

Um Processo para o Desenvolvimento de Serviços de Colaboração em um Ecossistema de Software Científico

Marcio G. Arakaki¹, Guilherme G. Martins¹, José Maria N. David¹, Regina Braga¹, Fernanda Campos¹, Frâncila W. Neiva²

¹Instituto de Ciências Exatas – Departamento de Ciência da Computação –
Universidade Federal de Juiz de Fora (UFJF)
Caixa Postal 36.036-330 – Juiz de Fora – MG – Brasil

²Programa de Pós-Graduação em Informática –
Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 68.530 – Rio de Janeiro – RJ – Brasil

{marcio.arakaki, guilherme}@ice.ufjf.br, {jose.david, regina.braga,
fernanda.campos}@ufjf.edu.br, fran.weidt@gmail.com

Abstract. *Software ecosystem capability to evolve by employees, enables the creation of an environment that can meet specific collaboration needs of each team. However, developing new services to an ecosystem can be a difficult task. This paper proposes a process for the development of collaboration services to a Scientific Software Ecosystem, E-SECO. To evaluate the process, a service was created and extended in accordance with the steps in the process. Thus, it was possible to obtain evidence regarding the validity of the process.*

Resumo. *A capacidade de ecossistemas de software de evoluírem por colaboradores possibilita a criação de um ambiente que possa atender às necessidades de colaboração específicas de cada equipe. Entretanto, desenvolver novos serviços para um ecossistema pode ser uma tarefa difícil. Este trabalho propõe um processo para o desenvolvimento de serviços de colaboração para um Ecossistema de Software Científico, o E-SECO. Para avaliar o processo, um serviço foi criado e, posteriormente, estendido de acordo com as etapas apresentadas pelo processo. Com isso, foi possível obter evidências que apontam para a validade do processo.*

1. Introdução

Experimentação científica é uma atividade colaborativa e complexa que muitas vezes exige um conhecimento diversificado por parte dos cientistas. Essa natureza da experimentação científica e o avanço da tecnologia possibilitaram o crescimento da *e-Science* que, segundo Mattoso *et al.*, (2008), pode ser definida como uma ciência em larga escala cujo objetivo é apoiar os cientistas no desenvolvimento de experimentos colaborativos e distribuídos por meio de uma infraestrutura computacional.

O E-SECO [Freitas *et al.*, 2015] é uma abordagem que utiliza os conceitos de Ecossistemas de Software Científico (ECOSC) a fim de auxiliar cientistas durante as etapas de um experimento científico [Belloum *et al.*, 2011]. Freitas *et al.*, (2015) definem um ECOSC como um Ecossistema de Software que relaciona fornecedores de software científico, institutos de pesquisa, cientistas, órgãos de fomento, instituições

financiadoras, e as partes interessadas nos resultados da pesquisa. Pela própria natureza de Ecossistemas de Software, o E-SECO é uma plataforma aberta, o que possibilita a sua evolução por meio da adição de serviços à infraestrutura e de melhorias sugeridas por desenvolvedores externos à equipe mantenedora. Com isso, é possível desenvolver os serviços necessários para oferecer suporte à colaboração entre os cientistas. Com o objetivo de promover as atividades neste contexto, os serviços de colaboração devem apoiar a comunicação, a coordenação, a cooperação e, por fim, a percepção, entre outros elementos que interferem diretamente no modelo de colaboração [Fuks *et al.*, 2003].

A forma como cientistas colaboram varia de acordo com o contexto do experimento, o ambiente de trabalho e as pessoas envolvidas. Por isso, durante o desenvolvimento, os aspectos específicos do contexto em que os serviços de colaboração serão inseridos devem ser considerados, de forma a atender os requisitos funcionais e não funcionais de cada instância do experimento. Mesmo contendo características específicas, estes serviços possuem elementos comuns e variáveis quanto aos conceitos de colaboração. Nesse cenário, o uso de Linhas de Produto de Software (LPS) possibilita reduzir o retrabalho e diminuir o esforço do desenvolvimento de novos serviços [Gurp *et al.*, 2001].

Ao desenvolver serviços de colaboração para o E-SECO, é necessário considerar as singularidades da infraestrutura do ecossistema, de modo a garantir que seus aspectos de qualidade, tais como flexibilidade, extensibilidade e escalabilidade [Freitas *et al.*, 2015], sejam mantidos. O problema é que, uma vez que o E-SECO é uma iniciativa *open-source* em que toda comunidade pode contribuir para sua evolução, fazer com que esta comunidade considere, durante o desenvolvimento de serviços de colaboração, às singularidades da infraestrutura do ecossistema se torna um trabalho difícil e complexo. A falta de uma iniciativa que apoie a comunidade durante o desenvolvimento dificulta a manutenção de um ecossistema de software aderente aos requisitos não funcionais e, ao mesmo tempo, prejudica a criação de serviços de qualidade que poderiam ser reutilizados.

Para enfrentar tal problema, este trabalho contribui com um processo de referência para o desenvolvimento de serviços de colaboração no contexto de um ecossistema de software científico, em específico, considerando a infraestrutura do E-SECO [Freitas *et al.*, 2015]. No contexto de LPS, processos também são utilizados para apoiar o desenvolvimento de software [Gimenes and Travassos, 2002] [Kang *et al.*, 2002]. Porém, no melhor do nosso conhecimento, não encontramos propostas de processos de desenvolvimento aplicados ao desenvolvimento de serviços de colaboração em ecossistemas de software científico.

O restante deste artigo no está organizado da seguinte forma: a seção 2 apresenta alguns trabalhos relacionados ao processo proposto, a seção 3 descreve as etapas do processo, a seção 4 descreve o funcionamento do processo enquanto a seção 5 apresenta sua avaliação, por fim, a seção 6 apresenta as considerações finais do trabalho.

2. Trabalhos Relacionados

Existem alguns trabalhos na literatura que apoiam a colaboração no contexto de *e-Science*. Zhang et al., (2014) apresentam Confucius, uma ferramenta de suporte à concepção colaborativa de *workflows* científicos por meio do apoio de uma ontologia de

proveniência. Confucius foi desenvolvido a partir do sistema Taverna¹, um sistema de gerenciamento de *workflow* de código aberto. Confucius se baseia no rastreamento das interações humanas durante o processo de concepção de um *workflow*, gravando-as com o apoio da ontologia de proveniência. Segundo os autores, essa ontologia é centrada nos conceitos de *workflow*, que compreendem, por exemplo, a organização de tarefas, a ligação entre os dados, e a definição de pré-requisitos. Em suma, esta ontologia armazena informações sobre as atividades, usuários envolvidos com a concepção, nível de acesso de cada usuário a cada atividade do *workflow*, e as provê ao sistema para apoiar a coordenação do processo. Apesar de Confucius fornecer apoio à colaboração, não oferece, por exemplo, serviços de comunicação para apoiar os cientistas em experimentos geograficamente distribuídos apoiados por uma infraestrutura que utiliza a abordagem de ecossistemas de software.

