

Tycho – Um Sistema para Apoio à Avaliação e Gestão Institucional da USP.

RESUMO

Apresenta-se neste documento a proposta de um sistema, denominado Tycho, para apoiar a avaliação e a gestão institucional da USP. Esse sistema deverá apoiar:

- as atividades de elaboração de relatórios pelos departamentos/unidades da USP;
- a avaliação institucional da USP na perspectiva da CPA (Comissão Permanente de Avaliação) e de outros órgãos centrais da USP;
- a avaliação individual de docentes na perspectiva dos departamentos, unidades e órgãos centrais da USP;
- dar subsídios para a gestão acadêmica e planejamento institucional; e
- a comunicação interna e externa das atividades e resultados da USP .

A coleta de informações será fortemente calcada na base de dados corporativos existentes no Departamento de Informática da Reitoria (DI), no Currículo Lattes do CNPq, eventualmente e, restrita apenas para condições efetivamente necessárias, em coleta de dados adicionais por meio de preenchimento de questionários. O princípio básico de concepção do Sistema Tycho é a integração das atividades de coleta de dados, avaliação, diagnóstico e planejamento que já existem na USP, porém ainda de forma isolada. O sistema deverá integrar os indicadores de forma multidimensional, considerando as metas da Universidade e os indicadores oriundos dos departamentos, das Pró-Reitorias, da CPA e ainda, da perspectiva de programas institucionais, a exemplo da meta de internacionalização das atividades e resultados da USP, de forma a facilitar o uso dos indicadores sob diversos interesses e segmentos da Universidade: avaliação das atividades-fim, gestão acadêmica; concessão de claros; concessão de recursos financeiros; indução de ações; planejamento da infra-estrutura; avaliação do impacto social, entre outros. Por outro lado, como as atividades desenvolvidas na Universidade estão intimamente relacionadas com aquelas desenvolvidas em instituições irmãs, tanto

as nacionais quanto as de fora do país, o sistema deverá contar também com recursos para coletar dados de fontes externas à universidade, de forma a viabilizar a comparação de indicadores entre universidades. O sistema será desenvolvido de forma incremental e evolutiva com base na chamada Representação Paramétrica Multidimensional (RPM) de Objetos, na filosofia de software livre, podendo contribuir para a gestão acadêmica de outras universidades, além da USP. Portanto, a integração desse sistema com os sistemas da CAPES, CNPq e FAPESP entre outros é também desejável.

INTRODUÇÃO

O Regimento Geral da USP define que “A Avaliação Institucional será coordenada por Comissão Permanente de Avaliação (CPA), a quem compete propor, ao Conselho Universitário (Co), diretrizes para essa finalidade, bem como fornecer a ele e ao Reitor análises qualitativas e quantitativas sobre o desempenho da Universidade, no que se refere às atividades-fim” [Res. 4928]. Cabe também a organização de Base de Dados para facilitar o processo de avaliação. Por outro lado, a Comissão Especial de Regimes de Trabalho (CERT), prevista no inciso XI do artigo 34 do Estatuto da USP, é a comissão supervisora e fiscalizadora dos regimes de trabalho do corpo docente da Universidade de São Paulo e também necessita de indicadores do corpo docente para nortear as suas decisões e recomendações. Mais recentemente, foi criada a Comissão de Planejamento Institucional, a qual, com base em diagnósticos a partir de indicadores bem estabelecidos deverá cuidar do planejamento institucional, estabelecendo ações, metas e indicadores para a gestão efetiva da USP. Desse cenário, também participam as Pró-Reitorias e outros órgãos centrais da USP, que estabelecem indicadores, diagnósticos, projetos e planejamentos dentro da perspectiva de suas dimensões de atuação: graduação, pós-graduação, pesquisa e extensão.

O procedimento tradicionalmente empregado no processo de elaboração de relatórios de avaliação consiste na solicitação de informações aos departamentos acadêmicos e aos docentes e a extração ad-hoc de dados da base de dados dos sistemas corporativos da instituição. Há uma dispersão de esforço para coleta, processamento e uso de dados em praticamente todos os departamentos e unidades, assim como nos órgãos centrais, com superposição de iniciativas não integradas, limitando possibilidades de relatórios gerencias amplos, completos e integrados, tornando todo o sistema ineficiente, inefetivo e oneroso, além de sujeito a erros.

Esse cenário motiva a proposição de um sistema para apoiar a avaliação e a gestão institucional da USP. Propomos aqui a criação de um sistema, denominado **Tycho**, que permitirá a obtenção, de maneira automática, de dados relevantes para o processo de avaliação acadêmica, quase sem a necessidade de preenchimento de questionários pelos departamentos acadêmicos e/ou pelos docentes.

O nome proposto homenageia o nobre dinamarquês Tycho Brahe, astrônomo, astrólogo e alquimista, que “percebeu que o progresso da ciência da astronomia poderia ser conseguido não com observações aleatórias e ocasionais, mas apenas através de sistemática e rigorosa observação, noite após noite e com o uso de instrumentos com a maior acuidade possível.” [Wikipedia2007]. Seus dados acumulados permitiram que seu assistente, Johannes Kepler, deduzisse as leis do movimento planetário. Como inspiração, é oportuno citar uma observação atribuída a Tycho Brahe, que teria dito em 1563, aos 17 anos, “Estudei todas as cartas disponíveis de planetas e estrelas, e nenhuma delas confere com as outras. Existem tantas medições e métodos quanto astrônomos e todos eles estão em desacordo. O que é necessário é um projeto de longo prazo, com a meta de mapear os céus, conduzido a partir de uma única localidade e por um período de muito anos.” Como diversos dos sistemas centrais que apóiam as atividades diárias da USP têm nomes de planetas, um sistema que os observa e coleta dados de maneira “sistemática e rigorosa” lembra o astrônomo que permitiu o entendimento da mecânica celeste.

