 2 - Gerenciar conversas; Capacitor 3 - Mobilizar ativistas do conhecimento; Capacitor 4 - Criar o contexto adequado e, Capacitor 5 - Globalizar o conhecimento local.

Além dos capacitores existem maneiras de compartilhar o conhecimento tácito. O compartilhamento pode ocorrer pela observação direta, observação direta e narração, imitação, experimentação e comparação e execução conjunta. (VON KROGH, ICHIJO; NONAKA, 2001, p. 106-107).

O processo de aprendizagem envolve uma mistura dos processos de observação, imitação, narração, experimentação e execução (VON KROGH; ICHIJO; NONAKA, 2001, p. 107).

As cinco condições capacitadoras e os modos de conversão do conhecimento e, considerando a variável tempo resulta em um modelo integrado das cinco fases do processo de criação do conhecimento organizacional: Compartilhamento do conhecimento tácito, Criação de conceitos, Justificação de conceitos, Construção de arquétipo ou protótipo e Difusão interativa do conhecimento.

Outro fator relevante observado por Von Krogh, Ichijo e Nonaka (2001, p. 217) é que "a criação do conhecimento está diretamente relacionada com a criação de um espaço compartilhado, denominado de contexto capacitante, ou *ba*,

palavra de origem japonesa, que tem tradução aproximada de local onde acontece um fato". O *ba* pode ser interpretado como o espaço, ou contexto, compartilhado para novos relacionamentos, servindo como uma base para a criação do conhecimento. O *ba* sustenta o conhecimento individual e/ou coletivo (OUTI e STRAUHS, 2004).

No entanto, é importante ter em mente que nem todos os contextos são adequados ao compartilhamento e à criação de conhecimento, por enfrentarem barreiras individuais e organizacionais. Davenport e Prusak (1998, p. 117-118) apontam sete barreiras que dificultam este processo: falta de confiança mútua; diferentes culturas, vocabulários e quadros de referência; falta de tempo e de lugar de encontro; idéia estreita de trabalho produtivo; falta de capacidade de absorção pelos recipientes; crença de que o conhecimento é prerrogativa de determinados grupos; síndrome do "não criado aqui" e intolerância com erros ou necessidades de ajuda. Starec, Gomes e Bezerra (2005, p. 55-61) também apontam as dificuldades enfrentadas pelas organizações com relação às informações e conhecimentos. Os autores denominam tais problemas como sendo "pecados organizacionais", são eles: avareza, concentração da informação somente a alguns, poder de reter a informação; gula, excesso de informação levando ao caos informacional; ira, insatisfação com a falta de informação, imprecisão e ineficácia; inveja, problemas relacionados à ética informacional; luxúria, problemas na estética da comunicação, falhas no processo; preguiça, obsolescência da informação e soberba relacionada ao poder de possuir a informação.

## **Metodologias que estimulam o compartilhamento de conhecimento**

O foco deste trabalho é analisar como as metodologias empregadas no Global Forum América Latina (GFAL), estimulam o compartilhamento de conhecimentos entre os participantes. Neste sentido, são descritas a Facilitação Gráfica e a Investigação Apreciativa.

## Facilitação Gráfica

A Facilitação gráfica é um estilo de facilitação de grupo que usa visualização e a escuta do modo que usamos a nossa língua para apresentações (SIBBET, 2006, p. 3). Segundo Sibbet (2006, p. 3) a metodologia tem como base fundamental um facilitador o qual tem o papel de registrar publicamente as sínteses de todo o contexto apresentado por grupos e palestrantes. Por meio de utilização de *flip charts*, papéis largos, murais, *templates* gráficos e outras mídias visuais. Ela também envolve documentação de processo grupal em tempo real e proporciona às pessoas caminhos para, visualmente, lembrar seus compromissos e trilhar os progressos em uma série de reuniões. A facilitação gráfica pode combinar a redação gráfica com a prática de planejamento e guia de grupos na busca e consolidação de objetivos comuns.

Seu objetivo é gerenciar informações, durante a aplicação de seu processo, usando linguagem gráfica consistente e apresentando desenhos por meio de documentos.

## Investigação Apreciativa

Em meados da década de 1980, David Cooperrider, Suresh Srivastva e alguns colegas na Case Western Reserve University, Cleveland, Estados Unidos, introduziram o termo *Appreciative Inquiry* (Investigação Apreciativa – IA) a partir de uma visão diferente de organização, de gerenciamento, de informação e de conhecimento.

A Investigação Apreciativa é uma busca cooperativa do melhor nas pessoas, nas suas organizações e no mundo ao redor. Envolve a descoberta sistemática do que dá “vida” a um sistema quando ele está no seu estado mais eficaz e capaz, em termos humanos, ecológicos e econômicos. A IA reforça a capacidade de um sistema de elevar o seu potencial positivo. Além disso, mobiliza a “investigação” artesanalmente, na preparação de uma “pergunta positiva

incondicional”, envolvendo normalmente centenas ou algumas vezes milhares de pessoas. As abordagens de mudança e desenvolvimento organizacionais utilizadas atualmente pelas organizações têm influência direta na aprendizagem das pessoas. Loures (2010, p. 9) comenta que a metodologia da Investigação Apreciativa (IA) foi utilizada pela empresa Nutrimental S.A., pelo fato de ser um efetivo instrumento de transformação organizacional que focaliza a aprendizagem a partir do sucesso.

A IA é a arte e a ciência de fazer perguntas no processo de intervenção, que fortalecem a capacidade de um sistema de aprender, antecipar e intensificar seu potencial positivo, porque dá espaço à imaginação e à inovação; em vez de negação, crítica e diagnóstico redundantes, acontecem a descoberta, o sonho e o planejamento. A IA efetiva a união construtiva entre pessoas integrais e a totalidade massiva do que os indivíduos mencionam como capacidades passadas e presentes: conquistas, recursos, potenciais inexplorados, inovações, pontos fortes, pensamentos elevados, oportunidades, marcas de nível, pontos altos, valores vividos, tradições, competências estratégicas, histórias, expressões de sabedoria, insights mais profundos sobre o espírito incorporado ou alma - e visões de futuro valorizados e possíveis (COOPERRIDER e WHITNEY, 2006).

Considerando todos esses fatores juntos, a IA trabalha tendo como ponto de partida os relatos deste cerne de transformação positiva, pressupondo que cada sistema vivo detém muitos relatos ricos e inspiradores não penetrados no positivo. Ao ligar a energia deste cerne à agenda de transformação, tornam-se possíveis mudanças antes nunca imaginadas, de forma súbita e democrática.

A IA fundamenta-se em cinco princípios que sustentam sua concepção e aplicação, são eles: construcionista, da simultaneidade, poético, antecipatório e positivo.

Além dos princípios norteadores, a metodologia de Investigação Apreciativa compõe-se de 4 passos

básicos. O Ciclo dos 4 D originou dos termos em inglês *Discovery, Dream, Design e Destiny*, e refere-se aos passos de Descoberta, Sonho, Planejamento e Destino (COOPERRIDER; WHITNEY, 2006).

## Análise dos resultados

Os resultados obtidos referem-se à análise da documentação do Global Forum América Latina (GFAL) e de questionário aplicado junto aos coordenadores da Área de Desenvolvimento Organizacional do Sistema FIEP. Com a análise da documentação, tornou-se possível apresentar a caracterização do ambiente de pesquisa. Já os dados obtidos pelo questionário foram tabulados e analisados nas categorias: importância do ambiente, externalização do conhecimento, fatores facilitadores e as metodologias como facilitadoras do compartilhamento de experiências.

### Caracterização do ambiente de pesquisa

O mundo atual tem exigido, cada vez mais, o comprometimento dos cidadãos, das instituições e das empresas para seu desenvolvimento harmônico e equilibrado. Neste contexto, empresas conscientes de sua responsabilidade têm assumido um papel de destaque na criação e implementação de iniciativas voltadas ao desenvolvimento sustentável das comunidades e regiões onde atuam. As universidades, por outro lado, responsáveis que são pela formação profissional de jovens e futuros líderes, têm procurado incluir, em seus currículos, conhecimentos voltados à ética, responsabilidade social e ambiental.

O ambiente onde aconteceu o Global Forum América Latina (GFAL) mescla propósitos entre a área de negócios e a educação em gestão. Trata-se de um movimento mundial de conversação sobre inovação da educação para a sustentabilidade.

O Global Forum é uma iniciativa espontânea de organização da sociedade na busca de ações inovadoras para a sustentabilidade da vida

humana em nosso planeta. Surgiu no ano 2000, a partir do *Global Compact*, na ONU, fundindo-se com o *Business as an Agent of World Benefit* - BAWB em um encontro promovido por Kofi Annan e organizado por David Cooperrider, da Case Western Reserve. Somente em 2006, o Global Forum ganhou as características de um movimento que se expande em conversações por toda a América, até chegar ao Brasil em 2008.

Duas metodologias são utilizadas no GFAL no decorrer de todo o evento, com a finalidade de estimular o diálogo entre os especialistas e o compartilhamento de conhecimento. Os diálogos são norteados pelos passos da metodologia da Investigação Apreciativa (IA) tendo como foco os princípios que envolvem uma educação significativa e transformadora, identificada com as evoluções tecnológicas, com o desafio da sustentabilidade e com as novas demandas de um mercado globalizado.

Os especialistas convidados são selecionados pela experiência empresarial, conhecimento técnico e relevância de sua pesquisa científica, ação empreendedora-inovadora no seu setor; assim como pela capacidade de pensar o futuro da educação.

O evento contou com palestras proferidas por personalidades brasileiras e estrangeiras do mundo empresarial e da área acadêmica, com painéis, apresentações expositivas e rodadas de diálogo entre os participantes.

A Facilitação Gráfica foi usada dentro do processo da Investigação Apreciativa para registrar o conhecimento que estava sendo compartilhado pelos participantes, ajudando a retratar sobre o que estava sendo explicitado pelos grupos e contribuindo para que as pessoas perceberem-se dentro do processo e como elas poderiam efetivamente contribuir. Esta metodologia ajuda de maneira simples e clara no gerenciamento das informações e conhecimentos compartilhados durante todo do GFAL, usando linguagem gráfica consistente e formatos de desenhos por meio de documentos.

Tendo em vista a compreensão do contexto em que ocorreu o GFAL fez-se uma análise da importância do ambiente, da importância da externalização das experiências e conhecimentos, dos fatores e das metodologias facilitadoras do compartilhamento do conhecimento.

### Importância do ambiente

O Global Forum America Latina – GFAL faz uso da Investigação Apreciativa e da Facilitação Gráfica a fim de estabelecer um ambiente agradável, que estimule o diálogo entre os seus diversos participantes, promovendo, assim, o compartilhamento de conhecimentos.

A Investigação Apreciativa, por meio do ciclo dos 4 D (*Discovery, Dream, Design e Destiny*), é utilizada para nortear os diálogos dos participantes durante o encontro. Já a Facilitação Gráfica é utilizada na composição do registro gráfico dos conhecimentos compartilhados no GFAL.

As metodologias mencionadas ajudaram na composição do ambiente, favorecendo o estabelecimento de um ambiente, o *ba*. Segundo as afirmações dos respondentes, o ambiente constituído durante o GFAL favoreceu o compartilhamento do conhecimento, experiências e informações e as metodologias utilizadas estimularam a participação das pessoas, as quais formaram grupos de até oito pessoas e nestes, durante o desenrolar das atividades, os grupos fizeram relatos de experiências e conhecimentos, sendo este o foco principal da primeira fase do ciclo dos 4 D da Investigação Apreciativa, a Descoberta.

### Importância da externalização

Ao se referir à importância de explicitar o conhecimento o entrevistado João Gratuliano afirma que os participantes dos encontros apresentam um forte interesse em documentar o seu conhecimento tácito, levando-o para o explícito, por meio de formulários físicos ou digitais, propostos pelos organizadores dos

encontros. Entretanto, observa-se que este forte interesse se deu durante o evento, mas considerando a continuidade do processo o interesse é fraco.

Dois fatores são indicados como os que mais estimulam a documentação, sendo:

- a possibilidade de apresentar casos de sucesso para os demais participantes dos grupos;
- o interesse pessoal de contar suas estórias de sucesso.

A entrevistada Lindamir Mocellin Mueller, também considera que o interesse seja forte em registrar o conhecimento tácito e aponta também dois fatores principais:

- o material disponibilizado para registro, ou seja os formulários bem elaborados, e a utilização de computadores nas mesas;
- uma boa elaboração do caderno de registros (de fácil entendimento) seja manual ou em computadores.

### Fatores que facilitam o compartilhamento

Relatar casos de histórias bem sucedidas, momentos em que as pessoas se sentiram mais integradas, atuantes e motivadas podem ser indicados como os mais estimulantes para a troca de conhecimento, experiências e informações. Em contrapartida o tempo das atividades, durante os encontros, não é muito extenso e cada um dos participantes acaba por compartilhar poucas histórias ou casos de sucesso.