Miranda *et al.*, (2014) apresentam CollabCumulus, uma abordagem que auxilia a análise e a consulta de dados de proveniência produzidos pela execução de *workflows* científicos em ambientes distribuídos. Seu objetivo é abstrair para os cientistas a estrutura e a linguagem de consulta ao banco de dados que os apoia nos experimentos. Para isso, CollabCumulus apresenta um portal, que funciona como uma rede social, que se comunica com a base de dados de proveniência, de modo a possibilitar que um cientista realize e compartilhe suas análises e opine sobre as análises de outros cientistas. Assim, como o trabalho anterior, este trabalho não fornece apoio à colaboração de forma distribuída em um contexto aberto, além disso, a abordagem não apresenta uma proposta para apoiar a criação de novos serviços de colaboração ou evoluir os serviços existentes.

Gadelha *et al.*, (2009) apresentam uma LPS para a derivação de serviços de colaboração. Essa LPS foi construída a partir da análise de algumas aplicações colaborativas de modo a representar as semelhanças e diferenças nelas encontradas. Esta abordagem se baseia no Modelo de Colaboração 3C [Fuks *et al.*, 2003] para orientar o desenvolvedor a selecionar os conceitos de colaboração desejados no serviço. Assim como a proposta Gadelha *et al.*, (2009), a pesquisa deste artigo se baseia no Modelo 3C para organizar os conceitos de colaboração, entretanto, estes conceitos são utilizados por meio de uma Ontologia de Colaboração [Martins *et al.*, 2015]. Além disso, Gadelha *et al.*, (2009) apresentam uma única LPS que deriva como produto uma aplicação de colaboração que contém vários serviços, por exemplo, uma aplicação que contém Lista de Discussões e Troca de Mensagens, enquanto este trabalho propõe a utilização de uma LPS no contexto de uma infraestrutura (aberta) que utiliza a abordagem de ecossistemas de software. Além disso, neste contexto, as pesquisas previamente apresentadas não oferecem um processo para a construção de serviços de colaboração considerando os requisitos que a infraestrutura necessita.

Diferente dos trabalhos encontrados na literatura, esta pesquisa contribui para os avanços das pesquisas em sistemas colaborativos no contexto de *e-Science* através de uma proposta de processo para o desenvolvimento de serviços de colaboração (criação e/ou extensão) aplicado ao domínio de ECOSC. A proposta de processo é apresentada a seguir.

¹ <http://www.taverna.org.uk/>

3. Processo para o Desenvolvimento de Serviços de Colaboração

Para desenvolver serviços de colaboração, no contexto do E-SECO, os desenvolvedores devem seguir o processo de desenvolvimento proposto de acordo com a infraestrutura do ecossistema. Este processo tem por objetivo padronizar a criação e a evolução de componentes de colaboração, de modo a apoiar a variabilidade, especificidade e interoperabilidade entre os serviços de colaboração desenvolvidos.

Ao se desenvolver um serviço de colaboração, o mesmo será integrado à infraestrutura do ecossistema e poderá ser utilizado por qualquer instância do E-SECO, de acordo com suas necessidades. Para isso, o processo de desenvolvimento envolve os conceitos de Linhas de Produto de Software (LPS), utilizando um repositório *online* como núcleo de artefatos, para apoiar a produção dos serviços de colaboração que são utilizados na instância.

A adoção de processos traz benefícios [Gimenes and Travassos, 2002], entre eles a documentação e o compartilhamento de boas práticas. No E-SECO, as boas práticas representadas por meio do processo proposto visam garantir serviços reusáveis e compatíveis, o contínuo atendimento aos requisitos não funcionais demandados por um ECOSC. Adicionalmente, pode diminuir a complexidade do desenvolvimento de serviços, uma vez que os desenvolvedores agora possuem um processo bem estabelecido para guiá-los.

Dessa forma, acredita-se que, sem um processo para auxiliar o desenvolvimento de serviços de colaboração, estes poderão ser construídos sem uma adequação aos processos já existentes na infraestrutura do E-SECO [Freitas *et al.*, 2015]. Por exemplo, serviços devem estar alinhados aos processos de suporte à interoperabilidade, de busca e recuperação de componentes, entre outros. A seguir, um processo para o desenvolvimento de serviços de colaboração será apresentado. Tal processo é inspirado em processos presentes na literatura ([Gimenes and Travassos, 2002] [Kang *et al.*, 2002]) e aborda elementos do contexto do E-SECO ([Costa *et al.*, 2015] [Freitas *et al.*, 2015] [Neiva *et al.*, 2015]).

3.1. Desenvolvimento de Serviços de Colaboração

A Figura 1 ilustra o processo proposto para o desenvolvimento de serviços de colaboração. Este processo é composto por seis atividades principais.

Etapa 1: Planejamento Inicial. Atividade inicial do processo por meio da qual é gerada uma especificação inicial sobre as características desejadas para um serviço de colaboração (Figura 1 - 1). Como resultado, tem-se uma especificação inicial de um serviço de colaboração.

Etapa 2: Busca. Utilizando as informações da etapa anterior, realiza-se uma busca na plataforma do ecossistema a fim de encontrar uma LPS que atenda total ou parcialmente à especificação criada (Figura 1 - 2). Como resultado da busca podem ocorrer três casos. No primeiro caso, nenhuma LPS que derive um serviço de colaboração, que atenda às necessidades desejadas, foi encontrada (Figura 1 - 3A), com isso, o serviço deverá ser desenvolvido desde o início (Processo de Criação). No segundo caso, encontra-se uma LPS que atende parcialmente às necessidades desejadas (Figura 1 - 3B). Com isso, a LPS será desenvolvida a partir dos artefatos existentes de modo a evolui-la (Processo de Extensão). Por fim, no terceiro caso, encontra-se uma

LPS que atende completamente às necessidades desejadas (Figura 1 – 3C). Com isso, não é necessária nenhuma alteração na LPS, uma vez que já existe a possibilidade de derivação de um serviço com as características desejadas. Como resultado desta etapa, tem-se o conjunto de artefatos pertencentes à LPS, caso alguma seja encontrada. É importante destacar que a atividade de busca é responsável por evitar retrabalho, ou seja, o desenvolvimento de serviços de colaboração redundantes. Contudo, não faz parte do escopo deste trabalho definir uma estratégia de busca e recuperação de serviços.