O princípio básico de concepção do Sistema Tycho é a integração das atividades de coleta de dados, avaliação, diagnóstico e planejamento que em boa parte já existem na USP, porém ainda de forma isolada. O sistema deverá integrar os indicadores de forma multidimensional, considerando as metas da Universidade e os indicadores oriundos dos departamentos, das Pró-Reitorias, da CPA e ainda, a perspectiva de programas institucionais, a exemplo da internacionalização das atividades e resultados da USP, de forma a facilitar o uso desses indicadores sob diversos interesses e segmentos da Universidade: avaliação das atividades-fim, gestão acadêmica; concessão de claros; concessão de recursos financeiros; indução de ações; planejamento da infra-estrutura; avaliação do impacto social, entre outros. Por outro lado, como as atividades desenvolvidas na universidade estão intimamente relacionadas com aquelas desenvolvidas em instituições irmãs, tanto as nacionais quanto as do exterior, o sistema proposto deverá também coletar indicadores que possibilitem o diagnóstico do índice de internacionalização das atividades e resultados da USP, assim como indicadores de outras perspectivas de avaliação, como, por exemplo, os de impacto social.

A geração de novos indicadores correlacionais e integradores inclui-se também entre as metas previstas. Um exemplo de possibilidade é mostrado pelo estudo da relação entre o impacto da colaboração internacional e o papel das redes de colaboração (Packer e Meneghini) e a importância de avaliação de programas (Pian e Meneghini).

A premissa é que os indicadores relevantes da produção em relação às suas atividades-fim podem ser obtidos das bases de dados corporativos, uma vez que as atividades cotidianas da instituição devem estar documentadas nessas bases, utilizadas principalmente para o controle e acompanhamento dos fluxos de trabalho da instituição, bem como de bases de dados de produção científica, a exemplo da base de Currículos Lattes do CNPq ou de bases de publicações como o Web of Science e Scielo, onde a produção acadêmica dos indivíduos que compõem a instituição já é declarada em função de outras finalidades, como a obtenção de bolsas ou auxílios e dados da Capes. O Currículo Lattes e as Bases de Dados dos Sistemas Corporativos da USP constituirão elementos centrais para os mais diversos relatórios institucionais. O Currículo Lattes deverá então ser institucionalizado incorporando os indicadores sinalizados pela USP.

O objetivo deste projeto é criar um sistema estável e evolutivo, que permita a amostragem regular de elementos e indicadores bem definidos a partir das bases de dados corporativos e de bases externas pertinentes e confiáveis. É fundamental ressaltar que sistemas dessa complexidade necessitam de uma equipe estável para a sua concepção, desenvolvimento e manutenção. Deve-se ressaltar também, que um sistema desse gênero, que possibilite a avaliação, gestão e planejamento institucional abrirá possibilidades de pesquisa, em termos de mestrado e doutorado, de forma que a equipe poderá ser motivada indicando-se a possibilidade desse sistema constituir estudo de caso para o desenvolvimento dos trabalhos científicos, além da implementação e evolução do sistema em si.

PROPOSTA

Equipe de Coordenação, Concepção Preliminar e Equipe de Desenvolvimento/Manutenção

Equipe de Coordenação

Em 2007 a CPA organizou o “*I Workshop sobre Sistemas Computacionais de Apoio à Avaliação Institucional da USP*”, no dia 05 de setembro de 2007, coordenado pelos Profs Lajolo e Maldonado, no Prédio da Reitoria (programa anexo), com participação do pessoal do DI. Nesse evento, identificaram-se diversas iniciativas em andamento na USP e em outras instituições, entre elas a UFMG e a UNICAMP. A partir desse evento identificou-se uma equipe para a gestão do Sistema Tycho, que vem atuando desde então visando seu projeto e implantação:

- Prof. Dr. José Carlos Maldonado – ICMC (coordenador geral - representante CPA)
- Silvio de Paula – DI (coordenador técnico)

- Luiz Moreira Gomes – DI
- Prof. Dr. Caetano Traina Jr – ICMC
- Profa. Dra. Cristina Dutra de Aguiar Ciferri – ICMC
- Prof. Dr. Roberto Marcondes Cesar Junior – IME
- Edna Maria Brazolim – GVR-CPA

Nas fases posteriores de especificação e implementação do Sistema Tycho, outros pesquisadores deverão ser agregados à equipe, provenientes, entre outras, das seguintes áreas: interface usuário-computador, sistemas *web*, estatística, engenharia de software, bases de dados, visualização de informação.

Concepção Preliminar

A proposta de criação do Sistema Tycho é realizar a implementação com um protótipo evolutivo, que será desenvolvido incrementalmente, e contribuir para a consolidação de um Modelo de Avaliação e Gestão Institucional para as Unidades, a ser utilizado pelos órgãos competentes da USP. O conceito que norteia a concepção desse sistema é a chamada Representação Paramétrica Multidimensional (RPM) de Objetos Complexos, uma técnica que permite estabelecer métricas relativas e absolutas para operações de comparação e classificação de objetos, cuja complexidade impede que essas operações sejam executadas diretamente. No caso desta proposta, os “objetos complexos” serão as Unidades de Ensino, Departamentos, Programas de Pós-graduação e quaisquer outras unidades organizacionais que se pretenda avaliar e administrar, como exemplo o próprio docente.