O entrevistado Gratuliano cita que “é possível aprender muito também com experiências que fracassaram”, mas o GFAL utiliza a Investigação Apreciativa e esta metodologia não abre esta possibilidade. Segundo Cooperrider e Whitney (2006, p. 3) a metodologia propõe incisivamente que as organizações não devem ser essencialmente problemas a serem resolvidos.

A entrevistada Lindamir afirma que o design específico da sala, o desenho da conferência e também o material criado nos encontros, bem como os questionamentos surgidos são os fatores que mais estimulam a troca de informações e conhecimentos. Contudo, aponta que a falta de vontade e abertura dos participantes, a dificuldade encontrada para dialogar e idéias pré-concebidas de alguns participantes são fatores que dificultam a troca de conhecimentos.

Durante os encontros do Global Forum, foram oportunizados contextos que facilitaram as atividades em grupo bem como favoreceram a criação e acúmulo de conhecimento em nível individual, esta colocação está sustentada nas palavras do entrevistado Gratuliano:

O ambiente e a agenda estão elaborados de forma a estimular a participação de pequenos grupos, mesas para 8 pessoas, distribuição do tempo valorizando a produção do grupo. Isto permite que todos tenham voz e que não fique cansativo escutar os outros, desta forma todos são estimulados a dar sua contribuição e exercitar o relato de seus conhecimentos, do tácito para o explícito.

Pelo que se percebe na fala do entrevistado, evidencia-se a socialização, a interação entre conhecimentos tácito e tácito; a externalização, interação entre conhecimentos tácito e explícito; a combinação interação entre conhecimentos explícito e explícito, e a internalização, interação entre conhecimentos explícito e tácito, bem como a sua relação com as cinco condições capacitadoras, em nível organizacional, que promovem a espiral do conhecimento (NONAKA e TAKEUCHI, 1997, p. 83-95).

As atividades realizadas durante o GFAL também contribuíram para a criação no sentido de proporcionar um ambiente favorável ao compartilhamento do conhecimento. Dentro do GFAL, os capacitores do conhecimento, atividades que contribuem para a criação do conhecimento e preparam o ambiente para o compartilhamento, na visão de Von Krogh, Ichijo e Nonaka foram trabalhados nas fases Sonho e Planejamento, da Investigação Apreciativa, as quais instigam os participantes à criação de algo inovador que faça uso de

conhecimento dos participantes na elaboração e preparação de propostas de projetos.

De acordo com a entrevistada Lindamir, a cada momento de diálogo a pessoa aprende, está socializando com os outros e o conhecimento está nas pessoas e nas suas experiências. Aponta que a disposição das mesas e a formação de grupos e os questionamentos também afetam positivamente a criação de conhecimento; é um ambiente favorável à troca.

As fases do processo de criação do conhecimento foram observadas durante o Global Forum América Latina, para o entrevistado Gratuliano:

embora nem todos os grupos alcancem um grau de estruturação consistente, quase sempre há uma construção coletiva que inclui diversas contribuições e possibilita a difusão desta construção com alguns integrantes. Boa parte da criação e das explicações dos protótipos não é registrada. Já a criação de conceitos, ou seja, o surgimento de algo completamente novo, isso não é muito comum, e sim a combinação de conceitos em uma proposta de um novo projeto, não necessariamente inovador.

Já para Lindamir, “Na visualização do futuro para a sustentabilidade, e na construção de protótipos as pessoas refletem sobre conhecimentos já existentes e juntas colocam à disposição todo o conhecimento que possuem para criar algo novo, aqui consiste o conhecimento”.

### Metodologias como facilitadoras do compartilhamento de conhecimento

Ao ser questionado como a metodologia Investigação Apreciativa pode contribuir para estimular o compartilhamento de conhecimento, o entrevistado Gratuliano ressaltou a importância, dentre os cinco princípios que sustentam e norteiam a metodologia, do **princípio construcionista**, pois ele estimula em todas as fases da IA o trabalho de criação coletiva, fazendo com que o compartilhamento do conhecimento aconteça no decorrer de todos os momentos de atividades em grupo. Nonaka e Takeuchi (1997) reforçam o princípio construcionista, pois pregam

que a organização não pode criar conhecimento, por si mesma. Isto é, sem que haja iniciativa do indivíduo e a interação que ocorre dentro do grupo. Portanto, os membros de uma equipe só criam novas perspectivas com o diálogo e o debate. A entrevistada Lindamir, em concordância afirma que as perguntas que são elaboradas, o construcionismo social e os aspectos positivos estimulam o compartilhamento, pois, sempre que a pessoa olha coisas positivas tende a ficar mais otimista e tem disponibilidade para construir o futuro e assimilar novos conhecimentos.

Quando se perguntou como a metodologia Facilitação Gráfica pode contribuir para o compartilhamento do conhecimento, o entrevistado Gratuliano comentou que, no GFAL, esta metodologia foi usada para que fosse um suporte ao registro dos conhecimentos por meio de desenhos e textos o que demonstra uma síntese das principais idéias apresentadas no evento. “Na minha experiência é muito útil para os participantes recordarem rapidamente do que aconteceu e recapitular frases e conclusões que nem sempre é possível anotar durante o evento”. Para a entrevistada Lindamir os registros feitos de forma colorida e em desenho estimulam a criatividade das pessoas e também o senso ético. O resumo faz com que as pessoas absorvam somente o que é essencial.

Pelos resultados da observação do contexto e das entrevistas pode-se afirmar que o compartilhamento de conhecimento no Evento se dá pelas metodologias empregadas, as quais privilegiam o indivíduo em constante interação com o grupo e também pelo fato de possibilitar o registro, a síntese dessa construção. A representação gráfica, como forma de síntese dos conhecimentos e experiências trocados durante o evento, pode estimular novas interações possibilitando um criar e recriar constantes.

## Considerações finais

O conhecimento passou a ser um dos elementos mais importantes para as organizações. Entretanto, gerenciar ambientes de conhecimento não é tarefa fácil, pois depende primeiramente de mudança cultural e de organizações que ofereçam espaços favoráveis à criação e compartilhamento do conhecimento individual para todos na organização. Somente oferecendo um contexto apropriado (*ba*) é que se torna possível ter criatividade, inovação e aprendizagem.

Conforme visto em Davenport e Prusak (1998) existem algumas barreiras que dificultam o compartilhamento de conhecimento e que devem ser transpostas. Para tal, é essencial a adoção de práticas e metodologias compatíveis com o processo de criação e compartilhamento do conhecimento individual e organizacional.

Para estudar como as metodologias podem auxiliar no processo de criação e compartilhamento de conhecimentos, analisou-se a experiência do Global Forum América Latina (GFAL), movimento mundial que tem como finalidade estimular o diálogo e a troca de conhecimentos sobre a inovação da educação para a sustentabilidade.

Verificou-se que as atividades desenvolvidas, durante o GFAL, foram norteadas pelos passos da Investigação Apreciativa (IA) e que a Facilitação Gráfica é utilizada a fim de compor um registro gráfico dos conhecimentos compartilhados entre os participantes do encontro.

Esta análise possibilitou observar que a IA pode ser considerada uma metodologia que facilita a construção de ambiente favorável (*ba*) para a criação de conhecimento, pois em todas as suas fases, o trabalho de criação coletiva é estimulado, fazendo com que o compartilhamento do conhecimento aconteça no decorrer de todos os momentos de atividade. Pela abordagem apreciativa, a metodologia consegue fazer com que as pessoas se sintam mais à vontade para contribuir e compartilhar o que sabem e também para assimilar novos conhecimentos. Percebeu-se que se o clima é positivo, as informações circulam de forma mais simples e efetiva.

Porém, não basta somente compartilhar o conhecimento tácito se de alguma maneira ele não é explicitado para que outras pessoas também possam utilizá-lo. Tendo essa preocupação o GFAL utilizou a Facilitação Gráfica. Esta metodologia contribuiu diretamente no processo de compartilhamento do conhecimento, pois serviu como suporte ao registro dos conhecimentos por meio de desenhos coloridos e textos que demonstram uma síntese das principais idéias apresentadas no evento. Caracterizando-se, assim, como uma metodologia possível para o registro e visualização do conhecimento.

Embora existam outras metodologias para o compartilhamento, este estudo teve como foco apenas aquelas metodologias aplicadas no GFAL. Percebeu-se também que não há estudos científicos sobre as metodologias aplicadas ao compartilhamento de conhecimentos. As formas de compartilhar têm sido utilizadas de maneira empírica pelas organizações, necessitando de aporte teórico para a sistematização dessas práticas. Portanto, esta pesquisa não se limita ao que foi exposto, tendo possibilidades de trabalhos futuros. Alguns pontos podem constituir-se em motivação para outras pesquisas, como: a elaboração de proposta metodológica para compartilhamento de conhecimento, considerando as metodologias estudadas; a identificação ou criação de ferramentas que auxiliem as tarefas de compartilhamento de conhecimento; a realização de novos estudos com diferentes metodologias para o compartilhamento de conhecimento.

## Referências

- BARBOSA, B. N. **A gestão do conhecimento aplicada a projetos:** proposta metodológica. Curitiba, 2005. 134 f. Monografia (Graduação em Gestão da Informação) – Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná, 2005.
- BARRETO, A. A. A oferta e a demanda de informação: condições técnicas, econômicas e políticas. **Ciência da Informação**, Brasília, v. 28, n. 2, p. 167-177, 1999.
- CARVALHO, H. G. **Texto básico da disciplina “Gestão de informação tecnológica”** (Draft). Mestrado em Tecnologia. Centro Federal de Educação Tecnológica do Paraná – CEFET-PR. Curitiba, PR, 2000.
- COOPERRIDER, D. L.; WHITNEY, D. **Investigação apreciativa:** uma abordagem positiva para a gestão de mudanças. Rio de Janeiro: Qualitymark, 2006.
- DAVENPORT, T. **Ecologia da informação:** por que só a tecnologia não basta para ter sucesso na era da informação. São Paulo: Futura, 1998.
- DAVENPORT, T. H.; PRUSAK, L. **Conhecimento empresarial:** como as organizações gerenciam o seu capital intelectual. Rio de Janeiro: Campus, 1998.
- GARVIN, D. A. Construindo a organização que aprende. In: HARVARD BUSINESS REVIEW. **Gestão do Conhecimento**. Rio de Janeiro: Campus, 2000.
- JANNUZZI, C. A. S. C.; MONTALLI, K. M. L. Informação tecnológica e para negócios no Brasil: introdução a uma discussão conceitual. **Ciência da Informação**, Brasília, v. 28, n. 1, p. 28-36, 1999.
- LOURES, R. C. da. R. **Evolução de T&D e aprendizagem nas empresas.** Disponível em: <<http://www.unindus.org.br/material/uploadAddress/Evolu%C3%A7%C3%A3o%20de%20T&D.doc>>. Acesso em: 04 jan. 2010.
- MCGEE, J. V.; PRUSAK, L. **Gerenciamento estratégico da informação:** aumente a competitividade e a eficiência de sua empresa utilizando a informação como uma ferramenta estratégica. Rio de Janeiro: Campus, 1994.
- MCINERNEY, C. R. Compartilhamento e gestão do conhecimento: profissionais da informação em um ambiente de confiança mútua. In: TARAPANOFF, K. (Org.). **Inteligência, Informação e Conhecimento**. Brasília: IBICT, 2006. p. 57-72.
- NONAKA, I.; TAKEUCHI, H. **Criação de conhecimento na empresa:** como empresas japonesas geram a dinâmica da inovação. Rio de Janeiro: Campus, 1997.
- OUTI, N. S; STRAUHS, F. R. Visão nacional do processo de criação do conhecimento japonês: estudo de caso AirCon do Brasil. In: KM BRASIL – CONGRESSO BRASILEIRO, 2004, São Paulo. **Anais...** São Paulo : KMBrasil2004, 2004. v. 1.
- SETZER, V. W. Dados, informação, conhecimento e competência. **DataGramZero:** Revista de Ciência da Informação, n. 0, dez. 1999.

SIBBET, D. **Graphic facilitation**: transforming group process whit the power of visual listening. San Francisco: Grove, 2006.

STEWART, T. A. **A riqueza do conhecimento**: o capital intelectual e a nova organização. Rio de Janeiro: Campus, 2002.

TERRA, J. C. C. **Gestão do conhecimento**: o grande desafio empresarial. São Paulo: Negócio, 2001.

VON KROGH, G. V.; ICHIJO, K.; NONAKA, I. **Facilitando a criação de conhecimento**: reinventando a empresa como poder da inovação contínua. Rio de Janeiro: Campus, 2001.