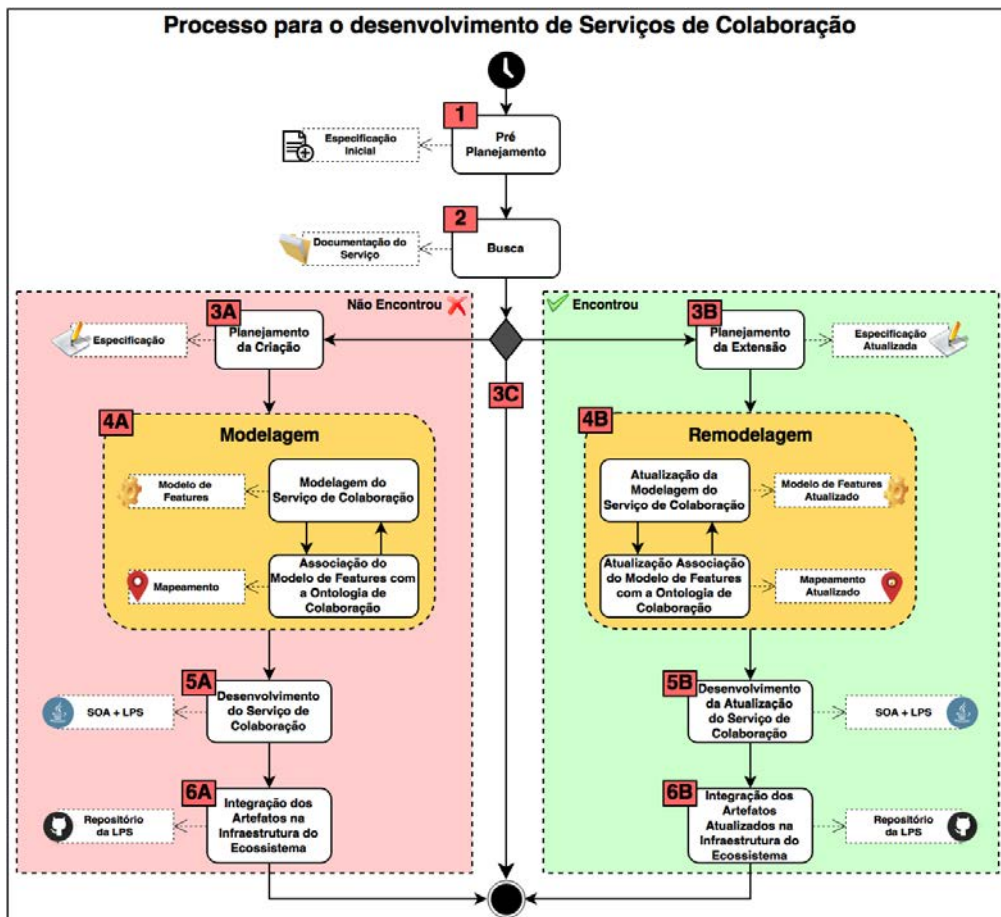


Figura 1. Processo para o Desenvolvimento de Serviços de Colaboração

As etapas posteriores do processo possuem semelhanças e diferenças entre a criação e a extensão de uma LPS que derive um serviço de colaboração. Por isso, serão divididas em dois casos, o primeiro referente à criação e o segundo à extensão.

Etapa 3: **Planejamento**. No caso da criação (Figura 1 - 3A), para desenvolver um novo serviço de colaboração deve-se fazer o seu planejamento, tendo como base a especificação inicial, de modo a definir as características do serviço. Além disso, deve-se analisar sua necessidade e sua contribuição para o ecossistema. Ao final desta etapa, tem-se um conjunto de documentos referentes à especificação do serviço de colaboração, por exemplo, descrição das funcionalidades e *features*, descrição de casos de uso, documento de requisitos e qualquer outro documento que a equipe desenvolvedora julgar necessário. O caso da extensão (Figura 1 - 3B), assim como no caso anterior, se refere ao planejamento do serviço de colaboração, a diferença é que o trabalho se concentra na evolução da documentação existente, e, caso seja necessário,

poderão ser adicionados novos documentos. Ao final desta etapa, tem-se um conjunto atualizado de documentos referentes à especificação do serviço de colaboração.

Etapa 4: **Modelagem**. Uma das propostas do E-SECO é que seus serviços sejam desenvolvidos de acordo com uma LPS. Para representar as variabilidades e semelhanças dos serviços presentes na LPS, utiliza-se o Modelo de *Features* [Kang *et al.*, 2002]. Como apresentado em Costa *et al.*, (2015), realiza-se o mapeamento do Modelo de *Features* com a Ontologia de Colaboração de modo a agregar mais semântica ao Modelo de *Features*. A Ontologia de Colaboração apresenta os conceitos que envolvem os serviços de colaboração em um ecossistema, e o relacionamento entre esses conceitos, possibilitando a verificação de consistência dos serviços gerados e a derivação de novos conhecimentos por meio do mecanismo de inferência [Martins *et al.*, 2015]. Como pode ser observado na Figura 1 (itens 4A e 4B), a modelagem é uma atividade cíclica que varia entre a criação do Modelo de *Features*, e seu mapeamento com a ontologia. Isso se deve ao fato de que a ontologia relaciona os conceitos de colaboração, e estes podem influenciar a criação e modificação do Modelo de *Features*, apresentando, por exemplo, conceitos que deveriam ser contemplados pelo modelo que, em um primeiro momento, não haviam sido incluídos. No caso da criação de serviços (Figura 1 - 4A), um novo Modelo de *Features* e um novo arquivo de mapeamento entre ele e a Ontologia de Colaboração devem ser criados segundo o planejamento da etapa anterior. Ao final desta etapa, tem-se um arquivo XML com o Modelo de *Features* do serviço de colaboração, e um arquivo XML com o mapeamento entre o Modelo de *Features* e a Ontologia de Colaboração. No caso da extensão (Figura 1 - 4B), tanto o Modelo de *Features* quanto o arquivo de mapeamento devem ser alterados segundo a evolução planejada na etapa anterior. Ao final desta etapa, tem-se o arquivo XML do Modelo de *Features* e o arquivo XML do mapeamento atualizados.