A RPM é baseada na definição de uma coleção de valores, a que chamamos “**características**”, que podem ser medidas sobre cada objeto complexo. Todos os objetos de um mesmo tipo têm um valor associado a cada característica de seu tipo. Assim, cada objeto passa a ser representado por um “**vetor de características**”, e todas as operações

de classificação, comparação, etc. entre os objetos passam a ser executadas de fato sobre seus vetores de características. A definição das características que compõem cada vetor corresponde ao “**esquema**” dos objetos daquele tipo. A comparação entre dois vetores de características é feita por uma função de comparação, a que chamamos “**métrica**”. Uma métrica realiza a comparação seguindo um determinado aspecto. Assim, sobre cada esquema, podem ser definidas diversas métricas, cada uma avaliando determinado aspecto.

O Sistema Tycho adotará como características dados que podem ser obtidos de bases, já existentes ou não. As fontes dos dados poderão ser tanto institucionais, como as bases dos sistemas corporativos Júpiter, Janus, Marte, etc., quanto externas à USP, como a coleção de currículo Lattes dos docentes da USP, a base de dados do Sistema de Grupos de Pesquisa do CNPq, FAPESP, CAPES, etc.. Dados que não estejam disponíveis em nenhuma base, eventualmente poderão ser obtidos utilizando-se módulos específicos para coleta de dados que venham a ser desenvolvidos como parte do projeto. Essa opção, no entanto, será utilizada apenas quando não houver maneira de se obter automaticamente os dados desejados, uma vez que consideramos fundamental minimizar ao máximo a necessidade das unidades/pesquisadores precisarem fornecer dados para o sistema.

A partir dos dados obtidos, o sistema deverá gerar as características que descrevem cada objeto, as quais, por sua vez, irão gerar os “**indicadores**” que sejam importantes para o processo de avaliação institucional. Consideramos que do ponto de vista das atividades-fim da Universidade, existem quatro conjuntos de indicadores que deverão ser supridos: Indicadores de Desempenho da Graduação, da Pós-graduação, de Pesquisa e de Extensão. Cada indicador é de fato constituído por um subconjunto do vetor de características, não necessariamente disjunto dos demais. Dentro da técnica RMP, cada indicador é de fato uma projeção do vetor de características nas dimensões de interesse. Dispondo dos indicadores já medidos, pode-se então utilizar ferramentas de apoio à tomada de decisão, que auxiliam as tarefas de avaliação e gestão institucionais. Por outro lado, como as atividades desenvolvidas na universidade estão intimamente relacionadas com aquelas desenvolvidas em instituições irmãs, tanto nacionais quanto de fora do país, o sistema

deverá contar também com recursos para coletar dados de fontes externas à universidade, de forma a viabilizar a comparação de indicadores entre as universidades.

A Figura 1 mostra a arquitetura geral do sistema previsto. As diversas fontes de dados são avaliadas periodicamente, e a partir dos dados brutos que elas disponibilizam, são extraídas as características importantes, gerando a coleção de vetores de características, os quais coletivamente formam o Armazém de dados do sistema (*data warehouse*). Os extratores de características filtram, compõem e consolidam os dados brutos – por exemplo, filtram os dados dos períodos de interesse, compõem informações individuais nas características que descrevem coletivamente os objetos complexos, e consolidam os dados gerados com os demais já disponíveis no Armazém. Nos períodos em que decisões devem ser tomadas a partir dos indicadores, as métricas são aplicadas sobre os dados disponíveis no Armazém para gerar os indicadores do período de interesse. Finalmente, sobre os indicadores gerados, ferramentas de apoio à decisão são utilizadas pelos avaliadores/administradores.

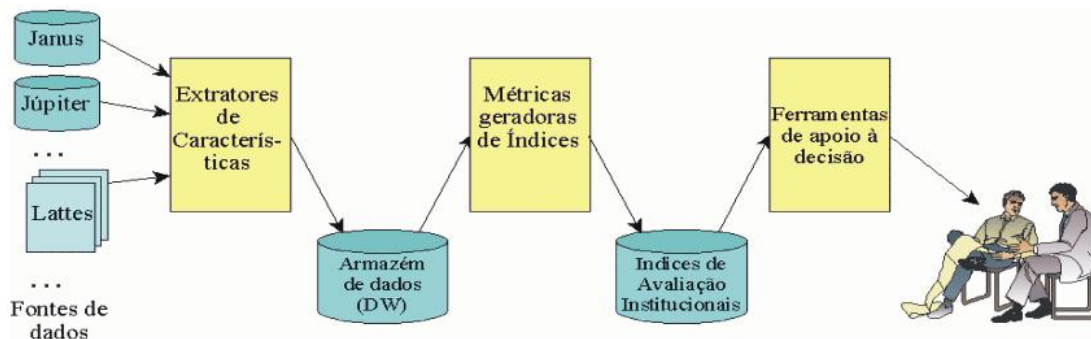


Figura 1 - Arquitetura geral do Sistema Tycho

Conforme mencionado anteriormente, o sistema deverá viabilizar a análise comparativa com outras universidades nacionais e internacionais, e deverá, portanto,

viabilizar a coleta de dados de fontes externas, conforme ilustrado na Figura 2. A título de exemplo, um aspecto importante e meta da atual gestão é a internacionalização das unidades, nos seus diversos níveis de atuação. A internacionalização visa comparar os indicadores obtidos no processo de avaliação institucional com indicadores equivalentes de instituições internacionais de renome. A partir desta comparação, será possível identificar metas a serem seguidas para melhorar a projeção internacional da USP. Essa preocupação é fundamental, pois a qualidade do ensino deve ser universal, respeitados, evidentemente, os aspectos culturais de cada nação.

Na pós-graduação, por exemplo, indicadores de internacionalização constituem condição necessária para que os cursos de pós-graduação possam receber conceito 6 ou 7. Entre esses indicadores estão os relativos à produção técnico-científica das áreas. Nessa perspectiva, o Comitê da Área de Computação da CAPES, com forte colaboração de pesquisadores da UFMG, realizou um estudo do perfil de publicações dos pesquisadores dos melhores programas do exterior e procederam a uma análise da inserção internacional dos programas 5, 6, e 7 do Brasil, de forma a embasar a recomendação daquele comitê ao CTC/CAPES (vide Anexo). Esse estudo, e sistemas de apoio associados, serão generalizados de forma a obterem automaticamente o perfil de publicações em nível internacional, em todas as áreas de conhecimento da USP. Isto contribuirá para o diagnóstico do nível de internacionalização dos programas de pós-graduação da USP, de acordo com esses indicadores, e para o planejamento de ações que visem o aprimoramento das atividades da Universidade nesse aspecto.

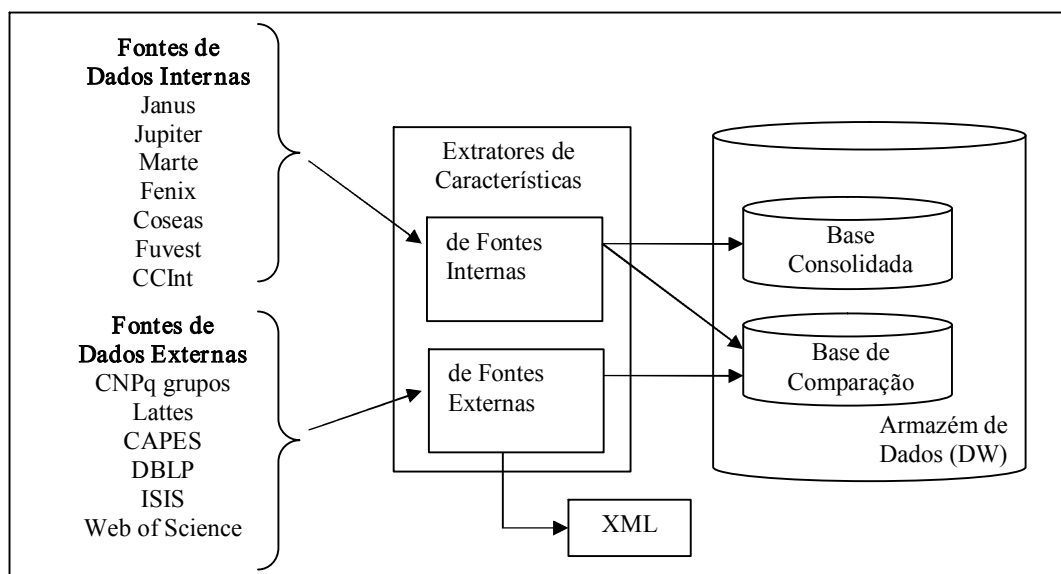


Figura 2. Detalhamento dos Extratores de Características

Na Figura 2, as Fontes de Dados Internas são os sistemas administrativos da USP ou então fontes que possuem dados consolidados. Estas fontes de dados oferecem, portanto, maior confiabilidade. Exemplos incluem: Janus, Jupiter, Marte, Fenix, Coseas, CCInt, e o banco de dados da Fuvest. As Fontes de Dados Externas são os diversos repositórios que possuem dados de cunho acadêmico, mas que não são tão confiáveis quanto os dados provenientes das fontes internas. Para as fontes de dados externas, a confiabilidade dos dados depende, por exemplo, de quem os digitam, de como são obtidos da fonte original e dos processos que os transformam. Exemplos de Fontes de Dados Externas incluem: CNPq-Grupos, Lattes, CAPES, DBLP, ISIS, e Web of Science. Os extratores de característica de fontes externas geram arquivos XML. Uma Base Consolidada armazena métricas absolutas, obtidas principalmente das fontes de dados internas. Uma Base de Comparação armazena métricas tanto da USP quanto de outras Instituições. Esta base permite a comparação, por exemplo, entre diferentes Instituições. Enquanto os dados da USP podem ser obtidos de fontes internas, os dados de outras Instituições precisam ser obtidos de fontes externas. Essa separação é importante para, no futuro, permitir que o

sistema seja utilizado, mesmo que de maneira parcial (sem a dependência do acesso às fontes internas), por outras universidades, seguindo o conceito de ampla disponibilidade do sistema, na filosofia de software livre.