### **Methodologies that stimulate knowledge sharing: The experience of The Global Forum Latin America - GFLA**

#### **Abstract**

*Introduction: This article looks into the contributions made by the methodologies employed in the Global Forum Latin America (GFLA) as a way to share knowledge, by means of Appreciative Investigation (AI) and Graphic Facilitation. It identifies the importance of the environment to this sharing, and also the importance of the externalization of knowledge, as well as the factors that make this sharing easier. Method: It performs a descriptive research with qualitative approach, departing both from questionnaires applied to coordinators in the field of Organizational Development of the Federation of Industries of the State of Paraná, which promoted the event in Brazil, and from the analysis of the documentation related to the methodologies used. Results: Describes the research environment and, based on the data collected through the questionnaire, the analysis was carried under the following categories: importance of the environment, outsourcing of knowledge, facilitating factors and methodologies to facilitate the experiences sharing. Conclusions: It concludes, from the analysis of the results, that AI can be considered as a methodology that facilitates the construction of a favorable environment (ba) to the creation of knowledge and that Graphic Facilitation directly contributes to the sharing process, since it provides an easier elucidation of one's knowledge.*

#### **Keywords**

*Knowledge sharing; Knowledge socialization; Appreciative Investigation; Graphic Facilitation; Knowledge Management; Global Forum Latin America.*

Recebido em 29 de março de 2011  
Aceito em 11 de abril de 2011

<sup>1</sup> Graduada em Gestão da Informação - UFPR, Especialista em Gestão do Conhecimento nas Organizações - UTFPR, Mestre em Ciência, Gestão e Tecnologia da Informação - UFPR. Gestora de Informações - Serviço Social da Indústria do Paraná (SESI-PR).  
isabeladrago@gmail.com

<sup>2</sup> Graduada em Gestão da Informação - UFPR, Especialista em Gestão do Conhecimento nas Organizações - UTFPR, Mestre em Ciência, Gestão e Tecnologia da Informação - UFPR. Assistente Administrativo - Banco do Brasil S.A.  
karoline.scroch@gmail.com

<sup>3</sup> Graduado em Administração de Empresas. MBA em Gestão de Projetos, Mestre em Ciência, Gestão e Tecnologia da Informação - UFPR. Professor - Universidade Positivo (UP).  
maxribeiro2003@gmail.com

<sup>4</sup> Graduada em Biblioteconomia, Mestre em Educação, Doutora em Engenharia de Produção. Professor Adjunto - Universidade Federal do Paraná/Departamento de Ciência e Gestão da Informação.  
helenanunes@ufpr.br

Como citar este artigo:

DRAGO, I.; SCROCH, K. A. S.; RIBEIRO, M. R.; SILVA, H. de F. N. Metodologias que estimulam o compartilhamento de conhecimentos: a experiência do Global Forum América Latina - GFAL. **AtoZ**, Curitiba, v. 1, n. 1, p. 38-49, jan./jun. 2011. Disponível em: <<http://www.atoz.ufpr.br>>. Acesso em:

Artigo originalmente publicado nos Anais do 9º Congresso Brasileiro de Gestão do Conhecimento - KMBrasil, 2010.

# Plano estratégico para o ambiente acadêmico: estudo de caso em grupo de pesquisa da Universidade Federal do Paraná

Ana Carolina Greef<sup>1</sup>, Francisco Daniel de Oliveira Costa<sup>2</sup>, Newton Corrêa de Castilho Junior<sup>3</sup>

## Resumo

**Introdução:** A competição entre concorrentes atuantes no ambiente acadêmico tem características distintas daquela entre agentes inseridos no mercado competitivo organizacional. Os objetivos estratégicos da primeira categoria de concorrentes referem-se à visibilidade e ao posicionamento no mercado de pesquisadores, grupos de pesquisa, Universidades. Assim, a aplicação de um plano estratégico ao Grupo de Pesquisa Aplicada em Ciência, Informação e Tecnologia (GP-CIT), poderá - potencialmente - agregar valor a processos e decisões neste ambiente. **Método:** Pesquisa aplicada envolvendo os seguintes passos: caracterização do ambiente; identificação da principal demanda em relação tipo de planejamento e voltada para a definição do foco; síntese do posicionamento do ambiente frente a seu mercado de atuação; diagnóstico de fatores internos e externos; detalhamento da competência central; definição de direcionamento estratégico; esboço da estratégia nos níveis de negócios, colaboração e crescimento. **Resultados:** As características da cultura interna e externa do GP-CIT são favoráveis à realização do planejamento estratégico com vistas a reforçar o potencial de recursos, competências e oportunidades. Tais aspectos fundamentaram o delineamento da direção estratégica a ser seguida pelo Grupo no ambiente acadêmico, com foco em uma visão de longo prazo e apoio em missão, objetivos e valores de negócio. Estratégias de diferenciação, colaboração e crescimento foram formuladas conforme objetivos de atuação do GP-CIT no ambiente acadêmico, considerando-se a dinâmica competitiva a ele inerente. **Conclusões:** A implementação das estratégias deve seguir passos cautelosos e decisões bem fundamentadas, agregando perspectivas sólidas e competitivas à atuação. O pensamento estratégico, enquanto visão holística, passa a ser inserido na cultura do Grupo de maneira a modificar não apenas processos a ele relacionados, mas todo o paradigma de atividades e competições no ambiente acadêmico.

## Palavras-chave

Planejamento estratégico; Universidade Federal do Paraná; Grupos de pesquisa; Competitividade; Gestão da informação.

## Introdução

Encaminhar uma organização “ao cumprimento de seus objetivos fundamentais” (HARRISON, 2005, p. 36) caracteriza um dos objetivos que compõem o planejamento estratégico, efetivado não apenas por diretrizes estabelecidas, mas também por “uma perspectiva integrada do empreendimento” (HARRISON, 2005, p. 40).

Ciclos analíticos e planos de ação integram o composto metodológico do planejamento e pensamento estratégicos, originando um plano que apresenta à organização a sua posição atual

no mercado de atuação e meios para atingir determinada posição futura, seja neste mesmo mercado ou em ambientes diferenciados.

O ambiente acadêmico apresenta, inerentemente, um cenário competitivo específico que se diferencia do contexto mercadológico quanto à natureza de competições entre os atores. Nesse, Universidades, grupos de pesquisa e pesquisadores atuam conforme objetivos de longo prazo e buscam visibilidade e posicionamento no mercado frente aos demais competidores.

Essa realidade reflete-se na atuação do Grupo de Pesquisa Aplicada em Ciência, Informação e Tecnologia (GP-CIT), da Universidade Federal do Paraná, cujos projetos e ações têm potencial considerável em termos de contribuição à comunidade científica e provimento de visibilidade ao Grupo em si. Em complemento, recursos e relações por ele mantidos têm quantidade, qualidade e características valiosas, raras, inimitáveis e organizáveis que encaminham o GP-CIT à vantagem competitiva.

Um plano estratégico desenvolvido no âmbito do Grupo viria a alavancar a administração desses potenciais em direção a uma posição futura desejada, determinando objetivos, indicadores e metas para atingir tal posicionamento e alinhando táticas e operações cotidianas a esses aspectos de maior abrangência.

A premissa acima apontada dá origem ao estudo de caso aqui descrito, no qual detalham-se aspectos teóricos e práticos relacionados à elaboração do plano estratégico, de modo que esse atenda necessidades estratégicas por parte do GP-CIT. Esta pesquisa caracteriza-se como aplicada – mesclando vertentes teóricas e práticas ao longo de sua descrição – utilizando técnicas para obtenção de resultados, adaptadas do tradicional contexto corporativo à academia. São elas, na sequência utilizada para o desenvolvimento da pesquisa:

- a) caracterização do ambiente a ser avaliado e ao qual deve ser aplicado o planejamento estratégico;
- b) identificação da principal demanda do referido ambiente em relação ao citado planejamento, definindo o foco deste último (competências e recursos internos ou a concorrentes e fornecedores, por exemplo);
- c) síntese do posicionamento do ambiente em estudo em seu mercado de atuação;
- d) diagnóstico de fatores internos e externos, mais especificamente as forças competitivas de concorrentes, fornecedores, substitutos e parceiros, os recursos e relações como pessoas,

informações, finanças e estrutura física, forças e fraquezas, oportunidades e ameaças ambientais;

- e) detalhamento da competência central do mesmo, que o diferencia de concorrentes e potencializa a procura do público-alvo por seus serviços;
- f) definição de direcionamento estratégico, composto por visão, missão, valores e objetivos que pretende alcançar no futuro;
- g) esboço propriamente dito da estratégia nos níveis de negócios, colaboração e crescimento, a ser seguida para o alcance do referido direcionamento, suportada pela competência central e demais recursos internos do ambiente estudado.

O artigo pretende apresentar componentes da prática do planejamento estratégico no ambiente acadêmico, bem como aspectos relevantes para sua efetiva inserção nesse mesmo contexto, viabilizando a replicação do processo aqui descrito por parte de qualquer entidade voltada à pesquisa.

## O grupo de pesquisa aplicada em ciência, informação e tecnologia

Criado em 2007, o GP-CIT está instalado no Setor de Ciências Sociais Aplicadas da Universidade Federal do Paraná, no âmbito do Departamento de Ciência e Gestão da Informação. Sua coordenação caracteriza-o como jovial e interdisciplinar, cujas linhas de pesquisa extrapolam a dinamicidade natural do ambiente científico e são, a saber: educação continuada e corporativa com inserção ou não de tecnologia; engenharia ergonômica de software: teoria, modelos e metodologia; fundamentos e práticas do direito sobre o trabalho intelectual; gerência de projetos e informações; workflow aplicado a sistema de gestão para projetos digitais. (BRASIL, 2009).

São subordinados ao GP-CIT, o Laboratório de Mídias Digitais (Labmidi) e o Centro de Estudos

em Realidade Virtual Aplicada (Cerva), que englobam, respectivamente, pesquisas em Educação e a relação dessa com Tecnologias da Informação, e; Arquitetura da Informação aplicada à Realidade Virtual.

A participação em editais de financiamento de projetos disponibilizados por entidades como o Conselho Nacional de Desenvolvimento Tecnológico (CNPq) e a Fundação Araucária permite ao GP-CIT obter recursos para aquisição de equipamentos, remuneração de colaboradores e desenvolvimento de produtos.

## **Escopo do planejamento Estratégico**

### Demandas identificadas

A recente formação do GP-CIT, paralela ao fato de que a cultura gerencial e de atividades desse encontra-se em construção, oportuniza a criação de medidas de planejamento estratégico ao grupo de pesquisa acadêmica. Tais medidas visam oferecer ao Grupo um posicionamento de destaque no ambiente em que atua, por meio de eficiência, eficácia e efetividade de ações e decisões.

Tais fatores vêm fundamentar as capacidades que o mercado competitivo (entendido como ambiente acadêmico) exige do GP-CIT. São eles: planejamento de objetivos, métodos, recursos financeiros, físicos, humanos e de conhecimento; agregação de inovação às atividades; encaminhamento adequado dos objetivos a serem atingidos.

Corresponder a tais capacidades demanda, essencialmente, que seja estabelecido um delineamento estratégico para estruturar e coordenar atividades em direção a um horizonte comum identificado como sendo a utilização de recursos e relacionamentos estratégicos para aquisição de vantagem competitiva, acrescida de decisões que agreguem o maior nível possível de certeza e de qualidade.

Considera-se que o delineamento supracitado é efetivado em quatro momentos:

- a) à medida que fatores ambientais internos e externos ao Grupo são diagnosticados;
- b) que a direção estratégica é estabelecida e difundida entre os integrantes;
- c) que problemas estratégicos são identificados e solucionados mediante ações exequíveis no âmbito do Grupo;
- d) que tais soluções são monitoradas por meio de controles específicos.

### Escopo: posicionamento no mercado

Avalia-se que oferecer ao GP-CIT um posicionamento destacado no ambiente onde atua significa definir internamente e esclarecer externamente sua colocação atual no mesmo contexto, de modo que as atividades por ele desenvolvidas atendam uma parcela considerável das demandas de seu universo potencial de usuários, modificando constantemente o desempenho atual de atividades (PORTER, 1999).

Em entrevista junto à coordenação do Grupo, identificou-se o desconhecimento desse posicionamento considerando os limites internos e externos à Universidade. Portanto, o planejamento estratégico teria como escopo fundamental posicionar o Grupo estrategicamente e fomentar sua visibilidade no âmbito da Universidade onde está localizado. Essa restrição de escopo não determina que o Grupo ignore ações em busca de parcerias, clientes e fornecedores externos, visto que o posicionamento é composto também da interação com esses atores. Enquanto espelho de uma estratégia caracterizada como ciclo analítico constituído de planos de ação, o plano estratégico advindo do escopo supracitado orienta o Grupo “ao cumprimento de seus objetivos fundamentais” (HARRISON, 2005, p. 36), identificado conforme as etapas de encaminhamento metodológico, apresentadas a seguir.

## Diagnóstico

### O ambiente de atuação do GP-CIT

Para Harrison (2005, p. 35), o ambiente de atuação da uma organização consiste no contexto sobre o qual essa exerce influências e é, por este, influenciada.

O ambiente onde o GP-CIT atua constitui-se na comunidade científica local (UFPR) e global (fora dos limites da Universidade), que, por suas características peculiares – quando comparada a ambientes de negócio convencionais - apresenta influências e exigências específicas. Contém, inherentemente, o “diálogo e a competição sobre o conhecimento cientificamente válido” (BRASIL, 2009) que, de forma conjuntural, segundo Oliveira (2001, p. 10), apresenta constante movimento e formação contínua, o que impede que aqueles que nela atuem tomem posicionamento estático e inerte.