Etapa 5: **Desenvolvimento**. Essa atividade representa o desenvolvimento do serviço de colaboração. Para seguir o padrão de desenvolvimento do E-SECO, os serviços devem ser desenvolvidos seguindo os conceitos de (i) uma LPS, o que possibilita maior variabilidade do serviço de colaboração, além de facilitar a evolução do serviço por meio do reuso, e (ii) os conceitos de Arquitetura Orientada a Serviço (SOA do inglês *Service Oriented Architecture*), de modo a facilitar a composição e a interoperabilidade entre os serviços do ecossistema [Neiva *et al.*, 2015]. É importante destacar que cada serviço de colaboração possui sua própria LPS, na qual os componentes são *features* e os produtos são serviços que apresentam um subconjunto dessas. No caso da criação (Figura 1 - 5A), o desenvolvimento deve ser realizado desde o início de acordo com o planejamento e a modelagem realizadas nas etapas anteriores. Ao final desta etapa, tem-se a LPS desenvolvida e apta a derivar serviços de colaboração. No caso da extensão (Figura 1 - 5B), o desenvolvimento da evolução deve ser realizado a partir da LPS existente de acordo com o planejamento e a modelagem realizadas anteriormente. Ao final desta etapa, tem-se a LPS atualizada e apta a derivar novos serviços de colaboração. É importante destacar que, nesta etapa, tanto na criação quanto na extensão, é desejável que se realizem testes à LPS de modo a verificar se o que foi desenvolvido está consistente com o que foi especificado. Estes testes devem ser documentados e anexados ao conjunto de artefatos da LPS do serviço de colaboração.

Etapa 6: **Integração**. Nesta última atividade, o serviço desenvolvido será integrado à Infraestrutura do ecossistema, juntamente com os artefatos gerados nas atividades anteriores. Neste trabalho, a integração ocorre através de um repositório no

GitHub², onde todos os serviços desenvolvidos, e os artefatos relacionados a eles, são armazenados e disponibilizados para serem utilizados por qualquer instância do E-SECO que o deseje. No caso da criação (Figura 1 - 6A), cria-se uma nova área no repositório para armazenar os artefatos gerados em todas as etapas anteriores. Ao final desta etapa, tem-se a LPS integrada à infraestrutura do E-SECO. No caso da extensão (Figura 1 - 6B), os novos artefatos, e os que foram modificados, são armazenados na área já existente do serviço estendido. Ao final desta etapa, tem-se a LPS integrada à infraestrutura do E-SECO. Em ambos os casos, os artefatos submetidos à integração com a infraestrutura do E-SECO serão avaliados e, caso sejam aprovados, serão anexados à infraestrutura.

Após a integração da LPS do serviço de colaboração à infraestrutura do ecossistema, ele estará apto a derivar serviços que poderão ser utilizados por qualquer instância do E-SECO. Como parte do trabalho, o processo proposto foi avaliado de duas formas, uma avaliando a criação e outra a extensão de serviços. Para a criação, o processo foi utilizado para a criação de um novo serviço de colaboração para o E-SECO. A extensão foi avaliada por meio de um estudo experimental onde um serviço de colaboração presente no ecossistema foi estendido. As avaliações são descritas a seguir.

4. Processo de Criação de Serviço em Ação

Com o objetivo de avaliar a viabilidade do processo quanto ao caso de criação de um novo serviço de colaboração para o E-SECO, esta seção apresenta a criação de um serviço de Lista de Usuários.

Seguindo o processo, a primeira etapa é o planejamento inicial, na qual cria-se uma especificação inicial da Lista de Usuários contendo uma breve descrição do serviço e das funcionalidades desejadas, neste caso, uma Lista de Usuários que apresente as seguintes informações sobre seus participantes: Nome, Função e *Status*.

Na segunda etapa do processo, realiza-se uma busca no repositório por uma LPS que consiga derivar um serviço de Lista de Usuário compatível com as informações descritas na etapa anterior. Como não há uma LPS que atente total, ou parcialmente, à especificação inicial, a busca não retorna nenhum conjunto de artefatos, o que indica que o caso de criação de um novo serviço deverá ser utilizado. É importante destacar que o processo prevê a etapa de busca, entretanto, não existe uma estratégia para tal. A princípio, a busca é realizada verificando a descrição de todos os serviços presentes no repositório manualmente. A especificação de uma estratégia de busca foge ao escopo deste trabalho.

Na terceira etapa do processo, realiza-se o planejamento do novo serviço de colaboração. Nesta atividade, foram descritos os requisitos desejados para a Lista de Usuários, são eles: (1) todos os usuários cadastrados devem ser listados; (2) deve-se informar qual a função ocupada por um usuário; (3) deve-se informar se um usuário está *online* ou *offline*, ou seja, seu status.

Na quarta etapa, cria-se o Modelo de *Features* do serviço e o mapeamento do modelo com a Ontologia de Colaboração. O modelo foi criado de acordo com o planejamento da etapa anterior, como ilustrado na Figura 2. Para criar o Modelo de

² https://github.com/pgcc/plscience-ecos/tree/master/web/files/Collaboration_Services

Features utilizou-se a ferramenta FeatureIDE³. Após a criação do modelo, realiza-se o seu mapeamento com a ontologia. Para analisar a Ontologia de Colaboração de modo a auxiliar a criação do arquivo de mapeamento, utilizou-se a ferramenta Protégè⁴.

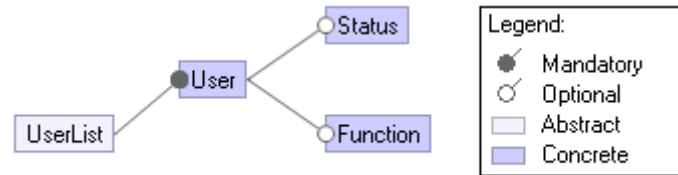


Figura 2. Modelo de Features do Serviço de Colaboração Lista de Usuários

Após a modelagem, na quinta etapa do processo, a Lista de Usuários foi desenvolvida baseando-se nos artefatos gerados anteriormente, principalmente do Modelo de Features, pois ele reflete a LPS da Lista de Usuários implementada. A lista foi desenvolvida na forma de um serviço como prevê o processo, e suas *features* foram implementadas como métodos do serviço, de forma a facilitar sua reutilização. Após desenvolver o serviço, criou-se um arquivo WSDL com sua descrição.

