

Uso de experimentação contínua para a instrumentação e o empacotamento do coorte PIPA UFRJ

Luciana Nascimento¹[0000-0002-0129-4042], Cecília Apa²[0000-0003-1588-709X], Diego Cerqueira¹[0000-0003-4861-3394], Michael Rasso⁴[0000-0002-2857-2671], Vladimir Erthal¹[0000-0002-0637-3479], Vitor Maia¹[0000-0001-9999-0195], Thatiana Fernandes³[0000-0001-5447-7531], Guilherme Horta Travassos¹[0000-0002-4258-0424]

¹ PESC/COPPE-UFRJ, Brasil

{lnascimento, dc, vladimirerthal, ght}@cos.ufrj.br, vcm@ufrj.br

² Fing/UdelaR, Uruguay
ceapa@fing.edu.uy

³ IESC-UFRJ, Brasil
thatianafernandes@iesc.ufrj.br

⁴ Avignon Université, França
michael.rasso@alumni.univ-avignon.fr

Abstract. A experimentação contínua permite apoiar a tomada de decisão no desenvolvimento de sistemas de software contemporâneos. Seu uso aparenta trazer vantagens quando a presença dos clientes e a definição de requisitos não ocorre de forma direta. Atualmente, diferentes tipos de sistemas de software apresentam desafios em sua construção, considerando a necessidade de encapsulamento de conhecimento, integração com dispositivos no ambiente e o acompanhamento dos eventos ocorridos durante sua utilização. Esse é o caso de software para apoiar estudos de coorte (um tipo de estudo longitudinal), os quais evoluem continuamente ao longo de seu ciclo de vida. Este artigo apresenta os resultados preliminares e reflexões sobre o uso de experimentação contínua na instrumentação e empacotamento de um estudo de coorte, o PIPA UFRJ, na área de saúde ambiental coletiva. Os desafios que observamos, relacionados com aspectos importantes no desenho experimental, ainda não têm sido tratados de forma mais detalhada na literatura. Acreditamos que tais desafios representam oportunidades de pesquisa e de contribuição para a utilização dos princípios de experimentação contínua em projetos de software para domínios de problema não convencionais.

Keywords: Continuous Experimentation, Longitudinal Study, Cohort, Laboratory Package, Environmental Health, Experimental Software Engineering.

1 Introdução

A experimentação contínua consiste no uso do método científico para a avaliação sistemática de ideias e opções de solução em projetos de software. Tais opções podem

ser vistas como hipóteses que precisam ser avaliadas com os usuários finais para inferir qual delas provoca o efeito esperado na aceitação do produto de software [1]. Encontramos abordagens genéricas, como o *Lean Startup* [2], difundindo estes princípios para a prática, principalmente em contextos de desenvolvimento onde é necessário continuamente evoluir o software para aumentar sua percepção de valor para os usuários. Assim, sistemas contemporâneos e intensivos em software são desenvolvidos de forma incremental e evolutiva, tendo as suas características definidas pelo conhecimento obtido do retorno dos usuários e/ou pela observação do software em uso [3].

A construção de software com o uso de experimentação contínua exige a disponibilidade de infraestrutura computacional adequada, que permita apoiar a inserção e entrega dos tratamentos previstos para a coleta de dados dos eventuais efeitos de uma ou outra opção. Propostas para organização de ambientes e atividades estão disponíveis para apoiar a experimentação contínua em engenharia de software [4] [5], sendo possível observar que existe grande cooperação entre a indústria e academia na investigação desta prática [1]. Notadamente, o uso de práticas e tecnologias associadas à experimentação contínua tem ocorrido, majoritariamente, em situações onde o uso do software envolve um volume considerável de usuários em testes A/B. Porém, diferentes contextos de desenvolvimento de software com alta incerteza podem se beneficiar de experimentação contínua mesmo sem ter usuários em larga escala, demandando que outros métodos de experimentação sejam aplicados.

Nesse artigo descrevemos as experiências iniciais e lições aprendidas com a utilização de experimentação contínua em um projeto de software contemporâneo e não convencional, a infraestrutura computacional para apoiar o estudo de coorte PIPA UFRJ. Buscamos aplicar o conhecimento de instrumentação e empacotamento previamente aplicados por nosso Grupo de Engenharia de Software Experimental da COPPE/UFRJ em estudos experimentais, combinados com práticas de experimentação contínua, a fim de criar uma solução para o PIPA UFRJ.

Esse artigo é composto de mais quatro seções além desta introdução. A próxima seção apresenta o conceito de estudo de coorte e descreve o projeto PIPA UFRJ. Na seção três é fornecida uma visão geral da solução projetada e as características do projeto que motivam a utilização de experimentação contínua. Um caso é descrito na seção quatro e os nossos primeiros resultados de pesquisa, que são os desafios observados para o uso de experimentação contínua, são descritos na seção cinco.

2 Domínio do Problema: O Estudo de Coorte PIPA UFRJ

Estudos de coorte são estudos de observação que segmentam os participantes da pesquisa no decorrer do tempo, potencializando as evidências. São mais comumente utilizados em estudos no domínio da saúde, em que os indivíduos possuem uma condição em comum e são classificados em expostos e não expostos à um tratamento cujo efeito é observado em sua saúde. Durante o tempo do estudo de coorte, que pode ser prospectivo ou retrospectivo, periodicamente são coletados dados dos participantes da pesquisa de forma a apoiar a análise da frequência de ocorrência dos supostos

efeitos do tratamento nos participantes [6]. Em engenharia de software o seu uso ainda é incipiente, embora algumas experiências iniciais tenham sido realizadas por pesquisadores da área [7]. Como limitação para a execução de um estudo de coorte temos seu alto custo financeiro e o risco da perda de participantes ao longo do seguimento.

O Projeto Infância e Poluentes Ambientais (PIPA UFRJ) é uma iniciativa pioneira no Brasil para a execução de um estudo de coorte de nascimentos de base hospitalar com o objetivo de avaliar a exposição de mães e crianças aos poluentes ambientais, sendo elegíveis para participação todas as crianças nascidas na Maternidade Escola da Universidade Federal do Rio de Janeiro (ME/UFRJ). Estudos científicos apontam para evidências de que exposições ambientais nos estágios iniciais do desenvolvimento humano (da concepção até a primeira infância) são determinantes para as condições de saúde e na influência para a ocorrência de doenças na vida adulta [8] [9]. Em várias partes do mundo são conduzidos este tipo de estudo, dentre os quais podemos exemplificar: o estudo The Norwegian Mother and Child Cohort Study (MoBa) [10] e o estudo The Maternal-Infant Research on Environmental Chemicals Study (MIREC)/Canada [11]. No projeto PIPA UFRJ, a avaliação da exposição ambiental será realizada pela análise de material biológico (metais/agrotóxicos/plastificantes) e monitoramento de marcadores ambientais.

O início do coorte de nascimentos do PIPA UFRJ está previsto para 2020, devendo ter a duração total de quatro anos durante os quais os participantes serão acompanhados por um grupo de pesquisadores em dez marcos de coleta de dados. O primeiro marco do estudo é o de abordagem de gestantes elegíveis da maternidade da UFRJ. As que aceitarem participar no projeto devem responder ao questionário de pré-nascimento e realizar a coleta de material biológico. O segundo marco é o de nascimento e compreende tanto a condução de um segundo questionário quanto a realização de coleta de materiais biológicos da mãe e da criança. A mãe ainda deve realizar exames referentes ao leite materno nos três próximos marcos: 1º mês, 3º mês e 6º mês de vida da criança. A partir do quinto marco a mãe será responsável apenas por responder aos questionários de acompanhamento. As coletas de material biológico e de dados antropométricos da criança ocorrerão até o final do quarto ano do projeto, nos seguintes marcos: nascimento, 1º, 3º, 6º, 12º, 18º, 24º, 36º e 48º mês de vida.

3 Uma solução para o PIPA UFRJ

Um piloto do coorte PIPA UFRJ foi realizado de setembro de 2017 até agosto de 2018, sem o uso de apoio computacional e tendo como foco o primeiro marco de coleta de dados. Os resultados do piloto têm sido usados como fonte de informação para o projeto de uma solução computacional semiautomatizada para o coorte PIPA UFRJ (apresentada na seção 3.1 a seguir). As incertezas decorrentes da aplicação de processos tradicionais de desenvolvimento aliados a experiência dos desenvolvedores com experimentação em engenharia de software direcionaram ao uso de experimentação contínua para o desenvolvimento da solução (discutido na seção 3.2).

3.1 Visão Geral de uma Solução para o PIPA UFRJ

O planejamento da solução, no primeiro momento, buscou compreender e estabelecer o fluxo de trabalho que pudesse ser automatizado (Figura 1), fornecendo ferramentas e instrumentos para captação das participantes elegíveis, aplicação dos instrumentos de pesquisa (questionários), coleta de dados biológicos, medição de dados antropométricos da criança e coleta de dados dos marcadores ambientais, organização dos dados do estudo em repositórios apropriados para sua análise, observação de requisitos de anonimidade e estruturação de dados, e integração e/ou exportação com ferramentas de análise estatística e de visualização de dados externas.

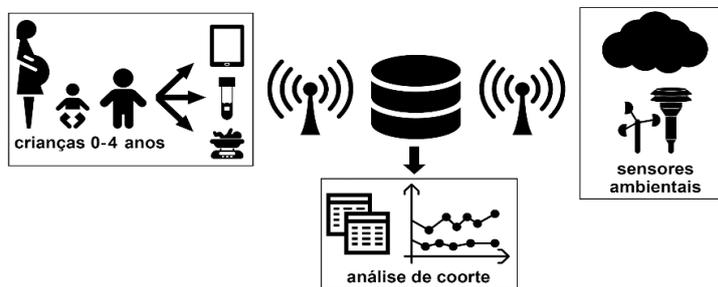


Fig. 1. Visão Geral de uma estratégia de solução para o PIPA UFRJ.

A solução proposta busca respeitar o direito à privacidade, como dispõe a Lei Geral de Proteção de Dados (LGPD) brasileira [12], que entra em vigor em agosto de 2020. Para garantir o direito à anonimização dos dados dos participantes, estes serão armazenados em bases de dados distintas, nomeadas como Controle Coorte e Dados Coorte. Na primeira base serão armazenados dados pessoais sensíveis e a segunda armazenará os dados coletados de forma anônima, constituindo a base longitudinal. O consentimento do armazenamento dos dados pelo participante, previsto em lei, é obtido através da assinatura de um Termo de Consentimento Livre e Esclarecido (TCLE).

3.2 Motivações para o uso da experimentação contínua no PIPA UFRJ

A solução proposta para o PIPA UFRJ, construída em processo de experimentação contínua, se apresenta como mais complexa que a construção de um sistema de gerenciamento de dados convencional. A principal característica do PIPA UFRJ que motiva o uso de experimentação contínua é a indisponibilidade de requisitos bem definidos para o software de apoio à instrumentação e ao empacotamento dos dados.

O coorte PIPA UFRJ é conduzido por diferentes pesquisadores em parceria com a Maternidade Escola da UFRJ, que recebe grávidas encaminhadas de diferentes unidades de saúde do município do Rio de Janeiro, fonte inicial de dados. Como resultado do piloto conduzido, os pesquisadores do PIPA UFRJ observaram a complexidade de estabelecer e orquestrar procedimentos e instrumentos de coleta e empacotamento de dados, preservando os dados para não perder participantes da coorte, e considerando os diferentes marcos do projeto que ocorrem para cada participante de forma paralela.

Assim, o processo de desenvolvimento para a solução PIPA UFRJ deve ser incremental e evolutivo, com algumas características de software a serem evoluídas conforme evidências da adequação de alternativas em uso e outras como resultado das futuras definições dos pesquisadores, visto que os mesmos também estarão em processo de aprendizagem durante a execução do estudo de coorte.

Observou-se também que o desconhecimento acerca do impacto que a eficiência de comunicação entre os diferentes atores do projeto (pesquisadores, técnicos de saúde, genitoras atendidas) é um risco ao estudo de coorte do PIPA UFRJ. Para evitar desistências de participação e garantir a qualidade das informações obtidas, é importante que os instrumentos dos questionários sejam fáceis de manusear e de entender e tenham tempo médio de execução razoável (inferior a 40 minutos) para a expectativa dos entrevistadores e das genitoras, considerando ainda que o grupo de entrevistadores pode mudar. O tempo de consumo e a qualidade da interpretação dos materiais coletados em exames ou dados antropométricos devem também ser avaliados.

Com isso, dentre os vários pontos que possibilitam a introdução da experimentação contínua ao longo do desenvolvimento da solução, envolvendo a instrumentação da coleta de dados dos participantes, temos decisões a serem tomadas sobre: (1) a lista de gestantes elegíveis, (2) o questionário de pré-nascimento, (3) o questionário de nascimento, e (4) o questionário de acompanhamento. A Figura 2 apresenta os materiais que serão obtidos em cada um dos marcos do coorte PIPA UFRJ e oportunidades de avaliação de características destes materiais. É importante notar que vários marcos podem ocorrer em paralelo no decorrer do projeto. Por exemplo, ainda no primeiro mês é possível que hajam casos dos marcos Gestação e Nascimento. Para cada ponto de experimentação existem diferentes opções de procedimentos e ferramentas a adotar para realizar o empacotamento dos materiais utilizados (objetos de estudo). A decisão sobre qual é a melhor combinação a utilizar depende das diferentes características de qualidade a serem observadas na experimentação, devido as múltiplas características de qualidade de um objeto de estudo.

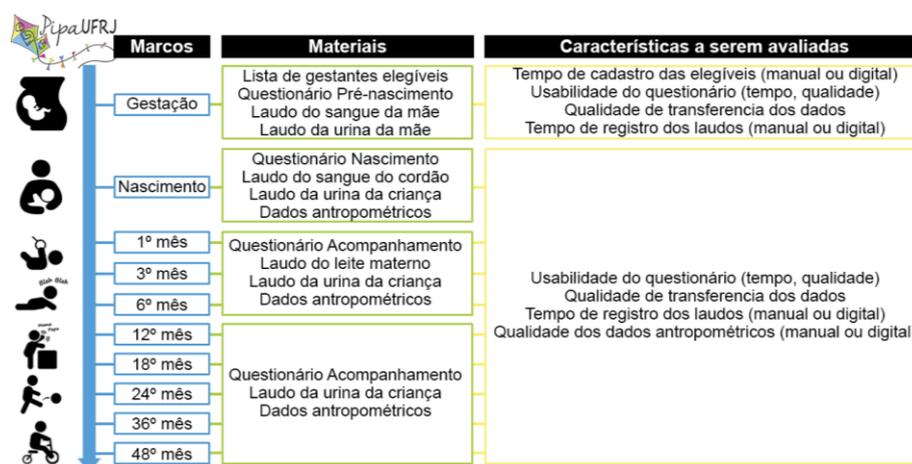


Fig. 2. Materiais e características a serem avaliadas a cada marco do Coorte PIPA UFRJ.

4 Caso de experimentação no PIPA UFRJ

Uma das possibilidades observadas para experimentação contínua foi o desenho de suporte ferramental para o questionário de pré-nascimento (Q1), que compõe a lista de materiais do primeiro marco. O Q1 é aplicado na primeira fase do coorte e coleta dados sobre o terceiro trimestre de gestação. Estrutura-se em 15 blocos, que contemplam dados da gestante (identificação, dados socioeconômicos, dados da gestação, hábitos, saúde), dados sobre o pai biológico e dados sobre a moradia.

O questionário foi modelado para permitir sua aplicação com as mães através do uso de *Tablets* pelos pesquisadores, visando facilitar a coleta dos dados e diminuir o tempo de abordagem do entrevistador com a gestante. A experimentação pode ser aplicada neste cenário, por exemplo, através da modelagem de diferentes formatos do questionário, para observar qual melhor se adapta ao tamanho e utilização dos *Tablets* e resultam em tempo de preenchimento em conformidade com o requisito do projeto (< 40 min). Para a primeira execução de experimentos, o instrumento foi desenvolvido utilizando a ferramenta *LimeSurvey*¹.

Diferentes alternativas de experimentação devem ser propostas visando melhorar as características de qualidade mais importantes. Para determinar as possibilidades de experimentação do Q1 (primeiro objeto de estudo) foi necessário entender quais características de qualidade deveriam ser avaliadas para o PIPA UFRJ, quais as prioridades de tais características e quais alternativas de instrumentação viabilizam a melhoria dessas características. Sob o ponto de vista do pesquisador do estudo de coorte, essas características são: tempo, qualidade do preenchimento do questionário, percepção da usabilidade, entendimento do questionário pela participante e capacidade de capturar mais e melhor informação. Também é preciso tratar outras características relacionadas à integração da solução, tais como a compatibilidade de migração dos dados coletados pelos questionários para um sistema de banco de dados (minimizar a perda de informação e semântica dos dados) e a integração da solução computacional com dispositivos construídos para coleta de dados com base no paradigma da Internet das Coisas (balança infantil, sensores ambientais, dentre outras).

Nesta etapa de planejamento da experimentação contínua, surgem questionamentos cuja discussão não foi encontrada nos relatos da literatura técnica que reportam experiências da utilização de experimentação contínua. Eles são relacionados a:

1. Quantidade de alternativas viáveis para avaliar: No estudo de coorte PIPA UFRJ é prevista a participação de 2000 gestantes (com base na execução do piloto) que serão abordadas nos primeiros cinco meses do projeto. Em média, 100 gestantes serão captadas por mês (assumindo uma distribuição normal). Do ponto de vista experimental, na execução de um experimento em um período de uma semana não é possível avaliar mais do que três alternativas em arranjo fatorial, já que a quantidade de observações por alternativa (25) não seria suficiente para garantir confiança nos resultados. É importante avaliar as alternativas propostas num período de tempo adequado para que a informação obtida permita a tomada de decisão e agregue valor ao coorte.

¹ *LimeSurvey* é uma plataforma de para aplicação de questionários online.

2. Definição das hipóteses e métricas associadas: As hipóteses do experimento precisam estar associadas com as características de qualidade a avaliar (variáveis independentes), tendo métricas e processos de medição para coletar a medida na condução do estudo. Algumas variáveis não apresentam complexidade nesse sentido (i.e., tempo de preenchimento), mas outras variáveis apresentam desafios em relação a definição da métrica e a coleta da medida (i.e., entendimento do questionário pela participante), sendo mais complexos os dados qualitativos. Outro desafio é como combinar as diferentes hipóteses para a tomada de decisão.

3. Arranjo experimental: Um aspecto a ser determinado é como as alternativas serão administradas entre os pesquisadores. No caso do PIPA UFRJ, a caracterização do pesquisador (quem aplica o questionário) é um aspecto importante a ser considerado. Por exemplo: existem pesquisadores que participaram do piloto e outros que não participaram? A experiência dos pesquisadores difere entre eles? Se alguma das respostas for sim, os resultados do experimento devem ser tratados de forma diferenciada (bloqueio) ou as alternativas devem ser administradas num grupo de pesquisadores com essas características balanceadas (balanço). Também é importante identificar outras causas que possam ter influência no resultado do experimento, como por exemplo, o perfil socioeconômico da mãe (pode influenciar no entendimento do questionário). Esses tipos de influências são “fatores de confusão” e é preciso identificá-los com o objetivo de minimizar ameaças à validade do experimento.

5 Considerações Finais: Desafios e Oportunidades de Pesquisa

A aplicação da experimentação contínua em domínios de problemas não convencionais não tem sido relatada na literatura técnica. Entretanto, acreditamos que é importante incluir essa discussão. Tomando o estudo de coorte PIPA UFRJ como um caso não convencional, no contexto de sistemas com usuários em pequena/média escala, ainda em desenvolvimento, e com tempo de avaliação de alternativas muito limitado, nos deparamos com vários desafios ainda não considerados e/ou relatados pela comunidade científica. Interessante observar que, mesmo no contexto de grande quantidade de usuários, a coleta de dados quantitativos através de testes A/B não garante que os mesmos contribuam para o direcionamento das decisões da organização de software. Um dos motivos é a falta de contextualização dos dados [13], que demandam a utilização de métodos para a coleta de dados qualitativos para entender a resposta dos usuários em relação as mudanças no software que lhe são propostas.

Os aspectos da aplicação da experimentação contínua com os quais lidamos, que não têm sido tratados na literatura técnica, e que acreditamos em sua importância são:

1. Determinação da quantidade de alternativas a avaliar tendo em vista a quantidade de observações por alternativa, o tempo máximo de execução do experimento, e o nível mínimo de confiança aceitável;
2. Adequação da definição das medidas a serem coletadas durante a execução dos experimentos com o processo de coleta e sua fidedignidade de representação em relação ao comportamento a ser observado;
3. Adequação na definição das hipóteses e viabilidade de sua combinação no momento da tomada de decisão; e

4. Utilização da caracterização do participante para identificação de fatores de confusão no planejamento do arranjo experimental.

Acreditamos que o aprofundamento e extensão das discussões dessas questões contribuem com o objetivo de apoiar, viabilizar, e encorajar a aplicação dos princípios de experimentação em engenharia de software e de práticas de experimentação contínua na construção de soluções para distintos domínios de problema.

Referências

1. Auer, F., Felderer, M.: Current state of research on continuous experimentation: a systematic mapping study. In : 44th Euromicro SEEA., pp.335-344 (2018)
2. Ries, E.: The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. Crown Business (2011)
3. Fagerholm, F. et al : Building blocks for continuous experimentation. In ACM : Proceedings of the 1st international workshop on rapid continuous software engineering., pp.26-35 (2014)
4. Fagerholm, F. et al .: The RIGHT model for continuous experimentation., 292-305 (2017)
5. Fabijan, A. et al.: The evolution of continuous experimentation in software product development: from data to a data-driven organization at scale. In IEEE-Press: Proceedings of the 39th ICSE, pp.770-780 (2017)
6. Blumenthal, U. J. et al.: Epidemiology - a tool for the assessment of risk.. In : Water quality: guidelines, standards & health. IWA publishing. (2001)
7. Fucci, D. et al.: A longitudinal cohort study on the retainment of test-driven development. In ACM, ed. : Proceedings of the 12th ACM/IEEE ESEM., p.p.18 (2018)
8. Barker, D. et al.: Fetal origins of adult disease: strength of effects and biological basis. International journal of epidemiology. 31(6), 1235-1239 (2002)
9. Gluckman, P. et al.: The fetal, neonatal, and infant environments - the long-term consequences for disease risk. Early human development 81(1), 51-59 (2002)
10. OLSEN, S. et al.: Possibilities and considerations when merging dietary data from the world's two largest pregnancy cohorts: the Danish National Birth Cohort and the Norwegian Mother and Child Cohort Study. Acta obstetricia et gynecologica Scandinavica. 93(11), 1131-1140 (2014)
11. Haines DA, A. et al.: Reporting results of human biomonitoring of environmental chemicals to study participants: a comparison of approaches followed in two Canadian studies. Journal of Epidemiology & Community Health, 191-198 (2011)
12. Instituto de Tecnologia e Sociedade: LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD) E SETOR PÚBLICO. (Acessado em 2019) Disponível em: <https://itsrio.org/wp-content/uploads/2019/05/LGPD-vf-1.pdf>.
13. Fabijan, A. et al.: Experimentation growth: Evolving trustworthy A/B testing capabilities in online software companies. Journal of Software: Evolution and Process. 30(12), e2113 (2018)