A análise desse ambiente constitui a primeira etapa metodológica da elaboração do plano estratégico do GP-CIT, que pode ser subdividida em enfoque externo e enfoque interno.

#### Enfoque externo: ambiente geral e forças competitivas

Segundo Chiavenato e Sapiro (2003), esse enfoque viabiliza a identificação de indicadores de tendências, a avaliação do ambiente de negócios, a análise concorrencial e a compreensão dos grupos estratégicos. Transparece, ainda, ameaças e oportunidades ambientais e os meios para contornar ou usufruir desses aspectos.

As forças competitivas (PORTER, 1999) do ambiente externo ao GP-CIT, ditam a rivalidade existente entre o núcleo de concorrentes diretos desse e viabilizam, segundo Harrison (2005, p. 61), a compreensão do posicionamento ideal desse frente a concorrentes atuais e potenciais, aos fornecedores, aos clientes e aos produtos substitutos. Oportunidades e fontes de vantagem competitiva surgem quando tais agentes são

considerados, visto que seu poder de barganha influencia sobremaneira as atividades do Grupo.

Nesses termos, o alto poder de barganha de fornecedores e clientes, enquanto qualificadores de atividades do Grupo, e a menor influência de concorrentes potenciais e produtos substitutos devido à sua presença reduzida no ambiente supracitado, encaminharam à constatação de que:

- a) a rivalidade entre concorrentes diretos no meio acadêmico-científico divide espaço com a colaboração, fator inerente ao próprio meio;
- b) a participação de mercado é tida como aspecto intrínseco às atividades, sem contudo ser estabelecida como conceito estritamente necessário à atuação na academia;
- c) a concorrência direta no meio acadêmico fundamenta-se em:
  - presença em veículos de comunicação científica;
  - recursos físicos e financeiros suficientes e abrangentes;
  - qualificação de integrantes;
  - contato consolidado com fornecedores.

Movimentações inerentes aos aspectos listados acima traduzem, segundo Harrison (2005), a aplicação da criatividade quando da busca em diminuir vantagens competitivas de concorrentes por parte das organizações. No contexto acadêmico, tal dinâmica competitiva reflete-se em fatores como as barreiras inerentes à científicidade quanto à imitação e a flexibilidade em termos de desenvolvimento de pesquisas inovadoras.

Em resposta à vantagem competitiva almejada pelo GP-CIT, fundamentada nos fatores acima, seus concorrentes e demais atuantes no ambiente acadêmico podem agregar forças às suas próprias vantagens (BARNEY e HESTERLY, 2007, p. 82), sustentadas por projetos, pesquisas e formação acadêmica de integrantes diversificados, tanto pela diferenciação do público-alvo, como pela

localização geográfica distinta. Contrariamente, esses atores podem reagir apoiando-se em elementos estratégicos variados em relação a recursos, comunicação, atendimento do público, entre outros fatores, visto que em ambiente de pesquisa científica, a competitividade é alcançada principalmente através de decisões e ações táticas (BARNEY e HESTERLY, 2007, p. 84).

### Enfoque interno: recursos e relações

Esse enfoque, por sua vez, permite a avaliação competitiva de potencialidades e fragilidades da organização em questão, bem como da localização dos recursos que fundamentam suas capacidades e suas competências originárias de valor (CHIAVENATO e SAPIRO, 2003).

Como passos iniciais da análise do enfoque interno à estratégia do Grupo identificaram-se e caracterizaram-se os recursos, as relações e a visão voltada a aspectos endógenos relativos ao seu potencial estratégico. A categorização do valor identificado junto aos recursos originou a definição da competência central do Grupo, insumo para o direcionamento estratégico apresentado ao final do estudo.

Influenciam diretamente as atividades do Grupo os relacionamentos com os seguintes públicos: recursos humanos internos; grupos e pesquisadores parceiros; instituições financeiras e comunidade científica, cujo foco é um conjunto de atividades pertinentes às suas áreas de atuação.

Considerando que relacionamentos internos e externos giram em torno da troca de benefícios entre as partes, no contexto do GP-CIT, tais fatores consistem em pesquisas e produtos desenvolvidos em parceria e pertinentes tanto às linhas de pesquisa integradas pelo Grupo quanto a outras que possam, porventura, criar aplicações multidisciplinares.

Quanto aos recursos mantidos pelo Grupo, aqueles de caráter intangível, humano, físico e financeiro agregam valor à atuação desse no

ambiente acadêmico, consistindo em fontes de competitividade devido à sua peculiaridade e dificuldade de identificação em outros ambientes. Na academia, recursos de quaisquer naturezas especializam-se conforme a atuação dos agentes e, logo, correspondem a certa exclusividade. Percebeu-se que a utilização desses na exploração de oportunidades e na mitigação de ameaças de mercado relaciona-se com a capacidade do Grupo em direcionar competências mantidas por seus integrantes ao desenvolvimento de novas pesquisas e ao gerenciamento do processo criativo e de planejamento. Ainda que tais capacidades possibilitem ao Grupo “aproveitar [...] outros recursos que controla” (BARNEY e HESTERLY, 2007, p. 64), estas carecem de otimização e organização para aproveitamento pleno destes (HARRISON, 2005, p. 79) em direção a um efetivo horizonte estratégico de atuação.

### Análise Swot

A partir dos conteúdos dispostos acima, realizou-se a análise SWOT do GP-CIT, onde aspectos endógenos e exógenos são categorizados quanto às forças, fraquezas, oportunidades e ameaças relacionadas ao Grupo e sua atuação no ambiente acadêmico, e apresentados em ordem de importância no Quadro 1.

Fraquezas e ameaças impedem o encaminhamento pleno do Grupo à vantagem competitiva. Parte da estratégia por ele empreendida deve, então, ser formulada de modo a minimizar ou, preferencialmente, eliminar tais fatores negativos e fomentar forças internas e oportunidades de mercado. Desta forma, estabelecer uma direção estratégica e definir objetivos de longo prazo é condição essencial à efetivação das oportunidades e necessidades acima apontadas, agregando qualidade às decisões tomadas e consolidando a vantagem competitiva do Grupo frente a seus competidores e às ameaças do ambiente.

### Competência central

A competência central do GP-CIT consiste no “conjunto de habilidades e tecnologias [que

Quadro 1 – Forças, fraquezas, oportunidades e ameaças ambientais do GP-CIT

Forças	Oportunidades
Ressaltam capacidades do Grupo e aspectos que viabilizam o planejamento estratégico	Determinam possibilidades de atuação e posicionamento do Grupo
<ul style="list-style-type: none"> <li>a) posicionamento relativamente destacado no Departamento de Ciência e Gestão da Informação e no Setor de Ciências Sociais Aplicadas, consequente horizonte para maior posicionamento na UFPR e fora dela;</li> <li>b) disposição à realização de planejamento estratégico;</li> <li>c) pertinência dos projetos e capacidade de diversificação do portfólio desses;</li> <li>d) formação acadêmica da coordenação do Grupo nas linhas de pesquisa integradas por esse;</li> <li>e) manutenção de contatos e pesquisas conjuntas com grupos parceiros.</li> </ul>	<ul style="list-style-type: none"> <li>a) atualidade e abrangência das linhas de pesquisa integradas pelo Grupo;</li> <li>b) pouca difusão de pesquisas similares na Universidade, proporcionando conjuntura para maior destaque do Grupo em termos de posicionamento;</li> <li>c) concorrência insignificante no âmbito da Universidade e de pouca influência direta no âmbito além dessa Instituição;</li> <li>d) constante disponibilização de editais de financiamento, para obtenção de recursos financeiros e certificação dos projetos.</li> </ul>
<ul style="list-style-type: none"> <li>a) ausência de objetivos de longo prazo, para posicionamento do Grupo;</li> <li>b) dificuldade de planejamento estratégico para alinhamento de atividades a objetivos, e de controle de fatores ambientais e previsão de tendências relativas às linhas de pesquisa;</li> <li>c) ausência de controle das informações sobre o Grupo, seus recursos e relações, e de disponibilização dessas em local específico e acessível a todos os membros;</li> <li>d) posicionamento de menor nível de destaque frente àquele possível, no ambiente de atuação;</li> <li>e) necessidade de tomada de decisões imediatas, sem a devida fundamentação informacional e com alto grau de incerteza.</li> </ul>	<ul style="list-style-type: none"> <li>a) planos políticos, legislações e restrições em nível Federal;</li> <li>b) consolidação de outros grupos de pesquisa com foco semelhante ao do GP-CIT, e recursos e relações diferenciados(as);</li> <li>c) mudanças tecnológicas e descontinuação dos tópicos desse gênero, estudados em projetos do Grupo.</li> </ul>

Fonte: os autores.

possibilitam ao Grupo oferecer benefícios a seus públicos interessados, com os quais o Grupo assume compromisso de criação e aperfeiçoamento” (PRAHALAD; HAMEL, 2005). Conforme os autores, tais habilidades e o uso da tecnologia permitem que essa competência sustente o Grupo em sua estratégia competitiva.

Advinda dos principais recursos mantidos pelo Grupo, a competência central surge do valor global a eles agregado em relação a cinco dimensões:

- a) valor em si;
- b) raridade;

- c) diferenciação frente a outros grupos de pesquisa;
- d) capacidade de organização interna;
- e) exploração pela organização.

O Quadro 2 sistematiza tais dimensões em relação aos recursos do GP-CIT, apontando aqueles de maior relevância para a competitividade e que dão origem a principal competência de suas atividades.

Considerando-se as pesquisas desenvolvidas no âmbito do Grupo e os conhecimentos adquiridos e mantidos por integrantes em relação a essas (tais como foram assinaladas nas colunas do Quadro 2), verifica-se que tais recursos

Quadro 2 – valor global de recursos do GP-CIT

Recurso	Valioso	Raro	Diferenciado	Organizado internamente	Explorado pelo Grupo
Pesquisas desenvolvidas	Amarelo	Azul	Verde	Roxo	Preto
Conhecimentos adquiridos	Amarelo	Azul	Verde	Roxo	Preto
Integrantes do Grupo	Amarelo	Azul	Verde	Roxo	Preto
Certificação por órgãos de fomento	Amarelo	Branco	Branco	Branco	Branco
Montantes financeiros obtidos	Amarelo	Branco	Branco	Roxo	Preto
Instalações físicas	Amarelo	Branco	Branco	Roxo	Preto

Fonte: adaptado de Barney e Hesterly (2007, p. 80).

agregam valor à atuação do GP-CIT no ambiente acadêmico. Portanto, podem ser definidos como estratégicos, ao mesmo tempo em que revelam a competência central a ser ressaltada pelo Grupo em suas atividades, potencializando a obtenção de vantagem competitiva sustentada.

## Direcionamento estratégico

Segundo Harrison (2005, p. 110), a direção estratégica determina a identidade da organização, comunicando-a a públicos internos e externos e orientando processos decisórios em relação à estratégia da mesma organização.

A partir do escopo identificado para o planejamento estratégico do GP-CIT e do diagnóstico supracitado, estabeleceu-se a direção a ser seguida visando atingir o posicionamento almejado. Esse direcionamento reflete-se nos fatores indicados a seguir, cujo nível de abrangência decresce a cada tópico:

### Visão

Ser referência em pesquisas multidisciplinares nas áreas de formação dos integrantes, com a capacidade de gerar e antecipar demandas e temáticas inovadoras relativas a tecnologias e aplicações nessas mesmas áreas de formação.

### Missão

Atuar em pesquisas inovadoras e multidisciplinares, relacionadas à aplicação de conceitos e tecnologias nas áreas de formação

dos integrantes, divulgando os resultados à comunidade científica e oferecendo benefícios também à comunidade em geral.

### Valores

- oferecer qualidade e compromisso no atendimento de demandas por pesquisas;
- promover o apoio da Universidade e de outros apoiadores das atividades do Grupo, em todas as oportunidades pertinentes;
- respeitar valores, idéias e críticas dos integrantes do Grupo, mantendo-o coeso;
- manter idoneidade junto aos órgãos de fomento/financiadores, fortalecendo o relacionamento com tais instituições.

### Objetivos

- gerar pesquisas inovadoras, bem como novos paradigmas de atuação sobre tecnologias e aplicações relativas às linhas de pesquisa;
- buscar e fortalecer contatos e parcerias pertinentes às linhas de pesquisa atendidas;
- estimular a coesão e o comprometimento da equipe;
- efetivar processos organizadamente e com boa estrutura informacional.

Estabelecer conteúdos para visão, missão, valores e objetivos do GP-CIT, contribui para que seu público-alvo compreenda sua atuação

no ambiente acadêmico, incentivando a criação de relações entre o Grupo e potenciais parceiros, uma vez que os objetivos comuns são claramente identificados (HARRISON, 2005). Estratégias podem ser formuladas a partir do direcionamento aqui estabelecido, de modo a manter uma linearidade entre o posicionamento pretendido pelo GP-CIT e os desdobramentos desse em longo, médio e curto prazos, partindo-se do posicionamento atual.