Etapas de Desenvolvimento

As etapas de desenvolvimento previstas para o Sistema Tycho são as seguintes:

1. Identificar os índices que deverão ser supridos para que as atividades de avaliação e administração institucionais possam ser executadas, dentro de cada um dos quatro conjuntos de indicadores previstos; (já iniciada)
2. Identificar as características a serem utilizadas para compor cada um dos indicadores previstos, definir e implementar as métricas que deverão ser usadas para gerá-los a partir dessas características;
3. Identificar as bases de dados, tanto institucionais quanto externas à USP, que contenham dados a serem utilizados para gerar cada característica, definir e implementar o procedimento a ser seguido para cada extrator;
4. Identificar e implementar as ferramentas e metodologias para apoio à Análise de desempenho, visando tanto a avaliação institucional quanto a gestão das unidades;
5. Identificar dados de atividades da USP, necessários à avaliação e administração, que não podem ser obtidos de fontes já existentes, e implementar módulos de coleta de dados específicos a serem fornecidos pelas Unidades;
6. Identificar fontes externas, sobre as quais serão necessárias intervenções humanas para filtrar os dados coletados, e implementar as ferramentas que permitam a especificação de filtros e efetuem a recuperação dos dados num formato compatível com o Armazém de Dados do sistema.

O sistema será desenvolvido como um protótipo que evoluirá incrementalmente. Assim, deve-se começar o desenvolvimento escolhendo um conjunto pequeno de indicadores, que possam ter métricas rapidamente definidas e implementadas, baseadas em poucas características obteníveis a partir dos currículos Lattes e da base de dados

dos sistemas corporativos. Ao mesmo tempo, serão desenvolvidas ferramentas de análise adequadas aos índices atendidos. A seguir, o protótipo será validado e, uma vez aprovado, será realizado o desenvolvimento de um outro conjunto de indicadores, repetindo-se o ciclo até que todos os índices necessários estejam suportados, e ferramentas de análise adequadas estejam disponíveis.

A Etapa 1 prevê reuniões da Equipe de Coordenação com a Presidência da CERT e com as Pró-reitorias, juntamente com seus respectivos técnicos, para definir os indicadores e fornecer os subsídios necessários à elaboração dos índices. Uma atividade-piloto prevista é utilizar técnicos do Instituto de Ciências Matemáticas e de Computação (ICMC) que auxiliem em um levantamento inicial, permitindo-se executar um primeiro ciclo de desenvolvimento. Enquanto essa atividade estiver sendo desenvolvida, deverá ser criada a equipe de desenvolvimento do projeto. Quanto à internacionalização com base em indicadores de publicações, procurar-se-á estreita colaboração com pesquisadores da UFMG e da UNICAMP, os quais participaram do workshop da CPA acima mencionado.

A Etapa 2 prevê o desenvolvimento do modelo geral de implementação do Armazém de Dados, com a definição das características fundamentais necessárias para que os indicadores sejam gerados.

A Etapa 3 prevê a participação dos técnicos responsáveis pela manutenção das fontes de dados com a equipe de desenvolvimento e com os técnicos das respectivas Pró-reitorias. Numa atividade-piloto, serão escolhidos os indicadores que irão compor o primeiro ciclo de desenvolvimento do protótipo e implementadas as métricas correspondentes.

A Etapa 4 envolve a atuação da Equipe de Coordenação junto aos membros da Comissão de Avaliação Institucional, para definir as ferramentas de análise que deverão ser disponibilizadas pelo sistema.

A Etapa 5 visa incorporar ao sistema todos os recursos necessários para a avaliação e administração, a partir de dados internos da Universidade. Nas primeiras versões do protótipo, não se pretende executar essa etapa.

A Etapa 6 visa completar o sistema, incorporando-lhe os recursos necessários para obter fontes de dados externas à USP e aos mecanismos oficiais de avaliação, tais como dados disponibilizados por outras universidades, veículos de avaliação de impacto de publicações, etc. Com isso, o sistema será capaz, dentre outros benefícios, de avaliar a posição relativa dos indicadores da USP com os de outras instituições, permitindo verificar sua inserção tanto no cenário nacional quanto no internacional.

Organização Administrativa: Equipe de Desenvolvimento/Manutenção

O projeto está previsto para ser concluído em 18 meses, na filosofia de software livre. As etapas acima mencionadas serão desenvolvidas iterativa e incrementalmente, abordando as diversas dimensões dos indicadores. A primeira versão do Sistema Tycho estará pronta em 12 meses, com 4 protótipos evolutivos.

Na primeira iteração abordar-se-ão os indicadores da pós-graduação, com a perspectiva nacional (Protótipo 1). Em uma segunda iteração serão abordados os indicadores de pesquisa, e graduação (Protótipo 2). Esses protótipos do sistema TYCHO estão planejados para setembro/2008 e dezembro/2008, respectivamente. Em 2009, no 1º sem, abordar-se-ão os indicadores de cultura e extensão (Protótipo 3), estando planejado para abril/2009. Paralelamente, serão abordados os indicadores internacionais, de forma a viabilizar a avaliação das unidades em função de seu desempenho em relação aos pares internacionais, possibilitando assim a avaliação do grau de internacionalização das unidades da USP e da própria USP. Assim, pretende-se disponibilizar uma versão integrada, com todas dimensões da avaliação no 2º semestre de 2009.

Nos primeiros 12 meses já serão disponibilizados indicadores para os quatro conjuntos e para internacionalização, com ênfase para os indicadores da Graduação, Pós-graduação e Pesquisa, que possam ser gerados a partir da coleção de currículos Lattes e das bases de dados corporativas da USP. Considera-se que os currículos Lattes constituem uma fonte ao mesmo tempo rica de informações e adequada para os objetivos do sistema, uma vez que seu uso pelos docentes da universidade está regulamentado pela Resolução USP-CoPq N° 4849 de 03 de agosto de 2001 [Res4849]. Serão também desenvolvidas ferramentas de análise que permitirão a interação dos avaliadores com o sistema, utilizando recursos gráficos.