Na última etapa, todos os artefatos (incluindo o serviço) foram integrados ao ecossistema. Para isso, criou-se um local com nome “*User List*” no repositório de LPS no GitHub, onde foram adicionados todos os artefatos gerados nas etapas do processo, incluindo o código e descrição do serviço. Com isso, o E-SECO foi evoluído e a Lista de Usuários desenvolvida poderá ser reutilizada, assim como seus artefatos. A utilização por uma instância do E-SECO do serviço de Lista de Usuários derivado da LPS desenvolvida através do processo está destacada na Figura 3.

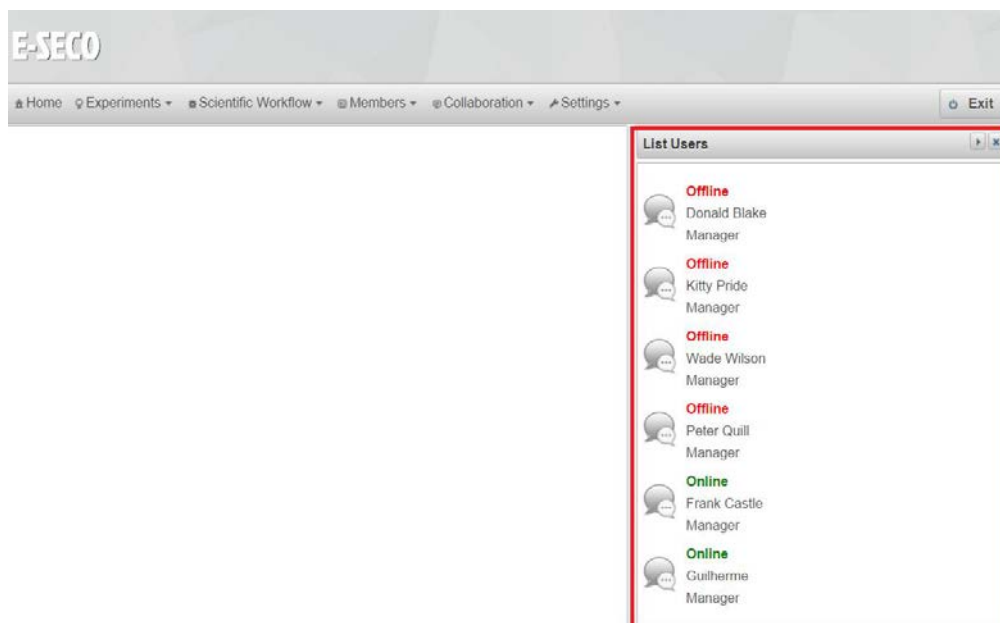


Figura 3. Serviço de Lista de Usuários Utilizado em uma Instância do E-SECO

Como pode ser observado na descrição da utilização do processo, ele se mostrou útil para padronizar a criação de serviços de colaboração, assim como para organizar os

³ http://www.witi.cs.uni-magdeburg.de/iti_db/research/featureide/

⁴ <http://protege.stanford.edu/>

artefatos gerados em cada etapa de modo a facilitar sua utilização, reutilização e evolução. Dessa maneira, as evidências geradas apontam para a viabilidade do processo. A seguir a avaliação do caso de extensão do serviço de colaboração será apresentado.

5. Avaliação do Processo: Caso de Extensão

Para avaliar o processo proposto quando ao caso de extensão, realizou-se um experimento formalmente definido de modo que os participantes aplicassem o processo para evoluir um serviço de colaboração do E-SECO. Primeiramente foi definido o escopo do experimento utilizando o GQM [Basili *et al.*, 2001]: “**Analisar o processo proposto para evoluir a infraestrutura do ecossistema com o propósito de avaliar seu suporte ao desenvolvimento de serviços de colaboração com respeito a sua usabilidade, compreensibilidade e completude do ponto de vista de desenvolvedores no contexto do desenvolvimento de serviços para o E-SECO.**”.

A partir do escopo foram definidas quatro hipóteses alternativas. Espera-se, através da avaliação destas, apresentar evidências sobre a utilidade, compreensibilidade, usabilidade e completude do processo. São elas:

- H0: O processo proposto auxilia na extensão de serviços de colaboração do E-SECO.
- H1: O processo apresenta todas as informações necessárias para a execução do mesmo.
- H2: Todas as informações presentes no processo são de fácil compreensão.
- H3: O processo pode ser utilizado sem a necessidade treinamento prévio.

O experimento consiste na aplicação do caso de extensão do processo por um participante voluntário com o intuito de evoluir a infraestrutura do E-SECO. Em específico, foi dado ao participante a tarefa de evoluir a LPS do serviço de Lista de Usuário existente no ecossistema, com isso, o participante deveria seguir as etapas apresentadas no processo, de modo a gerar os artefatos neles descritos. Essa evolução consistia na adição da possibilidade de acrescentar os usuários das Lista de Usuários em grupos. Para simular um contexto mais realista e avaliar a usabilidade do processo, não foi dado um treinamento prévio para o uso do processo.

A seleção do participante foi criteriosa, isso porque o processo foi criado visando apoiar a desenvolvedores que já possuem um certo conhecimento prévio, fator que teve de ser levado em conta na seleção de indivíduos, de forma a não prejudicar os resultados do estudo. O processo é direcionado a desenvolvedores com bom conhecimento em desenvolvimento com Java *Web*, SOA e LPS, além de conhecimento sobre edição e compreensão de Modelos de *Features*, e conhecimento básico em ontologias. Para caracterizar os participantes, utilizou-se um questionário⁵, onde foi selecionado apenas o participante que possuía o conhecimento previsto para o experimento.

A coleta de dados foi feita por meio de observação direta e por um questionário de avaliação⁶. No questionário encontravam-se perguntas fechadas referentes a cada atividade e duas perguntas abertas referentes ao processo como um todo. A análise dos dados coletados é apresentada na seção 5.2.

⁵ <https://drive.google.com/open?id=0B1uZYT7TKziBTWYxbmtlalhkNEk>

⁶ <https://drive.google.com/open?id=0B1uZYT7TKziBZWFaRmkxZ3ROQ3M>

5.1. Execução do Experimento

Este estudo considerou que as informações contidas na descrição do objetivo do experimento podem ser consideradas equivalentes às informações necessárias na primeira etapa do processo, a etapa do Planejamento Inicial. Portanto, na primeira atividade do estudo, o participante deveria buscar no repositório um serviço que estivesse de acordo com a descrição da tarefa. Esperava-se que o participante encontrasse o serviço de lista de usuários já disponível no ecossistema.

Como esperado, o participante conseguiu encontrar e recuperar as informações referentes a Lista de Usuários existente no ecossistema. Durante a execução da atividade foi possível constatar que o participante encontrou o serviço no repositório pelo seu nome (*User_List*), e baseado no Modelo de *Features* presente no repositório, ele identificou que o serviço se aproximava bastante do que ele precisava. Como o participante encontrou a LPS do serviço de Lista de Usuários disponível no ecossistema, ele seguiu o caso de extensão do processo proposto. É importante ressaltar que o serviço encontrado é o serviço que foi desenvolvido na seção anterior.

De posse dos artefatos referente à Lista de Usuários recuperada do repositório, a primeira ação do participante foi observar o Modelo de *Features* do serviço, posteriormente, ele analisou as descrições das *features* presentes no modelo e as comparou com os requisitos definidos para o serviço. Seguindo a descrição da etapa três do processo, o participante deveria criar o planejamento da extensão, para isso, ele estendeu as descrições existentes de modo a acrescentar o novo requisito de criação de grupos. O participante não sentiu a necessidade de criar nenhum documento adicional e avançou para a etapa seguinte.

Após o planejamento, o participante começou a quarta etapa do processo, a etapa de modelagem. Nessa atividade, o participante deveria atualizar Modelo de *Features* e o arquivo de mapeamento entre o modelo e a Ontologia de Colaboração. Para atualizar o modelo, ele utilizou o FeatureIDE e para verificar a ontologia para auxiliar a atualização do arquivo de mapeamento, ele utilizou o Protégè. Durante a atualização do modelo, o participante utilizou parte do seu tempo para analisar o modelo existente. Acredita-se que esse comportamento foi influenciado pelo fato de que, o processo enfatiza a consistência dos artefatos já existentes. Diante disso, é necessário observar as relações da nova *feature* com as já presentes no modelo. A Figura 4 apresenta o Modelo de *Features* atualizado.

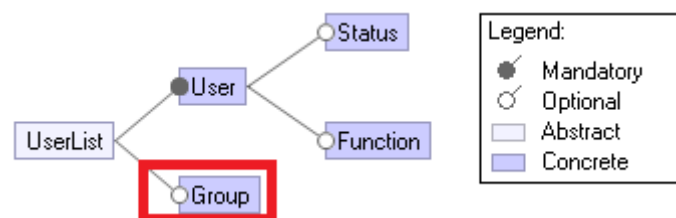


Figura 4. Modelo de *Features* do Serviço de Colaboração Lista de Usuários Atualizado

Na etapa de desenvolvimento, o participante deveria implementar a *feature* de grupo para a Lista de Usuários. Neste momento, ele deveria atentar-se aos princípios de SOA, de forma que a Lista de Usuários continuasse funcionando como um serviço após a implementação. Durante esta etapa, o participante estudou o código existente a fim de

entender seu funcionamento, para posteriormente, desenvolver a *feature* de grupo. Essa atividade tomou muito tempo, uma vez que não há nenhum elemento no processo que auxilie o desenvolvedor a nível de código, ou seja, essa etapa depende fortemente dos conhecimento e capacidades do participante. O participante conseguiu desenvolver a nova *feature* na forma de um novo método do serviço de Lista de Usuário e, em seguida, gerou a nova descrição (arquivo WSDL) para o serviço. Para testar a nova funcionalidade do serviço de Lista de Usuários, um serviço contendo a *feature* Grupo foi derivada da LPS e adicionada a uma instância local do E-SECO (Figura 5).

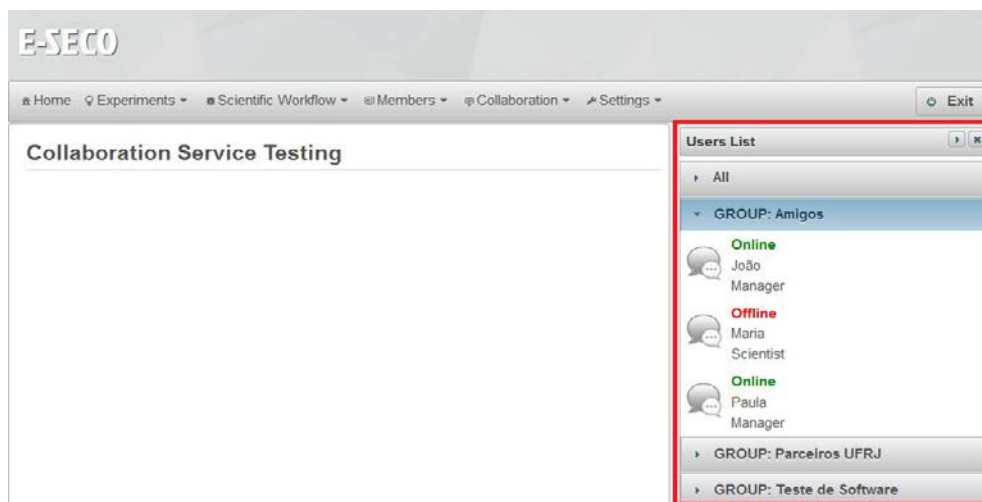


Figura 5. Serviço de Lista de Usuários com a *Feature* Grupo Utilizado em uma Instância do E-SECO

A última etapa do processo é a integração, nela o participante teria que atualizar o repositório da LPS da Lista de Usuários com os artefatos gerados nas etapas anteriores, disponibilizando assim o serviço no ecossistema. Para isso, o participante conferiu e reuniu todos os artefatos gerados durante a estudo, localizou a área do repositório no GitHub, referente a Lista de Usuários, e realizou o *upload* dos artefatos. Nessa atividade o participante não apresentou nenhuma dificuldade. Após a integração dos artefatos, o participante respondeu ao questionário. A seguir, a análise do resultado será apresentada.

5.2. Análise do Resultado

Para analisar os dados obtidos durante a realização do experimento, as respostas dos questionários foram fornecidas por meio de uma escala Likert de cinco níveis, contendo um nível intermediário como proposto por [Laitenberger and Dreyer, 1998]. Para cada resposta presente nas questões fechadas, foi atribuída uma pontuação: 1 para “discordo fortemente”, 2 para “discordo”, 3 para “não consigo e/ou não desejo responder”, 4 para “concordo” e 5 para “concordo fortemente”. Para cada etapa executada pelo participante, são apresentadas três questões, uma referente a utilidade do processo, uma sobre sua compreensibilidade e uma sobre sua completude. A Figura 6 apresenta um resumo dos resultados de cada etapa. É importante lembrar que a primeira etapa, referente ao planejamento inicial, não foi executada pelo participante, como discutido anteriormente.

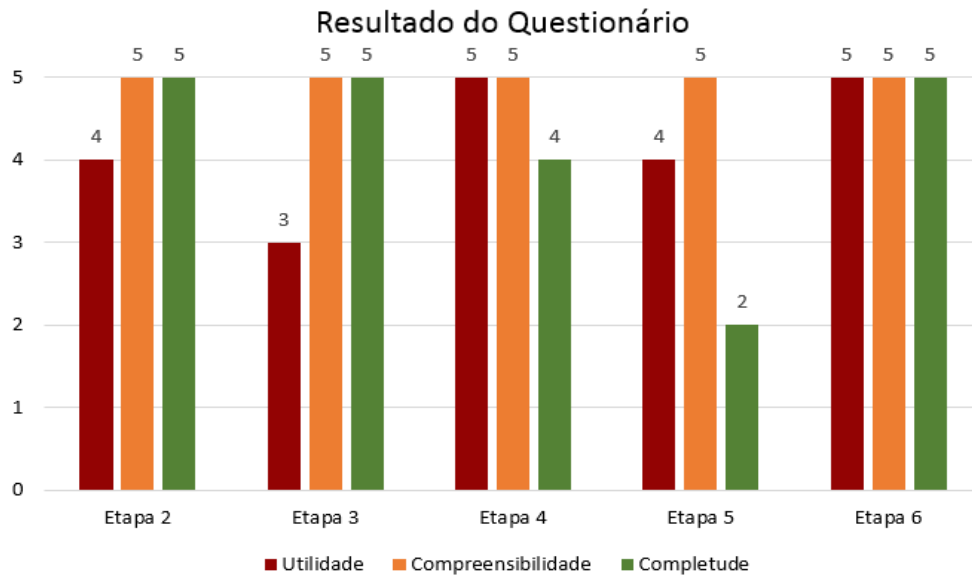


Figura 6. Resultado das Etapas do Questionário

Como pode ser observado na Figura 6, os resultados da etapa 2 indicam que as informações contidas no processo foram úteis para a realização da busca, foram de fácil compreensão e estavam completas. Nesta etapa, a observação direta indicou que as informações fornecidas pelo estudo foram suficientes para o participante realizar a busca, além disso, a complexidade da busca foi baixa, uma vez que existiam poucos serviços no repositório. Os resultados da etapa 3 indicam que as informações para a atualização do planejamento contidas no processo foram de fácil compreensão e estavam completas, contudo, o participante não concordou nem discordou que a descrição desta etapa foi útil. Nesta etapa, a observação direta indicou que o participante considerou esta etapa simples, pois se tratava de uma evolução simples de uma LPS já existente, e por não necessitar de nenhum documento adicional. Os resultados da etapa 4 indicam que as informações para a atualização do Modelo de *Features* e do mapeamento foram úteis, estavam compreensíveis e completas. Nesta etapa, a observação direta indicou que a descrição dos artefatos, a descrição dos cuidados que deveriam ser tomados e a indicação de ferramentas para executar as atividades foram muito importantes para o participante. Os resultados da etapa 5, onde o desenvolvimento é realizado, se mostrou útil e compreensível, entretanto, o participante discordou que a descrição desta etapa estava completa. Nesta etapa, a observação direta indicou que, apesar de descrever adequadamente os artefatos envolvidos, o processo não oferece suporte ao desenvolvedor a um nível de código, com isso, a atividade de implementação foi considerada bastante trabalhosa. Por fim, os resultados da etapa 6, responsável pela integração dos artefatos atualizados ao repositório da LPS, se mostrou útil, compreensível e completo. Nesta etapa, a observação direta indicou que o participante já estava familiarizado com a utilização do GitHub, o que facilitou esta atividade.

Além das questões fechadas, o participante respondeu a duas questões abertas. A primeira pergunta foi: “Como você descreve o suporte oferecido pelo processo em relação a realização do experimento?”. Nesta questão, o participante indicou que o processo foi de grande auxílio durante o experimento, enfatizando que provavelmente não teria conseguido completar a tarefa sem o auxílio do mesmo. A segunda pergunta

foi: “Você sentiu falta de um treinamento para sobre a utilização do processo?”. Nesta questão, o participante indicou que não houve necessidade de um treinamento prévio sobre o processo.

As informações analisadas apresentam evidências que indicam que as hipóteses levantadas podem ser confirmadas, apontando para a viabilidade do processo quanto ao caso de extensão. Com isso, pode-se dizer que o processo auxiliou a extensão da LPS do serviço de Lista de Usuários e que atendeu aos requisitos de usabilidade, compreensibilidade e completude, o que responde às questões de pesquisa apresentadas.

5.3. Ameaças à Validade

É preciso considerar as ameaças à validade presente neste experimento. Quanto à validade externa, (i) o processo está limitado ao contexto do E-SECO, mesmo dentro do contexto de *e-Science*, ecossistemas podem apresentar muitas diferenças, como, por exemplo, não utilizarem uma LPS; além disso, (ii) o processo se limita à criação e extensão de serviços de colaboração. Quanto à validade de resultados, o estudo não foi realizado com uma quantidade suficiente de participantes para se obter uma análise estatisticamente relevante, o que pretende-se realizar como um trabalho futuro.

6. Considerações Finais

Com advento da *e-Science*, a experimentação científica alcançou novas escalas, as equipes de cientistas passaram a estar geograficamente distribuídas, novas tecnologias foram inseridas e, com isso, a necessidade de sistemas que apoiassem a realização de experimentos e, principalmente, a colaboração nesse cenário. Pensando nisso, foi desenvolvido o E-SECO, um ecossistema de software científico (ECOSC) que, por meio de uma combinação de serviços, oferece mecanismos de apoio a colaboração e à realização de experimentos científicos colaborativos.