Pode-se dizer que a configuração do ambiente acadêmico apresenta uma estrutura bem definida e não tão flexível em termos estratégicos e, nesse sentido, a identidade do GP-CIT pode ser coordenada e também representada por recursos e relações apontados anteriormente, reforçando o conceito da competência central mantida pelo Grupo.

## Formulação da estratégia

A natureza do ambiente acadêmico tende a ignorar, como já comentado acima, a competitividade nos moldes de empresas e organizações. Modelos de dinâmica competitiva adotados por concorrentes diretos tendem a ser similares, mas ainda assim permitem a formulação de estratégias de atuação para, como aponta Harrison (2005), competir no mercado e adotar posturas condizentes com competências centrais.

Tal pressuposto determina a formulação de estratégias competitivas para o GP-CIT, as quais suportam a atuação desse no ambiente acadêmico considerando horizontes de longo prazo.

Visto que a atuação do GP-CIT não prevê a aquisição simultânea de vantagens competitivas em diversos setores – entendida como estratégia de nível corporativo (BARNEY e HESTERLY, 2007) - as decisões estratégicas relacionadas ao Grupo comportarão os níveis de negócios em um âmbito reduzido de “mercado” e de colaboração.

a) Estratégia no nível de negócio: diferenciação

A abordagem estratégica selecionada para o nível de negócios no GP-CIT é de diferenciação, ressaltando a busca por vantagem competitiva através de aumento ou destaque do valor oferecido por produtos e serviços desenvolvidos em seu âmbito. Enquanto “expressão da criatividade” (BARNEY e HESTERLY, 2007, p. 131-137), essa abordagem busca vantagem competitiva principalmente por meio da ressalva de recursos (HARRISON, 2005).

Constatou-se que a diferenciação almejada pelo Grupo poderia apresentar-se por meio das seguintes dimensões:

- a) características e complexidade dos projetos e pesquisas desenvolvidos, ou seja, o “foco direto nos atributos”, refletidas no conjunto de processos e atividades do Grupo, cuja organização e desenvolvimento únicos e variam em complexidade conforme fontes de pesquisa, métodos utilizados e resultados obtidos. Tais características e processos são difíceis de serem reproduzidos por concorrentes ou substitutos;
- b) localização geográfica do Grupo, considerando o índice mínimo de concorrência direta na Universidade onde esse está inserido e no âmbito de todo o estado;
- c) customização, uma vez que projetos e produtos desenvolvidos no âmbito do GP-CIT têm caráter inédito em relação a escopo, características e complexidade, ou método de desenvolvimento;
- d) reputação, determinante da “relação mais importante entre uma empresa e seus consumidores”, que diferencia o Grupo à medida que sua imagem no mercado é construída única e exclusivamente por suas próprias ações e produtos;
- e) atendimento e suporte, onde a associação entre ação e acompanhamento agrupa valor a determinadas pesquisas conduzidas pelo Grupo e oferece, consequentemente, visibilidade junto ao

mercado (BARNEY e HESTERLY, 2007, p. 133-137).

### b) Estratégia colaborativa

Espaços onde atores “combinam recursos em um esforço para conseguir uma posição mais forte” (HARRISON, 2005, p. 145) e “cooperam no desenvolvimento, na produção ou na venda de produtos ou serviços” (BARNEY e HESTERLY, 2007), constituem as estratégias colaborativas. Nesses termos, o ambiente acadêmico, oferece possibilidades inúmeras em termos de colaboração enquanto meio dinâmico de desenvolvimento de inovações.

Percebeu-se que o GP-CIT tem, como contexto colaborativo, parcerias com outros grupos ou pesquisadores quando da construção conjunta de conhecimentos sobre determinado tema de competência dos parceiros, do compartilhamento desses conhecimentos para desenvolvimento de resultados específicos, do acesso, disponibilidade e uso de meios de comunicação pela comunidade científica e pelos parceiros.

Segundo Barney e Hesterly (2007), o valor dessa estrutura colaborativa se dá pela melhoria de desempenho agregada aos processos do Grupo, considerando não o retorno econômico, mas o prestígio e a visibilidade adquiridos quando da entrega de produtos, serviços, ou da apresentação de resultados a toda a comunidade científica. Tal premissa determina que benefícios advindos de produções colaborativas onde o GP-CIT está inserido são obtidos à medida que os parceiros operam com “compromisso, coordenação e confiança” (BARNEY e HESTERLY, 2007, p. 268).

Conforme o contexto acima explorado, estabeleceu-se que a estratégia colaborativa a ser empreendida pelo GP-CIT deve ressaltar parcerias atuais e buscar novas relações visando um nível ideal de confiança, onde é refletida a capacidade do Grupo em comunicar-se com semelhantes e produzir resultados adequados às demandas. Resultados como produção

bibliográfica, técnica, científica, entre outros, devem também ser almejados ininterruptamente.

Tais objetivos foram formulados de modo a contribuir para o encaminhamento do Grupo ao posicionamento como referência na atuação relativa às suas linhas de pesquisa, considerando a Universidade Federal do Paraná e instituições parceiras como principais mercados alvo.

### c) Estratégia de crescimento

O crescimento do GP-CIT no mercado em que atua deve ser mantido como foco estratégico, traduzido pelo desenvolvimento de pesquisas e produtos em consonância com a direção estabelecida, no longo prazo. No entanto, o período de implementação deve abordar a chamada “estratégia temporária de não-crescimento” (HARRISON, 2005, p. 173), cujo intuito, segundo o autor, é de reforçar as bases estratégicas da organização visando obter maior crescimento futuramente.

## Considerações finais

Ao longo do estudo de cada uma das etapas acima detalhadas, identificou-se que o Grupo de Pesquisa Aplicada em Ciência, Informação e Tecnologia pode melhorar seu desempenho pela implementação das estratégias de negócio e colaborativa.

Tal processo de implementação é subsidiado pelo diagnóstico realizado e no fato de que a cultura do Grupo encontra-se em formação e pode alinhar-se à direção estratégica aqui estabelecida, de modo que a visibilidade e posicionamento de mercado possam ser estimulados principalmente pelo uso de recursos e competências estratégicos. Considera-se ainda que o diagnóstico empreendido é imprescindível ao desenvolvimento de um plano estratégico com foco no ambiente acadêmico, visto que as peculiaridades a pesquisa acadêmica exigem a identificação clara da realidade para que a estratégia torne-se viável e compatível com esse contexto.

Retomando questões culturais do GP-CIT, nota-se que as definições estratégicas estabelecidas a partir do estudo de caso integram empreendedorismo e inovação a uma organização que, cotidianamente, pouco volta esforços à atuação sustentada pela visão além do tático e operacional. Entende-se que, ao se buscar extrapolar esses limites se deve igualmente reconhecer um processo que se agregará paulatinamente às decisões tomadas no ambiente acadêmico, oferecendo-lhe perspectivas sólidas de atividade à medida que competidores e colaboradores têm noção do papel de sua atuação no mercado e do rumo que esse segue conforme pesquisas são desenvolvidas.

Especificamente em relação ao alcance dos objetivos embutidos nas estratégias formuladas, vale ressaltar que, apesar das características diferenciadas do ambiente acadêmico, riscos mostram-se presentes. No que diz respeito ao GP-CIT, há o risco de falha associada à estratégia de diferenciação, cuja concepção pode considerar demasiadamente características e benefícios implícitos de pesquisas desenvolvidas sem que esses sejam efetivamente percebidos pelos públicos-alvo. A atualidade dessas pesquisas também consiste em risco de alto potencial de interferência no posicionamento de mercado por parte do Grupo, visto que oferece espaços de atuação e novos mercados a concorrentes. Esses e outros aspectos interferem, ainda, na estratégia colaborativa estabelecida, uma vez que a confiança entre os parceiros pode ser abalada pela dinâmica competitiva do ambiente.

Ainda que se considerem os riscos supracitados, o plano estratégico formulado para o GP-CIT abarca uma perspectiva holística e integrada, tendo como pressuposto que não apenas ações de implementação da estratégia devem ser empreendidas, mas também a sensibilização para o chamado pensamento estratégico enquanto caráter inovador do Grupo frente à atual cultura da organização.

Entende-se que a inserção de tal pensamento nessa cultura poderá originar a formulação de novos objetivos estratégicos conforme o Grupo avança rumo à direção estabelecida

orientando-se a um posicionamento macro de mercado, à eficácia e efetividade relativas às dimensões informativas, de processos internos, de tecnologias, de aprendizado e crescimento, de clientes e agentes externos e de finanças.

Táticas e operações para implementação das estratégias e contínuo monitoramento de objetivos desse gênero são de responsabilidade do GP-CIT enquanto organização, uma vez que a internalização desses aspectos é fator sumariamente importante para o alcance de visão e missão de negócios a partir do mercado identificado.

## Referências

- BARNEY, J. R.; HESTERLY, W. S. **Administração estratégica e vantagem competitiva**. São Paulo: Pearson Prentice Hall, 2007.
- BRASIL. Ministério da Ciência e Tecnologia. Conselho Nacional de Desenvolvimento Científico e Tecnológico. **Diretório dos grupos de pesquisa no Brasil**. Disponível em: <<http://dgp.cnpq.br/buscaoperacional/detalhegrupo.jsp?grupo=01036072AJV4CL#identificacao>>. Acesso em: 6 abr. 2009.
- CHIAVENATO, I.; SAPIRO, A. **Planejamento estratégico**. Rio de Janeiro: Elsevier, 2003.
- HARRISON, J. S. **Administração estratégica de recursos e relacionamentos**. Porto Alegre: Bookman, 2005.
- OLIVEIRA, M. de. A pesquisa científica na ciência da informação: análise da pesquisa financiada pelo CNPq. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 6, n. 2, p. 143-156, jul./dez. 2001.
- PORTER, M. E. **Competição – on competition: estratégias competitivas essenciais**. 14. Reimp. Rio de Janeiro: Elsevier, 1999.
- PRAHALAD, C. K.; HAMEL, G. **Competindo pelo futuro: estratégias inovadoras para obter o controle do seu setor e criar os mercados de amanhã**. Rio de Janeiro: Elsevier, 2005.

## Strategic planning for the academic environment: a case study of a research group from the Federal University of Paraná

### Abstract

*Introduction: Competition on academic research environments differs from that in the competitive enterprise markets. Strategic objectives in the first case deal with visibility and positioning between researchers and in the research groups market. The Applied Research Group for Science, Information and Technology (GP-CIT) felt the urge for engaging into a strategic planning process, looking for aggregate value to its decision processes, establishing directions and strategies. Internal and external assessments to GP-CIT revealed favorable characteristics towards strategic planning, aiming at reinforcing the potential for its resources, competences and opportunities. Method: Applied research involving the following steps: characterization of the environment, identifying the main type of demand in relation to planning as well as directed the Group focus setting; synthesis of the positioning of the environment outside its market area; diagnosis of internal and external factors; detailment of central expertise; definition of the strategic direction; outlining the strategy in the levels of business collaboration and growth. Results: The characteristics of the internal culture and external GP-CIT are favorable to the realization of strategic planning focused to enhance its resources, skills and opportunities. These fundamental aspects of the design of the strategic direction can be pursued by the Group in the academic environment on long-term vision supporting its mission, goals and business values. Differentiation strategies, collaboration and growth objectives were formulated in order to foster the activity of the GP-CIT in the academic environment, considering its competitive dynamics. Conclusions: Implementation of strategies should follow careful steps and decisions, to achieve competitive prospect based on solid performance. Strategic thinking, while a holistic view, is to be inserted into the Group's culture in order to change not only processes related to it, but the whole paradigm of activities and its competitiveness in the academic environment.*

### Keywords

*Strategic planning; Universidade Federal do Paraná; Research group; Competitiveness; Information Management.*

Recebido em 21 de abril de 2011  
Aceito em 23 de abril de 2011

<sup>1</sup> Graduada em Gestão da Informação, Mestranda em Ciência, Gestão e Tecnologia da Informação - UFPR. Estudante de pós-graduação (Mestrado) - UFPR.  
ac.greef@gmail.com

<sup>2</sup> Graduado em Gestão da Informação, Mestrando em Ciência, Gestão e Tecnologia da Informação - UFPR. Estudante de pós-graduação (Mestrado) - UFPR.  
franciscodaniel@gmail.com

<sup>3</sup> Graduado em Ciências Econômicas, Especialista em Gestão Empresarial, Mestre e Doutor em Administração de Sistemas de Informação - FGV-SP. Professor adjunto - UFPR/DeCiGI.  
ncastilho@ufpr.br

### Como citar este artigo

GREEF, A. C.; COSTA, F. D. de O.; CASTILHO JR, N. C. de C. Plano estratégico para o ambiente acadêmico: estudo de caso em grupo de pesquisa da Universidade Federal do Paraná. **AtoZ**, Curitiba, v. 1, n. 1, p. 50-60, jan./jun. 2011. Disponível em: <[www.atoz.ufpr.br](http://www.atoz.ufpr.br)>. Acesso em:

# Mineração de dados: busca de conhecimento sobre a evolução do canto da família *Thamnophilidae*

Letícia da Costa e Silva<sup>1</sup>, Denise Fukumi Tsunoda<sup>2</sup>, Viviane Deslandes<sup>3</sup>

## Resumo

**Introdução:** Descreve a utilização de uma técnica de mineração de dados sobre o canto, a biologia e o micro-habitat da família *Thamnophilidae* (Aves) a fim de encontrar padrões que os relacionem. **Método:** Uma base de dados foi construída em planilha Excel® relacionando 82 espécies da família da ave *Thamnophilidae* com diversos atributos referentes às características do canto, da biologia e do micro-habitat em que são encontradas. Na análise utilizou-se o algoritmo Apriori no software WEKA 3.7.1. **Resultados:** Ao associar os diferentes atributos de 82 espécies diferentes considerando o suporte mínimo de 10% e a confiança mínima de 90% foram encontrados 172 padrões, dos quais 42 continham um dos atributos do canto: PC1 e PC2. Os padrões que relacionavam o atributo PC2 foram os mais significativos ao indicar a relação deste com o tamanho e gênero da família. Os resultados colaboraram para gerar a hipótese de que os atributos do canto possuem comportamentos não relacionados. **Conclusões:** O experimento demonstrou que o algoritmo pode ser melhor aproveitado em bases de dados maiores e/ou cuja padronização dos dados apresente menor número de categorias, o que pode ser uma limitação no campo da macroecologia. Mas, ao mesmo tempo, se mostrou um instrumento alternativo para o estudo exploratório de relações entre diversos atributos, cujos resultados podem servir de objetos de análises mais aprofundadas.