Em 2008 serão desenvolvidas funcionalidades básicas, mas que devem suportar, além dos recursos de visualização tradicionais, como tabelas e relatórios textuais, recursos de análise associativa, gerando, por exemplo, grafos de colaboração, visualização multidimensional e mineração visual de dados. Considera-se que isso seja importante para permitir o uso de indicadores relativos, além dos tradicionais indicadores absolutos, permitindo assim análises comparativas entre as unidades e entre as unidades e seus pares externos à USP (desde que dados adequados estejam disponíveis). Pretende-se também desenvolver um primeiro módulo de coleta de dados não obtíveis das fontes já existentes, a título de desenvolvimento da técnica dentro do Sistema Tycho. O alvo desse módulo será definido, em função da necessidade de algum indicador que venha a ser considerado importante durante o desenvolvimento. No segundo ano, todos os indicadores necessários serão disponibilizados, como também serão completadas as ferramentas de análise e os módulos de coleta de dados.

Recursos Humanos

Para o desenvolvimento do projeto serão necessários seis profissionais (analistas), com atribuições, responsabilidades e benefícios equivalentes aos dos bolsistas TT da FAPESP. Esses bolsistas (TT-IVa ou TT-V, da área de computação) atuarão em tempo integral no projeto, com possibilidade de realizarem atividades em programas de pós-

graduação do IME-USP ou ICMC-USP, se aceitos, com o envolvimento de $\frac{1}{4}$ do tempo na pós-graduação, e outros $\frac{3}{4}$ no projeto.

Gestão do TYCHO

Uma vez finalizado, o projeto será internalizado pelos órgãos centrais de informática da USP, prevendo-se que sua manutenção e continuação de desenvolvimento/atualização contará com dois técnicos em tempo integral, agora funcionários do quadro regular da USP. A execução permanente de coleta de dados de fontes externas à USP (correspondente à Etapa 6 das etapas de desenvolvimento) poderá requerer mais profissionais especializados em análise de dados, dependendo da intensidade com que essa coleta venha a ser executada.

A CPA deverá definir uma estrutura administrativa para apoiar essas atividades. O Departamento de Informática (DI) será o órgão coordenador do Sistema Tycho, gerenciando a equipe de desenvolvimento/manutenção, com alocação de pessoal de forma distribuída. O ICMC-USP e o IME-USP serão parceiros efetivos no sistema TYCHO. Está sendo considerado, por parte do DI, o treinamento de secretárias e de técnicos especializados das Unidades da USP, os quais darão apoio aos gestores de suas unidades no uso efetivo do Sistema Tycho, e sua eventual adaptação às especificidades das respectivas Unidades.

A Infra-estrutura de hardware necessária será oportunamente especificada pelo DI/USP, envolvendo as necessidades tanto do DI como do ICMC e IME.

REFERÊNCIAS

[Res4928] Resolução N° 4928, de 17/05/2002, que Altera dispositivos do Regimento Geral da Universidade de São Paulo, disponível em <http://www.usp.br/leginf/resol/r4928m.htm>, consultado em 15/01/2007

[Res4849] Resolução N° 4849, de 03/08/2001, que adota o Sistema Lattes como formato curricular de observância obrigatória pelos membros do corpo docente da USP, disponível em <http://www.usp.br/leginf/resol/r4849m.htm>, consultado em 24/01/2008

[Wikipedia2007] Tycho Brahe, disponível em http://en.wikipedia.org/wiki/Tycho_Brahe, consultado em 15/01/2007.

GLOSSÁRIO

Objeto complexo: são as unidades institucionais que devem ser avaliadas/administradas. Correspondem às Unidades de Ensino, Departamentos, Programas de Pós-graduação, Docentes, etc..

Características: Valores que podem ser medidos sobre cada objeto complexo, e representados como números ou categorias discretas. Por exemplo, uma Unidade de Ensino pode ter a característica “Numero de Professores Titulares”, “Valor total dos Projetos Financiados” e “Área de Atuação”, representados como valor inteiro, valor real e valor categórico, respectivamente.

Vetor de Características: Uma coleção indexada de características de um dado objeto complexo. Todo vetor de características tem uma dimensionalidade bem definida, e um dado índice de qualquer objeto complexo de um mesmo tipo representa sempre a mesma característica. Por exemplo, todas as Unidades de Ensino devem ser representadas pela coleção de todas as medidas feitas para qualquer Unidade, mesmo que para uma Unidade em particular, algumas características tenham valor nulo.

Métrica: Uma função de comparação de vetores de características (ou de Objetos complexos) segundo um determinado aspecto. Uma métrica pode ser relativa, indicando quão similares são dois vetores, ou absoluta (quanto também é chamada de **Norma**), resultando em um valor único para cada vetor.

Indicador: Valor resultante da aplicação de uma métrica sobre um ou mais vetores de características.

Fonte de dados: Um repositório de dados desenvolvido primordialmente para atender ao dia a dia do empreendimento. Armazena tipicamente grande quantidade de dados em formato tabular. São fontes de dados: as bases de dados corporativas dos sistemas Júpiter, Janus, Marte, etc. e a coleção de currícula Lattes dos docentes da USP.

Dados: Elemento básico de informação, representado por uma coleção de valores que definem as entidades representadas em bases de dados. Por exemplo, faz parte dos dados do Júpiter os horários em que determinada turma de determinada disciplina foi ministrada em um dado semestre.