Para evoluir o E-SECO e criar serviços de colaboração que atendam às necessidades das equipes de cientistas, é necessário considerar os princípios e tecnologias existentes no ecossistema. Diante disso, esse trabalho propôs um processo de desenvolvimento de serviços de colaboração para o E-SECO. Esse processo considera os conceitos de LPS de forma a possibilitar a variabilidade e a reutilização dos serviços. Duas atividades são descritas no processo: (i) a criação de novos serviços de colaboração e a (ii) extensão de serviços de colaboração já existentes. A primeira auxilia o desenvolvedor a criar novos serviços, enquanto a segunda auxilia o desenvolvedor a estender um serviço existente. Em ambos os casos, ao final do processo o novo serviço (ou extensão do serviço) é disponibilizado no ecossistema por meio do repositório de LPS.

Para avaliar o processo proposto foi desenvolvido uma Lista de Usuários seguindo o caso de criação do processo, e esse mesmo serviço foi estendido em um estudo experimental seguindo o caso de extensão do processo. Os dados coletados e analisados durante o estudo experimental apresentaram evidências de que o processo proposto atende aos requisitos de usabilidade, compreensibilidade e completude desejados. Entretanto, o estudo carece de resultados com relevância estatística, o que faz necessária a realização de mais experimentos e em uma escala maior. Espera-se também que o processo possa ser refinado conforme mais estudos sejam realizados.

Como trabalhos futuros em relação ao processo cabe citar: (i) realizar um estudo experimental sobre a criação de um novo serviço de colaboração, e não apenas sobre a extensão; (ii) avaliar a possibilidade de sua aplicação em outros domínios; (iii) criação de mecanismos de busca e gerência dos serviços desenvolvidos e disponibilizados no ecossistema, o que implica na definição de uma especificação inicial formal; (iv) evoluir o processo para que ele considere a composição de serviços de colaboração a partir dos já existentes. Cabe citar também trabalhos referentes aos serviços de colaboração desenvolvidos, como: (i) avaliar como os serviços de colaboração desenvolvidos utilizando o processo proposto, oferecem suporte a colaboração; (ii) avaliar a capacidade de interoperabilidade entre os serviços desenvolvidos utilizando o processo.

Agradecimentos

Aos participantes do Núcleo de Pesquisa em Engenharia do Conhecimento (NEnC). Este projeto é parcialmente financiado pela CAPES, CNPq e FAPEMIG. Guilherme G. Martins é apoiado pela CAPES.

Referências

- Basili, V. R., Caldiera, G. and Rombach, H. D. (1994) “The goal question metric approach”, Software Architecture. Working IEEE/IFIP Conference.
- Belloum, A., Inda, M. A., Vasunin, D., Korkhov, V., Zhao, Z., Rauwerda, H., Breit, T. M., Bubak, M. and Hertzberger, L. O. (2011) “Collaborative e-Science Experiments and Scientific Workflows”. *Internet Computing, IEEE*, v.15, n.4, pages 39-47.
- Costa, G. C. B., Braga, R., David, J. M. N. and Campos, F. (2015) “A scientific software product line for the bioinformatics domain”. *Journal of Biomedical Informatics*, v.56, n.C, pages 239-264.
- Freitas, V., David, J. M. N., Braga, R. and Campos, F. (2015) “Uma Arquitetura para Ecossistema de Software Científico”, In: IX Workshop em Desenvolvimento Distribuído de Software, Ecossistemas de Software e Sistemas-de-Sistemas (WDES 2015), CBSOft, Belo Horizonte, MG, pages 41-48.
- Fuks, H., Raposo, A. B. and Gerosa, M. A. (2003) “Do Modelo de Colaboração 3C à Engenharia de Groupware” In: WEBMIDIA 2003, Simpósio Brasileiro de Sistemas Multimídia e Web, Trilha especial de Trabalho Colaborativo Assistido por Computador, Salvador, BA, pages 445-452.
- Gadelha, B., Nunes, I., Fuks, H. and Lucena, C. J. P. (2009) “An approach for developing groupware product lines based on the 3C collaboration model”, In: 15th international conference on Groupware, Berlin, Germany, pages 328-343.
- Gimenes, I. M. S. and Travassos, G. H. (2002), “O Enfoque de Linha de Produto para Desenvolvimento de Software”, In: XXI Jornada de Atualização em Informática (JAI), Evento Integrante do XXII Congresso da SBC, pages 1-31.
- Gurp, J. van, Bosch, J. and Svahnberg, M. (2001) “On the notion of variability in software product lines”, In: Software Architecture, 2001. Proceedings. Working IEEE/IFIP Conference, pages 45-54.
- Kang, K. C., Lee J. and Donohoe, P. (2002) Feature-Oriented Product Line Engineering. In *IEEE Software*, v. 19, n. 4, pages 58-65.

- Laitenberger, O. and Dreyer, H.M. (1998). “Evaluating the usefulness and the ease of use of a web-based inspection data collection tool”. In: Software Metrics Symposium, pages 122-132.
- Martins, G. G., Neiva, F. W., David, J. M. N., Braga, R. M. and Campos, F. C. A. (2015) “Colaboração no Ciclo de Vida de um Experimento Científico no Contexto de um Ecossistema de Software”, In: XII Simpósio Brasileiro de Sistemas Colaborativos, Salvador, BA, pages 114-121.
- Mattoso, M., Werner, C., Travassos, G. H., Braganholo, V. and Murta, L. (2008) “Gerenciando experimentos científicos em larga escala”, In: Seminário Integrado de Software e Hardware, Belém. SEMISH, Congresso da SBC, pages 121-135.
- Miranda, G. C. V., Souza, J. A. B. V., Braganholo, V. and Oliveira, D. (2014) “CollabCumulus: Uma Ferramenta de Apoio à Análise Colaborativa de Proveniência em Workflows Científicos”, In: XI Simpósio Brasileiro de Sistemas Colaborativos, Curitiba, PR, pages 94-101.
- Neiva, F. W., David, J. M. N., Braga, R. M., Campos, F. C. A. and Freitas, V. (2015) “PRIME: Pragmatic Interoperability Architecture to Support Collaborative Development of Scientific Workflows”, In: IX Brazilian Symposium on Components, Architectures and Reuse Software, Belo Horizonte, MG, pages 50-59.
- Zhang, J., Kuc, D. and Lu, S. (2014) Confucius: A Tool Supporting Collaborative Scientific Workflow Composition. In *IEEE Transactions on Services Computing*, v. 7, n. 1. pages 2-17.