## Palavras-chave

Mineração de dados; Bases de dados; Aves florestais; *Thamnophilidae* (ave); Cantos dos pássaros.

## Introdução

Em diversas áreas do conhecimento, dados estão sendo coletados e acumulados em grande escala nos meios digitais. Há urgência na descoberta de teorias computacionais e ferramentas que auxiliem os humanos na extração de conhecimentos úteis a partir deste crescente volume de bancos de dados. A Descoberta de Conhecimento em Base de Dados (DCBD), processo criado por Fayyad, Piatetsky-Shapiro, Smyth (1997) surge como uma alternativa viável de atender parte desta demanda, e procura, em um nível abstrato, desenvolver métodos e técnicas que ofereçam significado a dados armazenados digitalmente.

Tudo o que é coletado do nosso ambiente são as básicas evidências usadas para construir teorias e modelos do universo em que vivemos (BAÇÃO e PAINHA, 2003). O computador tem auxiliado

os humanos nesta coleta ao armazenar dados em bancos que crescem continuamente de volume em dois sentidos: no número de registros ou objetos e no número de campos ou atributos relacionados ao objeto; além de permitir o relacionamento entre eles (FAYYAD, PIATETSKY-SHAPIRO E SMYTH, 1997). Dessa forma, o emprego do processo de DCBD, que inclui técnicas de mineração de dados, torna possível a extração de conhecimento a partir da navegação em variados ambientes informacionais.

Diversas ciências vêm utilizando a mineração de dados para o aprimoramento do saber em seus campos, entre elas: a administração, o marketing, a astronomia, a medicina, a física, a ciência da informação. No âmbito das ciências naturais, a astronomia destaca-se no uso de técnicas computacionais para adquirir conhecimento, além do recente interesse na utilização de bancos

de dados compartilhados sobre o genoma humano e de outros animais.

Na biologia e na geografia, a mineração de dados geralmente está associada ao sensoriamento remoto de imagens e no Sistema de Informações Geográficas (SIG). Ferramentas como o SIG, proporcionam aos usuários a visualização de fenômenos complexos que ocorrem na superfície terrestre por meio do armazenamento e geração de grandes quantidades de dados geo-referenciados (BAÇÃO e PAINHA, 2003, p. 6) que facilitam a sua classificação e sumarização. Estudos como o de Li, Di e Li (2000) aplicam algoritmos de mineração de dados sobre saídas de um sistema de informação gerencial e de ferramentas de sensoriamento remoto de dados para melhorar a classificação de imagens do uso de terra.

Recentemente, a macroecologia aplica técnicas de mineração utilizando metadados organizados por diversas pessoas e instituições a fim de descobrir padrões gerais em ecologia. Para Blackburn (2004, p. 1) entender padrões e processos na macroecologia é um desafio metodológico, uma vez que as escalas espaciais e temporais para identificar a distribuição e abundância de espécies são amplas e demandam muitos dados. Conforme Gotelli (2008, p. 4) o uso da mineração de dados na macroecologia (e na biogeografia) pode proporcionar correlações que revelem mecanismos em modelos estatísticos compreensíveis que incluem tantos dados possíveis de serem reunidos. Bekker *et al.* (2008) discorrem sobre o uso da ecoinformática em grandes bancos de dados que reúnem informações fitossociológicas das espécies buscando padrões de sua relação com o ambiente e de sua distribuição em escalas locais e regionais.

Apesar da aplicação da mineração de dados em diversos estudos na área de macroecologia ainda há um promissor cenário na sua exploração de forma a propiciar relevância, utilidade e validade na geração de novos conhecimentos. Este estudo tem o propósito de extraír conhecimento sobre a relação entre parâmetros do canto de aves *Thamnophilidae* com características de sua biologia e seu habitat. Para isso foi utilizada

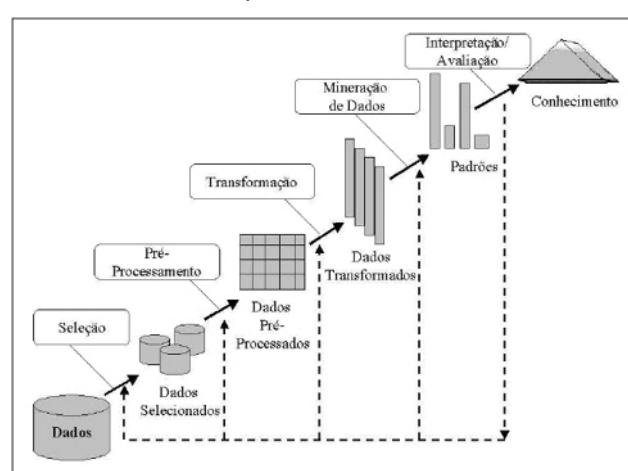
uma base de dados que reúne informações sobre a biologia, o canto e características do micro-habitat dessas aves.

O artigo divide-se em quatro partes. A primeira apresenta uma breve síntese sobre os conceitos de DCBD e mineração de dados, baseados em publicações de Fayyad, Piatetsky-Shapiro e Smyth, e indica os materiais e métodos utilizados a fim de explicitar como este trabalho foi empiricamente sistematizado. A segunda detalha como foi realizado o pré-processamento dos dados antes de submetê-lo ao algoritmo de mineração de dados. A terceira apresenta os resultados gerados pelo algoritmo. E por fim, a quarta, discute os conhecimentos obtidos após a mineração dos dados.

## DCBD e mineração de dados

Fayyad, Piatetsky-Shapiro e Smyth (1997, p.6) apresentam o processo de DCBD em cinco etapas:

Figura 1 – Uma visão geral das fases inclusas no processo de DCBD



Fonte: Neves, 2003, p. 28.

seleção da base de dados, pré-processamento (seleção e tratamento dos dados), transformação, mineração dos dados, interpretação e avaliação, conforme demonstrado na Figura 1.

A DCBD transforma dados de bancos tipicamente volumosos em produtos que facilitam o entendimento das relações entre eles. Vários são os produtos que podem se extraír deste processo como relatórios, modelagem de processos,

ou ainda modelos de previsão. Dessa forma, dependendo do que se pretende extrair da base de dados selecionada é escolhido um algoritmo que minerará os dados em busca da resposta procurada.

A mineração de dados envolve enquadrar modelos a (ou encontrar padrões em) dados observados. A maioria dos métodos de mineração é baseada em técnicas testadas de aprendizado de máquina e reconhecimento de padrões e estatística, como classificação, associação, regressão, *clustering*, summarização, modelagem de dependência, entre outros (FAYYAD, PIATETSKY-SHAPIRO, SMYTH, 1997, p. 8-10).

A escolha apropriada da técnica de mineração a ser utilizada é essencial para que sejam extraídos os conhecimentos pretendidos. Entender como será aplicada a mineração facilita a compreensão do usuário sobre sua contribuição e aplicabilidade na geração de conhecimento a partir de uma grande quantidade de dados.

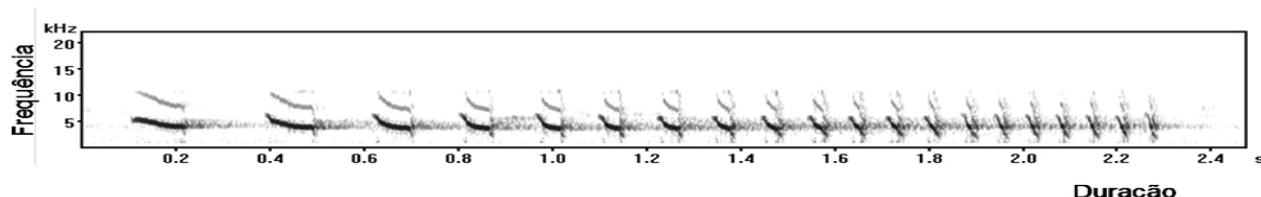
## Materiais e métodos

A base de dados utilizada no estudo relaciona 82 espécies da família da ave *Thamnophilidae* com diversos atributos referentes às características do canto, da biologia e do micro-habitat em que são encontradas. Esta base de dados foi construída em planilha Excel® do pacote Microsoft Office®.

Os cantos das 82 espécies de *Thamnophilidae* foram obtidos da base de dados do XENO-CANTO (2010). Respostas a *playback*<sup>1</sup>, apelos<sup>2</sup> e cantos com taxa de amostragem e resolução menores do que 22.050 Hz e 16 bits foram excluídos da amostra. A fim de evitar pseudo-repetição, foi medido somente um canto por indivíduo e localidade, com um total de 3 a 17 cantos por espécie, totalizando 676 cantos analisados. A partir das gravações foram gerados sonogramas, que são representações gráficas dos cantos, onde a frequência do canto é representada no eixo Y (em Hz) e a duração do canto é representada no eixo x (em segundos), conforme demonstrado na Figura 2.

Todos os sonogramas foram gerados no software RAVEN PRO 1.4 (CHARIF et al., 2010). As seguintes medidas acústicas foram obtidas: frequência maior (Hz), frequência menor (Hz), variação na frequência (Hz), frequência de pico (Hz), duração do canto (s), entropia agregada (s) e número de notas (unidades). Análise de Componentes Principais (PCA) foi aplicada aos parâmetros do canto, visando reduzir sua dimensionalidade. Dessa forma, os valores dos dois primeiros componentes principais (PC1 e PC2) do canto foram usados para representar o canto de cada uma das espécies do estudo, e os dados concernentes a biologia e aspectos do micro-habitat das espécies foram coletados dos trabalhos de Dunning (2008) e de Del Hoyio, Elliot e Smyth (2003).

Figura 2 – Exemplo da representação de um sonograma (ou audioespectrograma) do canto de *Drymophila malura*, um thamnophilídeo de sub-bosque



Fonte: os autores.

Obs.: A frequência (kHz) é representada no eixo y do gráfico e a duração do canto (segundos), no eixo x. Os traços mais escuros dispostos em sequência na figura representam as notas do canto, nesse caso, 20 notas.

<sup>1</sup> É a resposta da ave à gravação de voz de sua própria espécie.

<sup>2</sup> São as vocalizações menos complexas e não territoriais.

O intuito desse trabalho é a detecção de padrões que relacionem a biologia e micro-habitat das espécies aos atributos do canto. Devido à utilização de dois parâmetros do canto para a mineração de dados, optou-se por não utilizar atributo meta na busca de padrões. Dessa forma, a mineração dos dados foi realizada no *software WEKA 3.7.1*, por meio do algoritmo APRIORI (AGRAWAL *et al.*, 1993), instalado em um *notebook* Toshiba Satellite U305-S5127, processador Intel Core 2 Duo T7100 de 1.8GHz, memória 2GB RAM. O APRIORI é um método clássico de mineração de regras de associação baseado no princípio da antimonotonicidade do suporte, ou seja, “um *k-itemset* somente pode ser frequente se todos os seus *(k-1)-itemsets* forem frequentes” (GOLDSCHMIDT e PASSOS, 2005). Este algoritmo é um modelo não-paramétrico que tem por premissa a crença de que as relações que ocorrem de forma consistente no conjunto de dados repetir-se-ão em observações futuras. O método não exige conhecimento profundo do fenômeno a modelar (BAÇÃO e PAINHO, 2003, p. 5), sendo, dessa maneira, oportuno o seu uso na base utilizada neste trabalho.