Entidade: Conceito explicitamente definido, ao qual pode-se atribuir valores (medidas) em um determinado instante de tempo, por exemplo, aluno de graduação, disciplina, programas de pós, departamento, unidade, entre outros. A escolha de uma entidade deve ser bastante cuidadosa de forma a evitar interpretações dúbias.

Medida ou Amostragem: atribuição de um valor a uma característica a partir de consulta a uma base de dados. Esta amostragem deve ter o rigor científico e considerar fatores como confiabilidade e rigor.

Validade: Intervalo de tempo no qual uma característica está definida ou uma determinada entidade pode ser medida. Por exemplo, é possível obter com segurança uma contagem de docentes por departamento a partir de 1988. Uma contagem anterior esbarraria na disponibilidade e confiabilidade dos dados. A escolha de variáveis deve levar em conta validade ampla de modo a permitir a construção de séries históricas. É importante perseguir a observação de variáveis com significado mais permanente e cuja definição não mude ao longo do tempo.

Granularidade: Escopo de agregação na qual se faz a medida de uma característica. Por exemplo, a característica “Valor total dos projetos financiados” pode ter como granularidade a unidade, o departamento, o grupo de pesquisa ou o indivíduo.

Período Significativo: Intervalo de tempo convencionalizado sobre o qual a medida de uma variável tem significância. Por exemplo, não faz sentido indagar sobre o número de docentes ativos em um departamento em um determinado semestre, uma vez que o valor desta variável (docente ativo) nesta granularidade (departamento) pode variar diariamente. Neste caso, melhor seria expressar uma média (de acordo com uma fórmula dada) ou convencionalizar-se que o número de docentes ativos por semestre é o número de docentes medido em um determinado dia.

Totalizador: Valor de uma variável em relação a uma dada granularidade, como exemplo o número total de alunos de graduação na disciplina de Engenharia de Software no 2º. Semestre de 2005.

Relevância: Peso associado a uma característica avaliada por uma métrica, de maneira que o indicador resultante possa acomodar a diversidade do objeto complexo. Por exemplo, a relevância de artigos publicados em periódicos internacionais pode ser maior para a um departamento da área de Física do que para um departamento da área de Artes.

ANEXO: Resultados de Estudos sobre a Inserção Internacional da Pesquisa em Ciência da Computação no Brasil.

(extraído do Relatório do Comitê de Computação)

Este estudo foi realizado com o objetivo de comparar a produção científica dos programas nível 5, 6 e 7 da área de Ciência da Computação com a dos principais programas do mundo. Para isso, coletou-se da DBLP – Digital Bibliography Library Project (<http://dblp.uni-trier.de>) a produção científica dos docentes dos oito principais programas nacionais (PUC/Rio, URFJ/COPPE, UFMG, UFPE, UFRGS, Unicamp, USP/SP e USP/SC) e de 22 dos mais importantes programas da América do Norte e Europa assim distribuídos:

- América do Norte: Berkeley, Brown, Caltech, CMU, Cornell, Harvard, Illinois, MIT, Princeton, Stanford, UTexas at Austin, Washington, Wisconsin, British Columbia, Toronto e Waterloo;
- Europa: Cambridge, École Polytechnique de Paris, ETH Zürich, Imperial College, Oxford e Paris VI.

A escolha da DBLP como fonte de dados para este estudo se deveu ao fato de ser a mais abrangente biblioteca digital da área de Ciência da Computação, incluindo o registro de mais de 910.000 artigos publicados nos anais das principais conferências e nos mais importantes periódicos da área. Além disso, o conteúdo da DBLP é extremamente padronizado em relação aos nomes dos autores e títulos dos periódicos e conferências, facilitando a coleta e o tratamento de dados, já que a existência de homônimos entre os pesquisadores é bastante reduzida. Os artigos coletados foram então classificados conforme o Qualis da área e os dados resultantes tabulados de modo a permitir uma comparação da produção dos programas nacionais com a dos programas internacionais. Esses dados estão disponíveis em <http://www.latin.dcc.ufmg.br/publications> e podem ser acessados mediante senha a ser solicitada à equipe responsável pelo estudo (contato: Nivio Ziviani, nivio@dcc.ufmg.br).

Apesar da abrangência da DBLP, ela não cobre igualmente todas as sub-áreas da computação, focando mais naquelas de cunho tecnológico como Bancos de Dados, Engenharia de Software, Linguagens de Programação, Sistemas de Computação e Redes de Computadores. Mesmo assim, este estudo mostrou alguns resultados interessantes sobre o perfil da área:

1. A produção científica da área é fortemente focada em artigos completos e arbitrados publicados em anais de conferência em um relação próxima de 3 artigos em conferência para 1 em periódico. Essa relação de 3:1 se aplica tanto aos programas internacionais quanto aos nacionais.
2. Considerando a média de publicações (artigos em periódicos e conferências) por pesquisador no triênio 2004-2006, os 30 programas analisados (incluindo os 8 nacionais) podem ser divididos em dois grupos distintos:
 - a. O primeiro, com uma média no triênio entre 8,10 e 14,41 publicações por pesquisador, inclui 15 dos programas internacionais (Berkeley, Brown, CMU,

Cornell, ETH Zürich, Harvard, Illinois, Imperial College, Princeton, Stanford, U. of Texas at Austin, Washington, Wisconsin, Toronto e Waterloo).

- b. O segundo, com uma média entre 3,58 e 7,91 publicações por pesquisador, inclui os demais programas internacionais (British Columbia, Caltech, Cambridge, École Polytechnique de Paris, MIT, Oxford e Paris VI) e os 8 programas nacionais.
3. A média dos programas nacionais no triênio ficou entre 3,77 e 7,91, o que coloca esses programas em posição comparável à dos demais programas do segundo grupo.

I Workshop sobre Sistemas Computacionais de Apoio à Avaliação Institucional da USP.

Data: 04/setembro/2007
Local: Prédio da Reitoria, Sala A (antiga sala 4)
Horário: 8:15 - 18:30
Coordenação: Prof. Franco M. Lajolo (Vice-Reitoria, Presidente da CPA)
Prof. José Carlos Maldonado (ICMC-USP, membro da CPA)

Programa

8:15 - Abertura

9:00 - 12:00 FASE I: Experiências externas à USP

FASE I. 1 - Aspectos de Internacionalização da Produção Científica

- 9:00 - 9:30 Perfil-CC - Um Estudo sobre o Perfil da Produção Científica em Ciência da Computação
Alberto H. F. Laender e Nívio Ziviani
DCC-UFMG

Resumo: Nesta palestra apresentamos um relato dos primeiros resultados do Projeto Perfil-CC que tem como objetivo estudar o perfil da produção científica na área de Ciência da Computação. Para isso, foram coletados dados da produção científica dos oito principais programas de pós-graduação do país e de 22 dos mais importantes programas da América do Norte e Europa, contidos na DBLP - Digital Bibliography & Library Project, uma das mais abrangentes bibliotecas digitais da área. A análise desses dados coloca em perspectiva a inserção e relevância internacional dos programas de pós-graduação nacionais, possibilitando o planejamento de ações para consolidação da área no Brasil.

- 9:30 - 10:00 Um estudo comparativo entre departamentos de ciência da computação no país e no exterior baseado em publicações da Web of Science
Ricardo Anido
IC-UNICAMP

Resumo: Neste estudo foi comparada a produção científica de docentes de departamentos de Ciência da Computação do Brasil e do exterior. A metodologia permite o estudo comparativo não somente entre departamentos de computação, mas também entre os "perfis de publicação" entre diferentes áreas (como física e matemática).

FASE I. - 2 Indicadores Essenciais de Avaliação

- 10:00 - 10:30 Medição de Indicadores de Resultados na PG
João F. G. de Oliveira
EESC-USP

10:30 - 11:00 Coffee Break

- 11:00 - 12:15 I.2 Iniciativas da UNESP em Avaliação Institucional
Erivaldo Antonio da Silva
UNESP

12:15 - 13:15 - Lanche no Local

FASE II - Experiências e Iniciativas internas à USP

- 13:30 - 14:00 - Sistemas de Computação da USP: visão integrada dos dados e DATA USP
Luis Moreira e Silvio de Paula
DI-USP
- 14:00 - 14:30 - Articulação com o CNPq: CV LATTES Institucional
Luis Moreira e Silvio de Paula
DI-USP
- 14:30 - 15:00 - Desenvolvimento do sistema integrado JANUS para gerenciamento, acompanhamento e avaliação da pós-graduação: obstáculos, desafios e progresso
Prof. Dr. Fabio Kon - Assessor da Pró-Reitoria de Pós-Graduação
IME-USP

Resumo: A Pró-Reitoria de Pós-Graduação da USP iniciou há um ano uma reformulação dos seus sistemas informatizados de apoio à pós-graduação. Vários serviços relacionados tanto ao gerenciamento do dia-a-dia da pós-graduação quanto à sua análise sob um ponto de vista gerencial estão sendo integrados em um sistema único, o Janus. Nesta

apresentação, discorreremos sobre a visão que temos do sistema e de seus objetivos e descreveremos a metodologia ágil empregada para o desenvolvimento do sistema, que é baseado em software livre.

- 15:00 - 15:30 - Perspectiva da Qualidade - indicadores de produção
Dorival Leon
ICMC-USP

15:30 - 16:00 - Coffee Break

- 16:00 - 16:30 - Sistema URANO: Geração de Relatórios Departamentais a partir do CV Lattes e de dados institucionais.
Caetano Traina Jr e Cristina Ciferri
ICMC-USP

Resumo: O Sistema de Geração de Relatórios - URANO, tem por objetivo facilitar a tarefa de geração de relatórios departamentais a partir de dados já disponíveis, tanto na Plataforma Lattes como nos sistemas computacionais da USP.. Na primeira versão, ele usa apenas os dados do Lattes e deverá usar os sistemas centrais da USP na próxima versão.

- 16:30 - 17:00 - Coleta e análise dados de produção em pesquisa: relato de experiência
Roberto Marcondes César Jr
IME-USP

Resumo: Nesta palestra é relatada uma experiência sobre o desenvolvimento e aplicação de um Software livre que coleta informações de publicações do Lattes de um dado grupo de pesquisadores e gera alguns relatórios. Temos usado esse software para avaliar e planejar programas de publicações do departamento. Além disso, são mencionados alguns aspectos relevantes sobre programas como ISI e Scopus para atividades relevantes sobre avaliação de produção bibliográfica em pesquisa.

software: <http://scriptlattes.incubadora.fapesp.br/portal>

ex de saída: http://www.vision.ime.usp.br/~creativision/publications_dcc/0.html

17:00 - 18:00 - Definição de uma agenda de trabalho do grupo e planejamento.

18:00 – 18:30 - Encerramento.