Devido ao *software WEKA* ler arquivos em arff, foi necessário realizar duas conversões. Primeiramente de xls para csv e depois de csv para arff. A conversão de xls para csv foi realizada pelo comando salvar do Microsoft Excel®. A conversão de csv para arff foi realizada pelo site do pesquisador esloveno Marko Tkalcic (2008), o qual oferece este serviço gratuitamente.

Antes das duas conversões, os espaços e as vírgulas contidas no arquivo xls foram substituídos por “.” (um ponto), uma vez que a leitura em arff entende que os espaços e vírgulas separam um atributo do outro. No site que permite a conversão, os atributos foram selecionados como nominal conforme definido na seção ‘Categorização dos dados’.

O arquivo arff gerado foi salvo e importado para o WEKA 3.7.1 a fim de se aplicar o APRIORI. Para a definição do suporte mínimo fez-se necessário analisar como se daria a primeira etapa dos cálculos do algoritmo em relação aos atributos

do canto (PC1 e PC2), em que há a exclusão de categorias inferiores ao valor definido. Para evitar que fossem consideradas poucas categorias, ou até nenhuma, calculou-se qual era a porcentagem

**Tabela 1 – Porcentagem de ocorrência das categorias dos atributos PC1 e PC2**

PC1	N.º Ocorrências	Proporção rel. ao total	PC2	N.º Ocorrências	Proporção rel. ao total
-4a-2	15	0,183	-4a-2	1	0,012
-2a0	34	0,415	-2a0	24	0,293
0a2	17	0,207	0a2	49	0,598
2a4	11	0,134	2a4	8	0,098
4a6	5	0,061	Total	82	1
Total	82	1			

Fonte: os autores.

de ocorrência (ou seja, o suporte mínimo) das categorias de cada atributo do canto em relação ao total de linhas (82), conforme indica a Tabela 1.

A análise da Tabela 1 aponta que o suporte mínimo de 0,1 garante que um maior número de categorias seja considerado nos cálculos posteriores do algoritmo. Esta escolha possibilita a busca de um maior número de padrões entre os atributos do canto e os outros atributos da base de dados. Com este suporte mínimo somente as categorias 4a6 do PC1, 2a4 e -4a2 do PC2 serão desconsideradas. Aumentar o suporte, para 0,15 ou 0,2, por exemplo, excluiria categorias do PC1 com um número de ocorrência considerável, como 11 e 15 vezes.

Dessa forma, foi estipulado o suporte mínimo de 0,1 (ou 10%), o qual foi também considerado no segundo cálculo do algoritmo, em que verifica a ocorrência das categorias de todos os atributos da base em conjunto. E para o último cálculo, em que é considerado as ordens seqüenciais desses conjuntos acima da confiança mínima definida, foi utilizada a métrica de 0,9 (ou 90%).

Foram encontrados 172 padrões, das quais 42 apresentaram um dos dois atributos do canto (PC1 e PC2) utilizados para a mineração. Os campos vazios do banco de dados foram

preenchidos com 0 (zero) a fim de não serem considerados para análise. A versão 3.7.1 do WEKA oferece, para o algoritmo APRIORI, comandos que permitem a exclusão dos padrões retirados da base importada que relacionam os campos considerados vazios, quando preenchidos por zero. Esses comandos são os “*removeAllMissingCols*” e “*treatZeroAsMissing*”.

## Pré-processamento dos dados

A base de dados continha, inicialmente, informações de 200 espécies distribuídas em 200 linhas. O Quadro 1 apresenta os atributos presentes inicialmente na base e suas respectivas descrições.

**Quadro 1 – Descrição resumida dos atributos originais da tabela**

Atributo	Descrição
<b>Família</b> <i>Thamnophilidae</i>	Esta coluna lista as espécies da família <i>Thamnophilidae</i> que por sua vez são relacionadas aos parâmetros do canto e da biologia das espécies e características de seu micro-habitat.
<b>Parâmetros do canto</b>	PC1 e PC2 correspondem aos dois eixos principais que representam o canto de cada uma das espécies. Esses eixos foram obtidos da Análise de Componentes Principais, realizada a partir da média das seguintes variáveis do canto: frequência maior (Hz), frequência menor (Hz), variação na frequência (Hz), frequência de pico (Hz), duração do canto (s), entropia agregada (s) e número de notas (unidades). O PC1 representa os parâmetros relacionados à frequência e entropia dos cantos, enquanto o PC2 representa os parâmetros temporais do canto: número de notas e duração do canto.
<b>Morfologia do bico</b>	Neste atributo, parte dos dados da biologia das espécies, foi excluído, uma vez que não continha nenhum campo preenchido.
<b>Dados da biologia</b>	Peso, tamanho, micro-habitat, tipo de dieta, modo de forrageio, altitude, movimento, bando misto e dimorfismo de plumagem.

Fonte: os autores.

No atributo “dados da biologia”, o peso (g) e o tamanho (cm) indicados estão relacionados à estrutura física de cada espécie. O micro-habitat aponta características mais específicas da vegetação em que cada espécie é encontrada. A dieta indica a alimentação de cada espécie. O modo de forrageio apresenta o agrupamento das espécies na busca por alimentos. A altitude aponta a elevação do terreno máxima em que cada espécie comumente é encontrada. O movimento indica se a espécie é residente ou possui hábito migratório. O bando misto diz respeito ao comportamento de forrageio da espécie e, por fim, o dimorfismo de plumagem aponta se há diferença de coloração na plumagem entre macho e fêmea de uma mesma espécie.

Os dados referentes à coluna Movimento continham a informação residente para todas as espécies, pois a família não é migratória, dessa forma esta coluna foi excluída. Algumas espécies não possuíam os atributos de canto preenchidos, assim as linhas referentes foram excluídas também. Ao final dessas exclusões a planilha continha 82 linhas (82 espécies) e 15 colunas (1 coluna que relaciona as espécies da família e 14 colunas que relacionam os dados dos atributos).

A categorização dos dados realizada para preparar a base para o processo de mineração de dados está descrito por atributo.

### a) Família

Como não havia espécie repetida, os termos que indicavam a espécie foram suprimidos, restando somente os termos que designam o gênero da família. Por mais que fosse importante permanecer a indicação da espécie, se os dados não fossem padronizados não seria possível encontrar padrões por agrupamento ao utilizar o algoritmo APRIORI.

As espécies foram agrupadas em 32 gêneros diferentes: *Batara*, *Biatus*, *Cercomacra*, *Clytoctantes*, *Cymbilaimus*, *Dichrozona*, *Drymophilla*, *Dysithamnus*, *Epinecrophylla*, *Formicivora*, *Frederickena*, *Gymnocichla*,

*Herpsilochmus, Hypocnemis, Hypocnemoides, Hypoedaleus, Mackenziaena, Microrhopias, Myrmoborus, Myrmochanes, Myrmorchilus, Myrmotherula, Neoctantes, Pygiptila, Pyriglenam, Sakesphorus, Sclateria, Taraba, Terenura, Thamnistes, Thamnomanes, Thamnophilus.*

### b) Parâmetros do canto

PC1 e PC2 – os dados estavam indicados por números exatos<sup>3</sup>. O menor valor encontrado do PC1 foi -3,224 e o maior 4,370. Para se obter maior números de ocorrências iguais, optou-se por padronizá-las em intervalos de 2. Ex: -4 a -2; -2 a 0 e assim por diante. Para o PC2 o menor valor foi -2,120 e o maior 3,454. A padronização deste eixo seguiu o mesmo procedimento do PC1.

### c) Dados da biologia

**Peso** (g) – os dados foram retirados de Dunning (2008) e indicam a média do peso por espécie em número exato. Algumas espécies não possuíam informação, enquanto outras não indicavam a média do peso, apesar de apresentá-lo em intervalo de mínimo e máximo ou pelo sexo (macho e fêmea). Para fins de padronização todos os pesos foram convertidos para o seu valor médio. O menor peso encontrado foi 7 g e o maior 131 g. Não havia concentrações entre o menor e o maior valor deste atributo, assim optou-se por padronizá-los em intervalos de 5 g. Ex.: 6-10; 11-15. As médias com casas decimais que coincidiam com o início ou fim de um intervalo foram incluídas no intervalo padronizado anterior. Ex: 35,8 no intervalo 31-35.

**Tamanho** (cm) - os dados estavam indicados por intervalos e por números exatos. O menor tamanho encontrado foi 6 cm e o maior 28 cm. Não havia grandes concentrações entre o menor e o maior valor deste atributo, assim optou-se por padronizá-los em intervalos de 5 cm. Ex: 6-10; 11-15. Os números exatos com casas decimais contidos no início ou no final de um intervalo padronizado, como 10,5, e intervalos como

9-11, igualmente presentes em dois intervalos subseqüentes, foram enquadrados no intervalo anterior, nestes casos no 6-10.

**Micro-habitat** - como os dados deste atributo estavam descritos por extenso em diversas seqüências, optou-se por padronizá-los em siglas que posteriormente foram distribuídas em cinco diferentes atributos, apresentados a seguir:

Dados relacionados ao estrato da floresta:

- DS – Dossel, Sub-dossel e estrato superior;
- SB – Sub-bosque;
- EM – Estrato médio.

Dados relacionados à floresta em montanha:

- FM – Floresta montana;
- FSM – Floresta sub-montana e de Pé-de-montanha.

Dados relacionados ao ambiente:

- Aberto – caatinga, savana, cerrado, chaco, campo aberto, campinas, clareiras, deserto, mangue, áreas abertas e semi-abertas;
- Fechado - floresta tropical úmida sempre verde, floresta úmida, floresta semi-úmida, floresta árida, floresta seca, floresta de pinus, floresta ripária, floresta de transição, floresta de araucária, floresta decídua, floresta semi-decídua, floresta de terra firme, floresta de galeria, floresta alagada, buritizal, floresta com bambus, floresta temperada, floresta de várzea, floresta de igapó, capões.

Dados relacionados ao status de conservação da floresta:

- VP – vegetação primária;
- VS – vegetação secundária;

<sup>3</sup> Neste trabalho são considerados exatos os valores que não são indicados na forma de intervalos. Exemplo: 566,67 ou 35 são considerados exatos, já 155-655 ou 44,3-53 são considerados intervalos.

- AA - áreas antropizadas como plantações, manejo de madeira, parques e jardins.

Bambu:

- BA - espécies que são fortemente associadas à vegetação com a presença de bambus.

**Dieta** - os dados deste atributo foram padronizados em insetos, artrópodes, pequenos vertebrados, moluscos, sementes, frutos e ovos e filhotes de aves, conforme abaixo:

- Insetos: larvas, pupas e adultos de insetos pertencentes às diversas famílias das ordens *Orthoptera*, (larvas e adultos de) *Lepidoptera*, *Coleoptera*, *Dermaptera*, *Hymenoptera*, *Hemiptera*, *Homoptera*, *Isoptera*, *Diptera*, e *Formicidae*;
- Artrópodes: quilópodes, escorpião, outros artrópodes, aranha, e pseudoescorpião;
- Pequenos Vertebrados: sapo, lagarto, pequenas cobras, lagartixas, lagartos do gênero *Anoles*, e ovos e filhotes de aves;
- Moluscos: lesma e caramujo;
- Sementes: pequenas sementes e sementes em geral;
- Frutos: frutos de qualquer tipo.

**Modo de forrageio** - neste atributo havia informações redundantes contidas em outros atributos. Assim, optou-se por excluir as informações redundantes e padronizar os dados em solitário, par e familiar, conforme abaixo:

- Solitário: solitário;
- Par: em par, em par próximo ao solo e em par na vegetação densa;
- Familiar: pequenos grupos familiares e grupos familiares;

Algumas informações anexas na descrição do modo de forrageio foram excluídas: raramente, frequentemente, no solo, estrato médio.

**Altitude** (m) - os dados estavam indicados por intervalos identificados por números exatos. A menor altitude encontrada foi 0 m e a maior 3050 m. Optou-se por informar a altitude máxima considerando intervalos de 1000 em 1000 metros, obtendo a seguinte padronização para este atributo: 1 (de 0 a 1000m), 2 (de 1001 a 2000m) e 3 (de 2001-3000m).

**Bando Misto** - os dados foram padronizados em N, NA, SO e S, conforme detalhamento a seguir:

- N - não segue bando;
- NA - não há informação, sem informação precisa, sem informação;
- SO - ocasionalmente seguidor de correição, raramente associado, segue raramente, segue às vezes, segue eventualmente, segue ocasionalmente;
- S - segue, segue frequentemente, segue regularmente, segue rotineiramente.

**Dimorfismo Plumagem** - os dados indicados eram: forte, fraco e não. Por sua simplificação, a padronização deste atributo foi considerada como: forte, fraco e N para não.

## Resultados

Os padrões que consideram os atributos do canto estão distribuídos ao longo dos 172 padrões totais recuperados. Deste total, foram encontrados 42 padrões que relacionaram os atributos do canto aos outros atributos da base. O número de ocorrência em conjunto dos atributos de cada padrão variou entre 8 a 13 vezes das 82 possíveis, conforme indicado na Tabela 2.

Foram recuperados 32 padrões relacionados ao PC2 e dez ao PC1. Contudo, esses atributos não ocorrem em conjunto em nenhum dos padrões. As categorias do canto presentes são -2 a 0 do atributo PC2, em especial, e as -4 a -2, 0 a 2 e 2 a 4. A Tabela 2 resume as informações encontradas em 24 dos 42 padrões recuperados de forma decrescente: do padrão com maior confiança (100%) para o de menor (90%).

**Tabela 2 – Padrões encontrados pelo APRIORI**

Se	Então	Confiança
PC2=-2a0 e Ambiente=A.F e Dieta=artropodes.insetos	Tam(cm)=11--15	13/13 ou 100%
PC2=-2a0 e Dieta=artropodes.insetos e Alt(m)=2	Tam(cm)=11--15	12/12 ou 100%
PC2=-2a0 e Peso(g)=16-20	Tam(cm)=11--15	10/10 ou 100%
PC2=-2a0 e Peso(g)=6-10	MF=solitario.par.familiar	9/9 ou 100%
PC2=-2a0 e Peso(g)=16-20 e Dieta=artropodes.insetos	Tam(cm)=11--15	9/9 ou 100%
PC1=0a2 e Estrato=SB	Bambu=0	8/8 ou 100%
PC1=2a4 e MF=solitario.par.familiar	Dieta=artropodes.insetos	8/8 ou 100%
PC1=-4a-2 e Tam(cm)=11-15 e Bambu=0	Genero=Thamnophilus	8/8 ou 100%
PC1=-4a-2 e Tam(cm)=11-15 e MF=solitario.par	Genero=Thamnophilus	8/8 ou 100%
Genero=Thamnophilus e PC1=-4a-2 e Alt(m)=2	MF=solitario.par	8/8 ou 100%
PC2=-2a0 e Peso(g)=6-10 e Dieta=artropodes.insetos	MF=solitario.par.familiar	8/8 ou 100%
PC2=-2a0 e MF=solitario.par.familiar e DP=2	Bambu=0	8/8 ou 100%
PC2=-2a0 e StConservacao=VS e Dieta=artropodes.insetos	Tam(cm)=11--15	12/11 ou 92%
PC2=-2a0 e Dieta=artropodes.insetos e BM=SO	Tam(cm)=11--15	12/11 ou 92%
PC1=2a4	Dieta=artropodes.insetos	11/10 ou 91%
Genero=Thamnophilus e PC1=-4a-2	MF=solitario.par	11/10 ou 91%
PC1=-4a-2 e Ambiente=A.F	MF=solitario.par	11/10 ou 91%
PC1=0a2 e Dieta=artropodes.insetos	MF=solitario.par.familiar	11/10 ou 91%
PC2=-2a0 e Tam(cm)=11-15 e DP=2	Bambu=0	11/10 ou 91%
PC1=-4a-2 e Tam(cm)=11-15	Genero=Thamnophilus	10/9 ou 90%
PC2=-2a0 e Peso(g)=16-20	Dieta=artropodes.insetos	10/9 ou 90%
StConservacao= VS e Alt(m)=1	PC2=-2a0	10/9 ou 90%
PC2=-2a0 e Peso(g)=16-20 e Tam(cm)=11-15	Dieta=artropodes.insetos	10/9 ou 90%
PC2=-2a0 e Peso(g)=16-20	Tam(cm)=11-15 e Dieta=artropodes.insetos	10/9 ou 90%

Fonte: os autores.

Tomando como exemplo o primeiro padrão da Tabela 2, o resultado deve ser interpretado da seguinte forma: se PC2 = -2a0, o ambiente= A.F (aberto e fechado) e a dieta= artrópodes.insetos, então o tamanho deverá ser 11-15 (entre 11 e 15 cm) com 100% de confiança. Considerando que os atributos PC2, ambiente e dieta ocorreram 13 vezes juntos, e nessas 13 vezes, o atributo tamanho foi encontrado em todas essas ocorrências.

Dos padrões com 100% de confiança encontram-se as razões de ocorrência em conjunto dos atributos: 13/13, 12/12, 10/10, 9/9 e 8/8; com 92% a razão 12/11; com 91% a razão 10/11; e com 90% a razão 10/9.

Vale ressaltar que os padrões que apresentaram o atributo Bambu estão todos indicando a

informação 0 (zero). Isto se explica pelo fato do atributo ter sido categorizado em duas opções somente: 0 ou BA e, destes, 61 serem 0. Apesar da baixa ocorrência do BA, esta categoria foi incluída na segunda etapa de cálculos do algoritmo por ter confiança mínima de 0,25. Contudo, não se apresentou suficientemente em conjunto com outros atributos para ter a confiança mínima de 0,9. Ao contrário da opção 0, a qual está presente em grande parte dos padrões encontrados, apesar de terem sido utilizados os comandos “removeAllMissingCols” e “treatZeroAsMissing”.

## Considerações finais

A grande porcentagem de ocorrência de algumas categorias do canto não garantiu sua recuperação nos 42 padrões encontrados, apesar do baixo

suporte mínimo (0,1) escolhido. Tanto a categoria -2 a 0 do atributo PC1 (41%) e a 0 a 2 do atributo PC2 (59%) são relacionadas de forma muito variável aos atributos de biologia e de micro-habitat das espécies.

Devidos a todos os padrões encontrados relacionando o atributo bambu terem apresentado informação zero, este atributo deve ser desconsiderado nas interpretações. A inclusão deste atributo foi feita com o intuito de estabelecer a relação da presença de bambu no ambiente com algum atributo do canto.

Apesar de o atributo PC2 ter ocorrido 32 vezes no total dos padrões obtidos, as relações geradas não foram as comumente esperadas ao considerar as hipóteses de evolução do canto. Entretanto, os 10 padrões recuperados para o PC1 permitem, em parte, essa associação. Esse atributo do canto representa os parâmetros de frequência e entropia, e sua associação com a segunda menor classe de tamanho (11-15 cm) e o gênero *Thamnophilus* sugerem um aprofundamento no estudo dessas relações.

A forma da padronização dos dados da base e o seu tamanho, em linhas, não permitiram que fossem gerados padrões conclusivos ou que indicassem tendências marcantes em relação à evolução do canto da família *Thamnophilidae*. Porém, a padronização dos dados não pôde ser reduzida para menos categorias, caso contrário haveria perda significativa de informações dos atributos, e a amostra utilizada relaciona já um número considerável de espécies da família.

Entretanto, o APRIORI se mostrou útil como uma ferramenta exploratória, pois foi possível identificar indícios de interessantes relações entre os atributos do canto e os da biologia e micro-habitat para uso em novos estudos. Além disto, foi também verificado que há uma possível tendência no direcionamento distinto dos atributos do canto PC1 e PC2, uma vez que não foram encontrados em conjunto em nenhum dos padrões resgatados.

Dessa forma, apresenta-se como sugestão a aplicação do APRIORI em base de dados com

maior número de linhas e que seja possível a padronização em um menor número de categorias, recomendando-se a sua aplicação, na macroecologia, para estudos exploratórios, a fim de aprofundar os estudos das relações entre atributos ao utilizar outros algoritmos, como os de análise multivariada.

## Referências

- AGRAWAL, R.; IMIELINSKI, T.; SWAMI, A. Mining association rules between sets of items in large databases. ACM SIGMOD Conference Management of Data, 1993. **Proceedings...** Disponível em: <<http://www.rakesh.agrawal-family.com/papers/sigmod93assoc.pdf>>. Acesso em: 10 nov. 2010.
- BAÇÃO, F.; PAINHO, M. Aspectos metodológicos da utilização do data mining no âmbito da geografia. *Finisterra*, v. 38, n. 75, p. 135-147, 2003.
- BEGON, M.; TOWNSEND, C. R.; HARPER, J. L. **Ecology: from individuals to ecosystems.** 4 ed. Oxford: Blackwell, 2006.
- BEKKER, R. M. et al. Long term datasets: from descriptive to predictive data using ecoinformatics. *Journal of Vegetation Science*, v. 18, n. 4, p. 457-462, 2007. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.1654-1103.2007.tb02559.x/abstract>>. Acesso em: 10 nov. 2010.
- BLACKBURN, T. M. **Method in macroecology.** Disponível em: <<http://wolfweb.unr.edu/~ldyer/classes/blackburn.pdf>>. Acesso em: 20 abr. 2011.
- CHARIE, R. A.; WAACK, A. M.; STRICKMAN, L. M. **Raven Pro 1.4 user's manual.** Cornell Lab of Ornithology, Ithaca, NY. 2010. Disponível em: <<http://www.birds.cornell.edu/brp/raven/RavenOverview.html>>. Acesso em: 14 set. 2010.
- DEL HOYIO, J.; ELLIOT, A.; CHRISTIE, D. A (Ed.). **handbook of the birds of the world.** 8 v. (Broadbills to Tapaculos). Barcelona: Lynx, 2003.
- DUNNING, J. B (Ed.). **CRC handbook of avian body masses.** 2 ed. Boca Raton: CRC, 2008.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. **Knowledge discovery and data mining:** towards a unifying framework. 1996. Disponível em: <<http://www.aaai.org/home.html>>. Acesso em: 10 out. 2010.

\_\_\_\_\_. **From data mining to knowledge discovery in databases.** 1997. Disponível em: <<http://www.kdnuggets.com/gpssubs/aimag-kdd-overview-1996-Fayyad.pdf>>. Acesso em: 10 out. 2010.

GLUSMAN, G. et al. The olfactory receptor gene superfamily: data mining, classification and nomenclature. **Mammalian Genome**, NY, 2000.

GOLDSCHMIDT, R.; PASSOS, E. **Data mining**: um guia prático. Rio de Janeiro: Elsevier, 2005.

GOTELLI, N. J. Perspectives in biogeography: hypothesis testing, curve fitting, and data mining in macroecology. **International Biogeography Society Newsletter**, v. 6, n. 3, p. 1-7, 2008.

LI, D.; DI, K.; LI, D. Land use classification of remote sensing image with GIS data based on spatial data mining techniques. **Archives of Photogrammetry and Remote Sensing**, Amsterdam, v. 33, parte b3, 2000.

NEVES, R. de C. D das. **Pré-processamento no processo de descoberta de conhecimento em banco de dados**. 2003. Dissertação (Programa de Pós-graduação em Computação) - Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.

TKALCIC, M. **csv2arff**. 2008. Disponível em: <<http://slavnik.fe.uni-lj.si/markot/csv2arff/csv 2arff.php?do=instructions>>. Acesso em: 11 abr. 2011.

XENO-CANTO. 2005-2010. Disponível em: <<http://www.xeno-canto.org/>>. Acesso em: 20 abr 2010.

## Data mining: the search for knowledge about the singing evolution of *Thamnophilidae* family

### Abstract

*Introduction:* Describes the use of a data mining technique about the song, biology and micro-habitat of the *Thamnophilidae* bird family in order to find patterns which relate them. *Method:* A database was built in Excel ® spreadsheet listing 82 species of the family of the bird *Thamnophilidae* comprising various attributes related to bird calling features, biology and micro-habitat in which they are found. For the analysis it was used the algorithm *APRIORI* in the *WEKA* 3.7.1 software.. *Results:* The association of the different attributes of the 82 different species, considering 10% of minimum support and 90% of minimum confidence, allowed the rescue of 172 patterns, from which 42 contained one of the song's attributes: PC1 e PC2. The patterns which related the attribute PC2 were the most expressive ones due to its relation to the size and gender of the family. *Conclusions:* The experiment demonstrated that the algorithm could be better suited in larger databases and/or when the data standardization presents a lower number of categories, what could be a limitation in the macroecology

field. Nonetheless, it has presented itself as an alternative instrument to the exploratory study of the relations among diverse attributes, which results could serve as objects for further analysis.

### Keywords

*Datamining; Database; Forest birds; Thamnophilidae (bird); Bird songs.*

Recebido em 9 de maio de 2011

Aceito em 12 de maio de 2011

<sup>1</sup> Graduada em Administração - UEA, Mestre em Ciência, Gestão e Tecnologia da Informação - UFPR. Estudante de pós-graduação (Mestrado) - Bolsista PROF - UFPR.  
leticia.csilva01@gmail.com

<sup>2</sup> Bacharel em Informática - UFPR, Mestre em Informática Industrial, Doutora em Engenharia Biomédica. Professor adjunto - UFPR/DeCiGI.  
dtsunoda@ufpr.br

<sup>3</sup> Graduada em Ciências Biológicas, Mestre em Ecologia (INPA), Doutoranda pelo Programa de Pós-graduação em Ecologia e Conservação (UFPR). Estudante de pós-graduação (Doutorado) - UFPR.  
viviane.deslandes@gmail.com

Como citar este artigo:

COSTA E SILVA, L. da; TSUNODA, D. F.; DESLANDES, V. Mineração de dados: busca de conhecimento sobre a evolução do canto da família *Thamnophilidae*. **AtoZ**, Curitiba, v. 1, n. 1, p. 61-70, jan./jun. 2011. Disponível em: <[www.atoz.ufpr.br](http://www.atoz.ufpr.br)>. Acesso em: