

**Universidade de São Paulo
Instituto de Matemática e Estatística**

Estudo de usabilidade utilizando eye tracking

Autora: Beatriz Sonzzini Ribeiro de Souza
Orientador: Carlos Hitoshi Morimoto
Co-orientador: Vagner Figueredo de Santana

**São Paulo-SP
Dezembro/2015**

Agradecimentos

Gostaria de agradecer ao professor Carlos Hitoshi e ao Vagner Figueredo de Santana por terem me orientado e coorientado e por terem me ajudado no decorrer do ano com os testes e com os equipamentos. Muito obrigada por todo esforço que me ajudaram a tornar o trabalho melhor.

Resumo

Testes de usabilidade são importantes pois quando é dada uma maior atenção aos aspectos de usabilidade, são menores os custos de manutenção e suporte, a vantagem competitiva é maior, entre outros. Para este trabalho foi utilizado um sistema de visualização de dados GRC (Dados de Governança, Riscos e Controles). Dados GRC são difíceis de serem interpretados e sumarizados para apoio à tomada de decisão; assim é necessário termos sistemas especializados para os usuários conseguirem compreender esses dados. Um sistema para esses dados foi desenvolvido e esse trabalho consiste no estudo da usabilidade desse sistema para validação das visualizações desenvolvidas. O estudo foi feito com o uso de um dispositivo para *eye tracking*, pois auxilia na identificação de quais locais os usuários estão olhando. No estudo de visualizações, observar os pontos que as pessoas passam os olhos e fixam o olhar é fundamental para estudos de usabilidade. Foram realizados testes com 4 especialistas em dados GRC. Entre as 5 visualizações estudadas, as que foram mais utilizadas e que despertaram maior interesse por parte dos participantes foram as seguintes: (1) visualização de taxonomias de GRC baseada em gráfico Sunburst; (2) visualização de riscos e controles usando grafos.

Abstract

Usability tests are important since as more usability aspects are considered, costs related to maintenance and support are reduced, advantages over competitors are gained, among others. In this work a visual analytics system for GRC (Governance, Risk, and Compliance) data has been used. GRC data are hard to be interpreted and summarized to support decision-making processes, thus there is a need to have specialized systems so that specialists can understand these data in a whole. A system for visualizing GRC data was developed and this study involves the usability test of it to validate the developed visualizations. The study was performed using an eye tracking device, since it helps identifying areas that the users are looking at. When studying visualizations, observed points where people perform gazes or fixations is fundamental. User tests performed counted on the participation of 4 GRC specialists. Considering the 5 visualizations developed, the ones that were most used and that participants had shown to be more interested with were the following: (1) GRC taxonomies visualization based on Sunburst visualization; (2) risk visualization using graphs.

Sumário

- [1. Introdução](#)
 - [1.1 Motivação](#)
 - [1.2 Objetivos](#)
 - [1.3 Organização](#)
- [2. Apresentação dos conceitos](#)
 - [2.1 GRC e sistemas GRC](#)
 - [2.2 Usabilidade e testes de usabilidade](#)
 - [2.3 Dispositivo de eye tracking](#)
 - [2.4 Heat map](#)
- [3. Trabalhos relacionados](#)
- [4. Experimento](#)
 - [4.1 Materiais utilizados](#)
 - [4.2 O sistema GRC utilizado](#)
 - [4.3 Como o teste foi realizado](#)
- [5. Resultados](#)
- [6. Conclusões e trabalhos futuros](#)
- [7. Referências](#)
- [8. Parte Subjetiva](#)

1. Introdução

1.1 Motivação

Usabilidade é um termo utilizado para medir a facilidade com que as pessoas podem usar uma ferramenta, um site, um sistema para realizar uma tarefa. Na área da Interação Humano-Computador (IHC) e na Ciência da Computação, usabilidade normalmente se resume à simplicidade e facilidade com que uma interface, um programa de computador ou um *website* pode ser utilizado.

Teste de usabilidade é uma técnica de pesquisa utilizada para avaliar um produto ou serviço através de testes com usuários. Com esse teste conseguimos identificar se os usuários estão tendo sobrecarga cognitiva, física ou perceptiva quando utilizam o sistema e fazer as mudanças necessárias.

Há várias maneiras de se realizar um teste de usabilidade. Um método interessante é testes utilizando dispositivos de *eye tracking*. Esse método acrescenta uma outra dimensão ao estudo pois observando o que as pessoas estão olhando obtemos um nível de dados a mais do teste que poderá ser utilizado para obtermos uma conclusão mais fundamentada sobre o experimento. Utilizando o dispositivo podemos observar os pontos que mais chamam a atenção na página.

GRC (*Governance, Risk and Compliance*) é uma tendência recente de integração das áreas de conhecimento de Gestão de Riscos, Governança Corporativa e práticas de auditoria e controle que busca garantir a conformidade com leis, regulamentos e imposições de padrões. Reflete um novo caminho para organizações integrarem áreas e atividades que normalmente manipulam as mesmas informações com objetivos distintos, como auditoria, controle dos riscos e conformidade.

Iremos fazer um estudo de usabilidade utilizando *eye tracking* sobre um sistema desenvolvido por uma empresa de TI que foi feito para a visualização de dados GRC. Esses dados são difíceis de serem interpretados e serem mostrados de maneira compreensível aos usuários por terem uma estrutura complexa e esparsa. Por esse motivo utilizaremos esse sistema para fazer o estudo com o *eye tracking*.

1.2 Objetivos

Testes de usabilidade com *eye tracking* são interessantes e mostraremos a utilidade de se usar esse dispositivo e como conduzir um teste com ele. Além disso, com o teste, verificaremos se o sistema desenvolvido possui uma boa usabilidade e se ele atendeu ao objetivo pelo qual foi desenvolvido, de mostrar os dados GRC de maneira clara e via múltiplas perspectivas. Mostraremos os resultados e conclusões obtidos pelo teste.

1.3 Organização

O trabalho foi organizado nas seguintes seções:

- Seção 1 - Introdução: Introdução ao trabalho realizado e suas motivações.
- Seção 2 - Apresentação dos conceitos: Explicação dos conceitos que foram utilizados no trabalho.
- Seção 3 - Trabalhos relacionados: Apresentação de trabalhos que já foram feitos e são relacionados com este trabalho e explicação da diferença deste trabalho para os outros.
- Seção 4 - Experimento: Desenvolvimento de todo o experimento, método e material utilizado, assim como o detalhamento do teste com usuários utilizando *eye tracking*.
- Seção 5 - Resultados: Apresentação dos resultados obtidos nos testes.
- Seção 6 - Conclusão e trabalhos futuros: Conclusão do trabalho, do teste e respostas sobre a usabilidade do sistema.
- Seção 7 - Bibliografia
- Seção 8 - Parte Subjetiva

2. Apresentação dos conceitos

2.1 GRC e sistemas GRC

Governance ou Governança aborda os processos, sistemas, controles pelos quais as organizações operam. *Governance* é normalmente administrada por governos, comumente inclui o exercício da autoridade legal e regulamentar, e isso considera o uso de recursos institucionais para gerenciar organizações (Tarantino, 2008).

Risk ou Risco se refere à possibilidade de perda ou prejuízo gerado por uma atividade ou por uma pessoa. O gerenciamento de riscos procura identificar, avaliar e medir o risco e assim desenvolver medidas para lidar com isso, normalmente minimizando o impacto. Riscos não devem ser vistos como algo inteiramente ruim, já que oportunidades vem com uma certa medida de risco. Os tipos de riscos variam dependendo do país de origem, indústria, nível de globalização e vários outros fatores (Tarantino, 2008).

Compliance ou conformidade envolve agir de acordo com leis estabelecidas, regulamentos, protocolos, padrões e especificações. Conformidade corporativa normalmente inclui conformidade com leis e regulamentações externas e protocolos internos (Tarantino, 2008).

A complexidade do mercado e as relações entre as organizações guiam a abordagem da tríade *Governance, Risk and Compliance* (GRC). Além disso, de acordo com Tarantino (Tarantino, 2008), GRC é tratado de uma forma holística, porque atacar os três separadamente é caro devido aos esforços duplicados e grandes chances de falhas nas três áreas.

Nesse contexto, o papel dos sistemas GRC é prover uma ferramenta integrada para combinar múltiplas fontes de dados, múltiplas funções, e múltiplas metas relativas a Governança, Risco e Conformidade. Desafios em tais sistemas envolvem estruturas de dados, relacionamentos entre estruturas diferentes (ex. riscos, controles, avaliações), e de prover múltiplas visualizações para as várias partes interessadas envolvidas no domínio GRC.

A questão crítica no contexto de GRC está relacionada ao custo de falhar, que pode ser civil, criminal, de reputação, financeiro, ou baseado em mercado. Por exemplo, recentemente a Petrobras sofreu um sério dano na imagem e investimento depois de uma série de casos de corrupção e um dos caminhos de mostrar para os investidores e governos que a companhia estava agindo foi anunciar uma nova composição no departamento de GRC.

2.2 Usabilidade e testes de usabilidade

Em projetos de conteúdos digitais um critério de qualidade é importante: a usabilidade. A usabilidade, de acordo com ISO 9241 1998, é a medida pela qual um produto pode ser usado por usuários específicos para alcançar objetivos específicos com efetividade, eficiência e satisfação em um contexto de uso específico. Ela aumenta a qualidade de um produto de conteúdo digital, e é um direito do consumidor.

Walter Cybis 2010 diz que aplicativos com baixa usabilidade sobrecarregam os usuários em três aspectos, sobrecarga cognitiva, perceptiva e física. A sobrecarga cognitiva acontece quando o usuário passa a não reconhecer, dentro da interface, os meios necessários para conseguir executar a sua tarefa com eficácia e eficiência. A sobrecarga perceptiva acontece quando o usuário encontra diante de si uma interface que, de alguma forma, prejudica a leitura dos seus itens. E a sobrecarga física acontece quando o acionamento de alguma função é prejudicado por um mecanismo muito complexo.

Os benefícios obtidos quando é dada uma maior atenção aos aspectos de usabilidade durante o ciclo do desenvolvimento de aplicativos são menores custos de desenvolvimento, redução do tempo de projeto, menores custos de manutenção e suporte, vantagem competitiva e fidelização do cliente.

A seguir temos a definição de testes de usabilidade de acordo com o site Usability.gov :

“Teste de usabilidade é uma técnica de pesquisa utilizada para avaliar um produto ou serviço. Os testes são realizados com usuários representativos do público-alvo. Cada participante tenta realizar tarefas típicas enquanto o analista observa, ouve e anota.”

Nos testes de usabilidade os usuários a serem recrutados devem ser pessoas que se encaixam no público-alvo do aplicativo avaliado, para que os resultados reflitam a possível realidade de uso do mesmo.

2.3 Dispositivo de *eye tracking*

O dispositivo de *eye tracking* monitora o movimento ocular das pessoas, permitindo saber exatamente para onde a pessoa olhou, quanto tempo e quantas vezes olhou, qual foi o caminho percorrido por seus olhos e mais. Devido a quantidade de informações geradas pela ferramenta, pode-se fazer várias análises.

Esse dispositivo funciona através de raio infravermelho dirigido aos olhos do usuário, que provoca reflexo na pupila. Esse reflexo é captado por um sensor que calcula as coordenadas para onde a pessoa está olhando. Esse dispositivo diz também se foi uma fixação dos olhos, quanto tempo fixou, etc.

2.4 *Heat map*

O *heat map* (mapa de “pontos quentes”) é um tipo de visualização que mostra concentração de fixações em áreas da tela. Essa visualização é uma das mais usadas em testes com *eye tracking* pois sumariza um grande volume de dados sobre as interfaces de usuários avaliadas (Pernice e Nielsen, 2009).

3. Trabalhos relacionados

Matsunaga et al. (data desconhecida) tratam da investigação dos movimentos dos olhos, suas características e a aplicabilidade da tecnologia de interação baseada na visão (*eye tracking*). Eles dizem que o uso do *eye tracking* para detecção dos movimentos dos olhos é de fundamental importância para compreendermos como funciona o sistema visual.

Granka et. al. (2004) investigaram como os usuários interagiram com a página de resultados de um sistema de busca web usando *eye tracking*. O objetivo era ter uma ideia de como os usuários navegavam os resumos e links apresentados e como eles selecionavam os links a fim de fazer uma maior exploração.

Salvucci et. al. (2000) propuseram uma taxonomia de algoritmos de identificação de fixação que classificam algoritmos em termos de como eles utilizam a informação do espaço e tempo em protocolos de *eye tracking*. Usando essa taxonomia, eles descreveram cinco algoritmos que representam classes diferentes e são baseadas em técnicas comumente empregadas. Com isso avaliaram e compararam esses algoritmos de acordo com o número de características qualitativas.

O departamento de psicologia da Universidade da Califórnia 2008 ofereceu um novo quadro que usa regressão logística multi-nível para analisar dados de experimentos de *eye tracking* usado em pesquisa psicolinguística.

Li D. et. al. (2006) desenvolveram um sistema que consiste em um design *open-hardware* para *eye tracker* digital que pode ser construído por componentes com baixo custo, e um conjunto de ferramentas de software *open-source* para capturar imagem digital e para manipulação e análise em aplicações *eye tracking*.

Poole e Ball (data desconhecida), primeiramente, introduzem o básico da tecnologia para detecção do movimento dos olhos, e apresentam também os aspectos chaves do guia prático para interessados em usar *eye tracking* para pesquisa. Eles examinam também vários caminhos em que o movimento dos olhos podem ser sistematicamente mensurado para examinar a usabilidade da interface.

Jacob e Karn (2003) descrevem as promessas da tecnologia de *eye tracking*, suas limitações, e os obstáculos que ainda devem ser superados.

Goldberg et. al. (2002) conduziram um estudo utilizando *eye tracking* para avaliar características de design específicas de um protótipo de uma aplicação de portal web. Este software serve o conteúdo da web independente através de *portlets* separados que podem ser modificados pelo usuário em uma página web.

Goldberg e Wichansky (2003) disponibilizaram um guia prático para o uso de *eye tracking* para avaliação da usabilidade.

Abaixo estão alguns sistemas desenvolvidos para dados GRC e suas principais características.

BWise GRC¹ oferece várias soluções baseadas em funções, permitindo assim departamentos a trabalharem com uma taxonomia de risco comum, focada em objetivos de negócios e com uma linguagem comum para auditoria interna, riscos e *compliance*. As principais funcionalidades oferecidas pelo BWise são as seguintes:

- Auditoria interna;
- Gerenciamento de riscos;
- Gerenciamento de conformidade e política;
- Controle interno;
- Tecnologia da informação (TI) GRC;
- Gerenciamento de desempenho de sustentabilidade.

Compliance 360² oferece módulos configuráveis que ajudam a identificar lacunas e riscos, eliminar esforços duplicados e manter os registros necessários para demonstrar o controle de conformidade, riscos e programas de auditoria. As principais funcionalidades oferecidas pelo Compliance 360 são as seguintes:

- Sistema de registro de conformidade, risco e gestão de auditoria;
- Gerenciamento de estruturas centralizado e distribuído para eficiência e controle;
- Mantém um estado de prontidão para auditorias, inquéritos e exames;
- Reage a mudanças nas leis, regulamentos, medidas de aplicação e de eventos externos;
- Aumenta a responsabilidade e a visibilidade do gerenciamento;
- Diminui processos e documentação de trabalho intensivos;
- Impõe padrões e consistência em toda a organização;
- Identifica e de forma proativa riscos em toda a empresa;
- Agiliza e direciona auditorias internas diretas para áreas de maior risco.

IBM OpenPages³ oferece um conjunto de serviços essenciais e componentes funcionais que estendem riscos e domínios de conformidade incluindo riscos operacionais, política e

¹ BWise GRC. <http://www.bwise.com/>

² Compliance 360. <http://www.compliance360.com/>

³ IBM Open Pages. <http://www03.ibm.com/software/products/pt/openpagesgrcplatform>

conformidade, gerenciamento de controles financeiros, governança de TI e auditorias internas. As principais funcionalidades oferecidas pelo IBM OpenPages são as seguintes:

- Adapta-se a metodologia de risco e gerenciamento de conformidade;
- Reduz os custos de desenvolvimento e implementação;
- Flexível e configurável;

MetricStream GRC⁴ oferece soluções para o gerenciamento de risco e atende às exigências normativas enquanto reduz os custos associados e utiliza módulos para dar suporte a relatórios, integração de dados e fluxo de trabalho. As principais funcionalidades oferecidas pelo MetricStream GRC são as seguintes:

- Integra risco, auditoria e programas de conformidade agregando e correlacionando dados de forma lógica;
- Estabelece governança estruturada, gestão de riscos e processos de conformidade que estão alinhados com os objetivos corporativos;
- Evolui à medida que o cenário dos negócios muda devido à sua tecnologia flexível;
- Permite atualização e modificação de módulos existentes;
- Fornece recursos de relatórios padrão e um assistente de relatórios para permitir a criação de relatórios ad-hoc;
- Adota arquitetura de TI aberta e tecnologia existente.

Outras ferramentas são OneSumX GRC, RSA Archer GRC, GRC SAP, SAS GRC.

As plataformas GRC apresentadas contam com um grande número de características, capacidades e benefícios fornecidos a organizações e várias partes interessadas. A seguir apresentamos uma lista de características comuns achadas:

- Abordagem integrada aos dados GRC;
- Flexibilidade de gerenciamento de dados GRC;
- Fácil gerenciamento de ferramentas;
- Arquitetura de plataforma modular;
- Plataforma extensível;
- Relatórios GRC padrão.

No entanto, as plataformas de GRC pesquisadas contam com as seguintes limitações:

⁴ MetricStream GRC. <http://www.metricstream.com/>

Os relatórios são geralmente centrados em risco e é difícil estender e incorporar visualizações além do padrão de relatórios gráficos de GRC (por exemplo, matriz de riscos, controles vs. riscos, gráficos representando taxonomias, resumos de classificação de risco).

A fim de conectar diferentes elementos apresentados na estrutura de dados geralmente presente em plataformas de GRC, consultas complexas normalmente são exigidas.

Considerando o *Visual Design Seeking Mantra* (Shneiderman, 1996), que sugere em primeiro lugar visão geral, zoom e filtro, e depois detalhes sobre encomenda, a plataforma que será estudada conta com um padrão de visualizações GRC para promover, principalmente, uma visão geral e aprimoramento das tarefas de filtragem e de consumo de conteúdo GRC detalhado a partir das análises visuais fornecidas.

4. Experimento

4.1 Materiais utilizados

O estudo conta com o sistema que será testado, o formulário do observador, o formulário de adesão do teste para ser assinado pelo participante, o *eye tracker* (Figura 1), um MacBook Pro com tela 13" usado para capturar os dados do *eye tracker*, e uma tela extra 19" usado pelos participantes. A figura 2 mostra a disposição da sala.

O formulário do observador conta com espaços para coleta dos dados durante o teste pelo observador, como por exemplo o horário de início de cada tarefa.

O dispositivo usado para capturar a interação dos olhos foi o *Eye Tribe eye tracker* (<https://theeyetribe.com/>) (Figura 1). Os dados capturados pelo *eye tracker* contam com coordenadas de cada olho, coordenadas média dos olhos, centro da pupila, tamanho da pupila para cada olho, e um valor booleano que indica se foi fixação ou não.

A visualização *heat map* mostra as áreas que os usuários olharam na tela. Os *heat maps* podem ser criados utilizando a duração da fixação (tempo que o usuário gasta olhando uma região específica) ou o número de fixações. Nesse estudo, o número de fixações foi considerado para tentar identificar todas as áreas que foram analisadas, mesmo quando ocorreu por um breve momento.



Figura 1. O *eye tracker Eye Tribe*.



Figura 2. Ambiente dos testes realizados

4.2 O sistema GRC utilizado

O sistema GRC utilizado no experimento foi desenvolvido pela IBM e é uma ferramenta nova que ainda não está no mercado. Esse sistema utilizado contribui com técnicas de recolha e manipulação de múltiplas perspectivas sobre dados GRC pouco conectados. Cada perspectiva conta com uma ou mais visualizações criadas para suportar a identificação de valores extremos, padrões visuais, clusters ou inconsistências no conjunto de dados predefinido. Além disso, considerando a dificuldade de suportar a decisão de tarefas que podem fazer parte das tarefas GRC, o sistema proposto também oferece várias visualizações para trabalhar sincronizadamente debaixo da mesma seleção de conjunto de dados, suportando filtragem e interação com os dados para várias visualizações.

Esse sistema foi desenvolvido para os gerentes corporativos de riscos. Ele trabalha com uma grande e tabular base de dados com centenas de tabelas contendo dados altamente relacionados, mas com poucos dados de risco financeiro.

A visualização do sistema oferece filtros que os clientes podem aplicar para explorar os dados. Algumas das visualizações fornecem vários componentes web de visualização na mesma página. Esses componentes são atualizados em tempo real quando os filtros são aplicados.

O sistema utiliza *tabs* para permitir várias visualizações e ao mesmo tempo usar um bom espaço na tela, enquanto preserva qualquer filtro mesmo com a mudança de *tab*.

Ao registrar os eventos que refletem mudanças no foco dos dados, nós podemos guardar a história que mostra como os dados foram explorados na tela quando as mudanças foram feitas.

As análises visuais desenvolvidas servem para diferentes propósitos.

Na figura 3 vemos um exemplo da estrutura dos dados normalmente encontradas em plataformas GRC, contando com uma ou mais taxonomias e camadas de riscos e controles conectados com os nós das taxonomias localizados nas folhas.

As figuras 4, 5, 6, 7 e 8 são imagens das telas que foram utilizadas para realizar o teste.

A figura 4 é a imagem da aba Risk Map. Essa visualização mostra a relação entre os riscos inerentes (eixo x) e a sua mudança (eixo y). Como exemplo, nesse gráfico para identificar uma boa situação, quando um risco inerente é muito alto e sua mudança é grande, basta observarmos o canto direito superior do gráfico, que indica que os círculos que estiverem nessa área tiveram controles aplicados efetivos. O Risk Map também coloca em destaque situações no qual classificação de riscos erradas foram introduzidas mostrando valores extremos no eixo y comparados com as localizações esperadas no gráfico dos grupos de risco.

A figura 5 é a imagem da aba Risk Clusters. Nessa visualização conseguimos ver os riscos (círculos maiores) e os controles aplicados a eles (círculos menores ligados aos círculos maiores). A cor dos círculos maiores representa a classificação desse risco, metade inerente e a outra metade residual. A cor dos círculos menores representa a classificação do controle, vermelho se ele não é efetivo e verde se é efetivo. Esse gráfico de força direcionada ajuda a visualizar padrões de riscos destacando situações com potencial de controles excessivos associados com riscos, facilitando a tomada de decisões sobre questões de controles redundantes ou a possibilidade de automatizá-los para salvar tempo e dinheiro.

A figuras 6a e 6b são as imagens da aba risk Overview. A visualização usada foi o design “sunburst”-- um tipo de diagrama de adjacência -- para resumir as posições de mitigação de risco para várias taxonomias, mostrando o risco inerente inicial assim como o risco residual restante baseado no antecipado impacto dos controles associados com riscos. Também está em destaque o impacto da mudança.

A figura 7 é a imagem da aba Sankey. Esse gráfico liga os riscos com suas entidades e com os controles que estão sendo aplicados a ele. É possível visualizar o relacionamento dos riscos individuais com suas taxonomias “pais”, como exemplo entidades de negócios, processos de negócios, e avaliação de riscos, assim como os controles usados para atenuar os riscos.

A figura 8 é a imagem da aba Control Treemap. Esse gráfico mostra os controles e se eles são efetivos ou não. O tamanho de cada quadrado pode ser representado pela data de criação do controle. Na área da legenda a direita o usuário escolhe entre representar os quadrados maiores pelos controles mais novos ou pelos mais velhos.

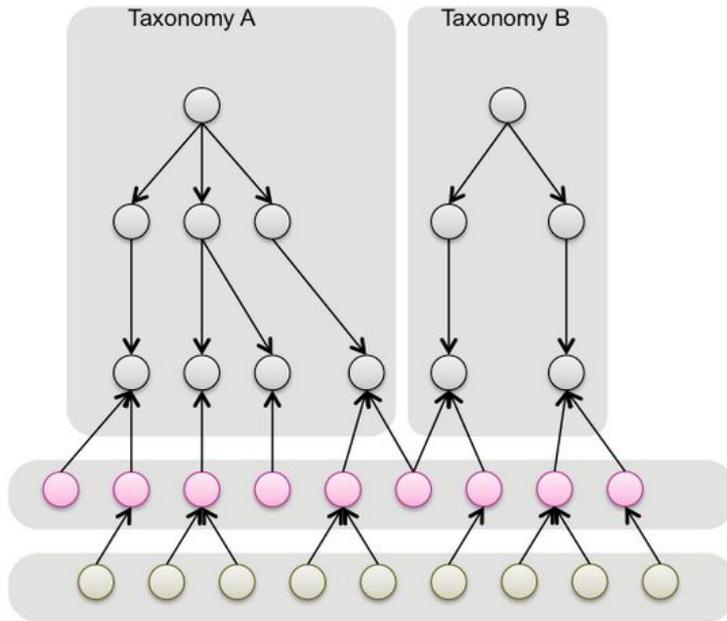


Figura 3. exemplo da estrutura dos dados normalmente encontradas em plataformas GRC, contando com uma ou mais taxonomias e camadas de riscos e controles conectados com esses nós das taxonomias localizados nas folhas

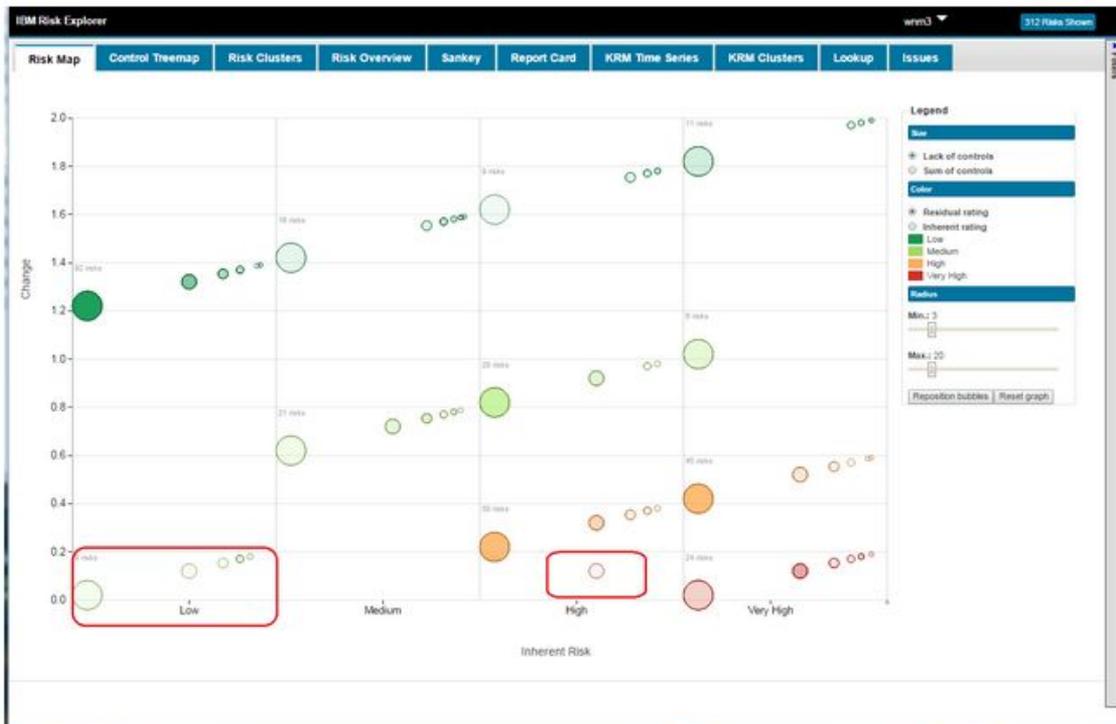


Figura 4. Aba Risk Map.

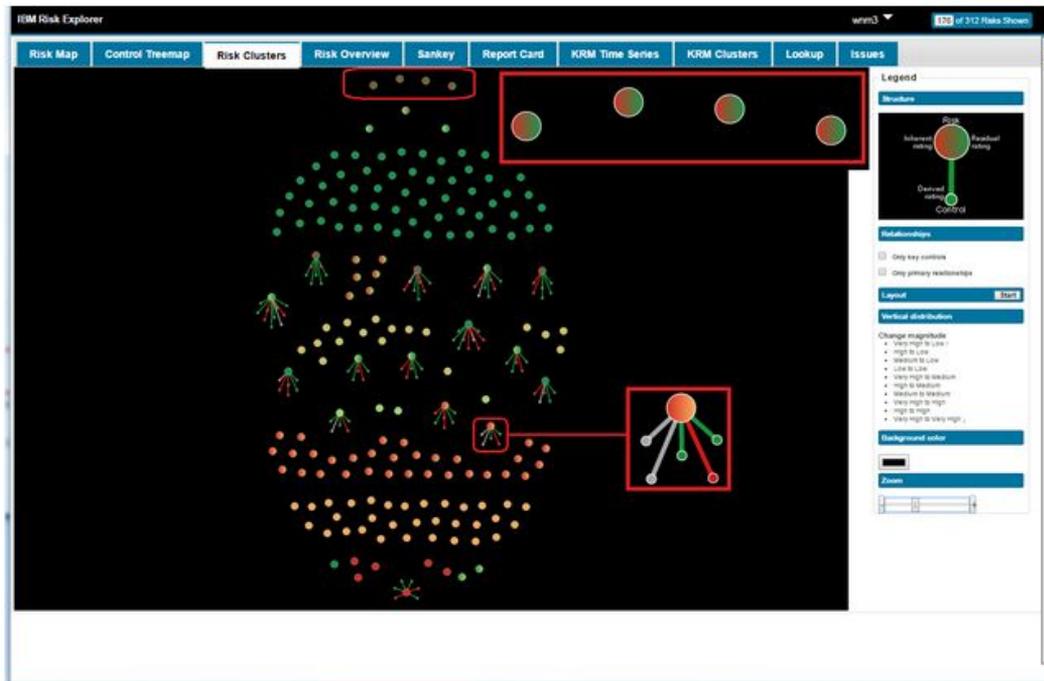


Figura 5. Aba Risk Clusters.



Figura 6a. Aba risk Overview e as legendas dos gráficos.



Figura 6b. Aba risk Overview.

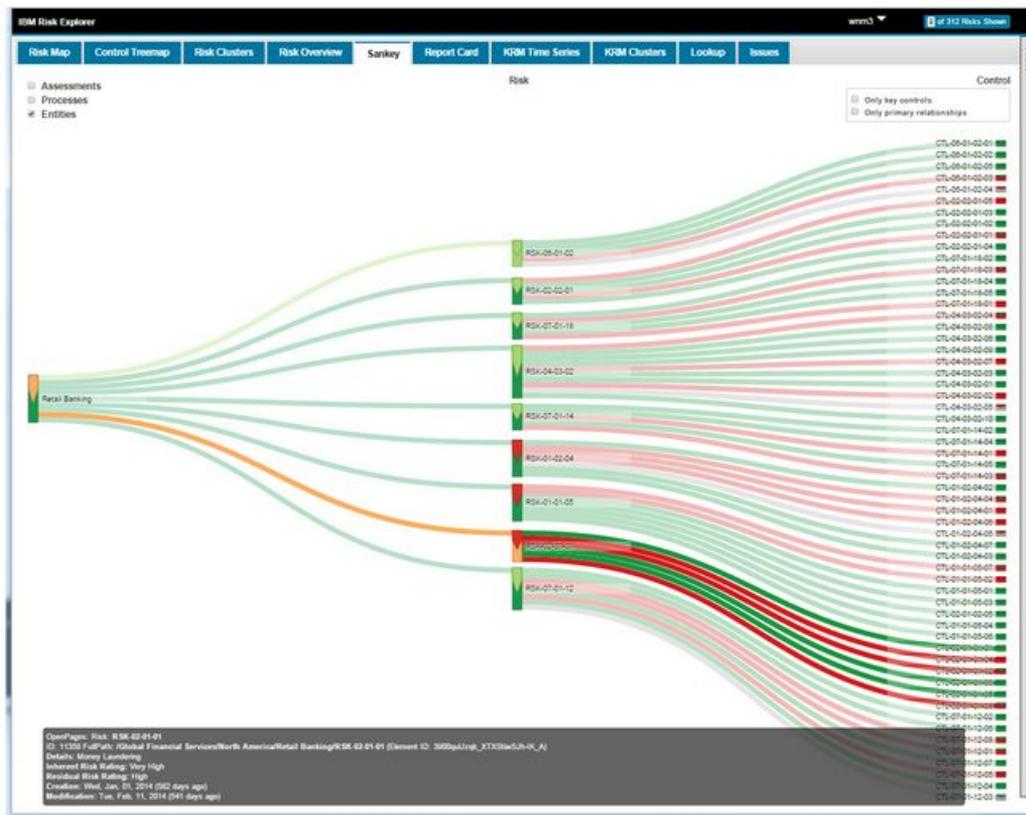


Figura 7. Aba Sankey.



Figura 8. Aba Control Treemap.

4.3 Como o teste foi realizado

Esse teste foi realizado para fazer uma análise qualitativa. Assim, ele foi feito com usuários que têm intimidade com dados GRC e as tarefas relacionadas a plataformas GRC. O teste foi feito com 4 pessoas que trabalham com esses dados. Fizemos um teste piloto para validar o plano de teste com uma pessoa que não era especialista em GRC e foi visto que muitas tarefas não faziam sentido para o participante do piloto, o que nos direcionou para realizar o teste apenas com profissionais da área de GRC. Dentro da empresa em que foram realizados os testes recebemos uma lista com 18 pessoas especializadas. Contatamos todas, mas apenas 6 pessoas aceitaram fazer parte do estudo. Destas 6, 5 agendaram um horário para a realização do teste, mas apenas 4 compareceram. Utilizamos um termo de consentimento (Figura 9) para ser assinado por cada um dos participantes. O convite feito por e-mail e aceito pelos 6 profissionais também perguntava sobre tarefas comuns executadas por eles no dia-a-dia. As respostas foram utilizadas para compor a lista de tarefas e definir a ordem entre os passos do teste de usuário.

Termo de Consentimento Livre e Esclarecido

Avaliação do [REDACTED]

Você está sendo convidado a participar de um estudo para avaliar o quanto o protótipo [REDACTED] ajuda em tarefas relacionadas à exploração de riscos, controles e estruturas de processos, entidades e *risk assessments*.

Neste estudo serão realizadas tarefas usando o [REDACTED]. A estimativa é que toda a avaliação levará aproximadamente 1 hora.

Não há riscos envolvidos nesse estudo e todas as dúvidas serão esclarecidas conforme solicitado. Em caso de dúvida, o pesquisador [REDACTED] poderá ser questionado em qualquer etapa do estudo.

A qualquer momento você poderá desistir da pesquisa, sem qualquer prejuízo a você.

As informações obtidas serão analisadas em conjunto com as informações resultantes de avaliações de outros participantes da pesquisa, não ocorrendo a identificação de nenhum participante.

Não existe despesa para os participantes da pesquisa. Também não há compensação financeira relacionada à sua participação.

As informações e dados coletados serão utilizados somente para a avaliação de como pessoas utilizam o [REDACTED].

Acredito ter sido suficientemente informado a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo de "Avaliação do [REDACTED]".

Eu discuti com o pesquisador [REDACTED] que conduziu a avaliação, sobre a minha decisão em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados e as garantias de confidencialidade e de esclarecimentos permanentes.

Ficou claro também que minha participação é isenta de despesas e compensações.

Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo.

Data ____/____/____

Assinatura do participante do estudo

Data ____/____/____

Assinatura do condutor da avaliação

Figura 9. Termo de consentimento.

Utilizamos também o formulário do observador (Figura 10) no qual este anota o horário do início da calibragem e das tarefas. Também foi utilizado para registrar observações importantes feitas durante o experimento e em qual momento ocorreu.

O experimento começou com o facilitador entregando o termo de consentimento para o participante assinar, esclarecendo eventuais dúvidas e depois assinando o termo também.

Depois o facilitador fez uma breve explicação de como o sistema funciona, deixou o participante explorar o sistema por cerca de 5 minutos e tirou eventuais dúvidas.

A próxima parte foi a calibragem do *eye tracker* e a gravação dos dados devolvidos por ele, assim como a gravação dos dados dos eventos, da tela e da voz.

Depois disso começa o experimento. O facilitador informou uma tarefa de cada vez para o participante e aguardava uma resposta por parte do participante ou uma eventual desistência. O experimento contou com 8 tarefas descritas a seguir:

1. Qual *assessment*/área/processo que chama mais sua atenção em relação a **maior risco inerente**?
2. Qual *assessment*/área/processo que chama mais sua atenção em relação ao **menor risco inerente**?
3. Qual *assessment*/área/processo que chama mais sua atenção em relação ao **maior risco residual**?
4. Qual *assessment*/área/processo que chama mais sua atenção em relação ao **menor risco residual**?
5. Identifique um risco que tem impacto em mais de uma estrutura (e.g., processo e *business entity*)?
6. Identifique um controle que é reutilizado em mais de um risco
7. Identifique um controle com problema de preenchimento.
8. Identifique um ou mais riscos com problema de preenchimento em relação ao rating.

Utilizamos o protocolo *Thinking aloud protocol* (Nielsen, 1993) em que o participante é incentivado a falar em voz alta o seu raciocínio para realizar a tarefa.

5. Resultados

De acordo com a Tabela 1 vemos que as cinco primeiras tarefas foram realizadas em menos de um minuto e poucos segundos. Nas três últimas tarefas os participantes demoraram mais para concluir, indicando que elas eram ou mais trabalhosas para realização ou mais complexas.

Na Tabela 2 vemos que as tarefas 5, 6, 7 e 8 tiveram um índice de acerto baixo, indicando a complexidade delas.

Tabela 1. Tempo de execução das tarefas

	Média	Desvio Padrão
Tarefa 1	1:07	1:07
Tarefa 2	0:44	0:35
Tarefa 3	0:36	0:32
Tarefa 4	0:31	0:26
Tarefa 5	1:07	0:50
Tarefa 6	2:39	1:21
Tarefa 7	2:44	1:54
Tarefa 8	3:03	1:41

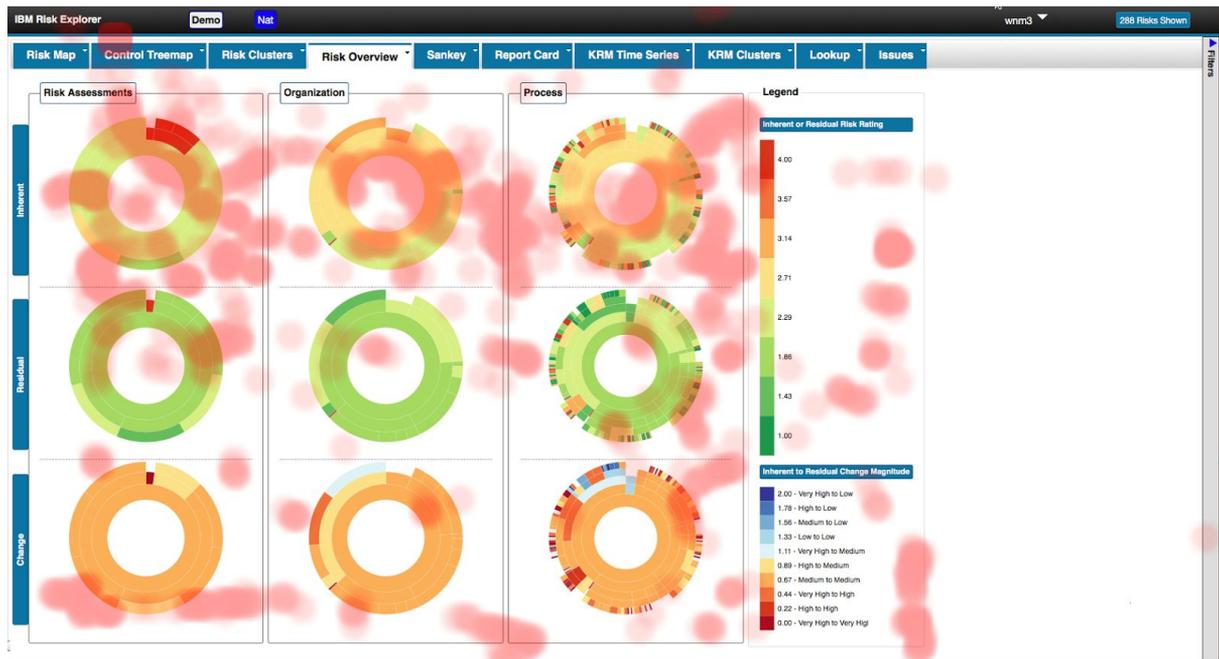
Tabela 2. Conclusão de tarefas

	Participante 1	Participante 2	Participante 3	Participante 4
Tarefa 1	Sim	Sim	Sim	Sim
Tarefa 2	Sim	Sim	Sim	Sim
Tarefa 3	Sim	Sim	Sim	Sim
Tarefa 4	Sim	Sim	Sim	Sim
Tarefa 5	Sim	Não	Não	Não
Tarefa 6	Sim	Não	Não	Não
Tarefa 7	Sim	Não	Não	Não
Tarefa 8	Sim	Sim	Não	Não

A seguir estão as imagens que mostram os resultados do teste e foram geradas com as telas e os dados colhidos com *eye tracking*. Os *heat maps* foram geradas usando Javascript e uma biblioteca em Javascript para criação de gráficos chamada D3⁵.

Participante 1

Tarefa 1



De acordo com o *heat map* vemos que o participante olhou para vários pontos da tela além do lugar onde estava a resposta (Primeira linha de sunbursts), mas conseguiu responder.

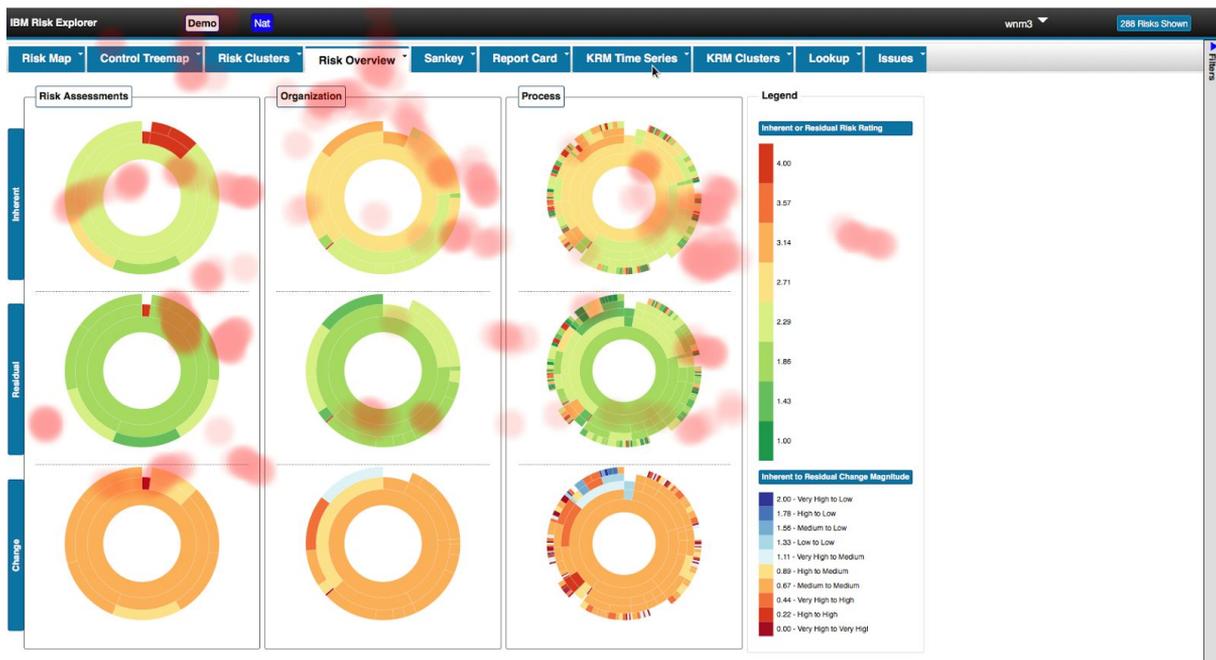
⁵ <http://d3js.org/>

Tarefa 2



A resposta estava na mesma região da tarefa anterior. Dessa vez o participante foi mais direto e olhou só para a região da resposta.

Tarefa 3



Assim como na tarefa 1 o participante olhou para vários pontos da tela além daquele onde estava a resposta (segunda linha de sunbursts), mas conseguiu responder.

Tarefa 4



A resposta estava na mesma região da tarefa 3. O participante foi mais direto e olhou só para a região da resposta.

Tarefa 5

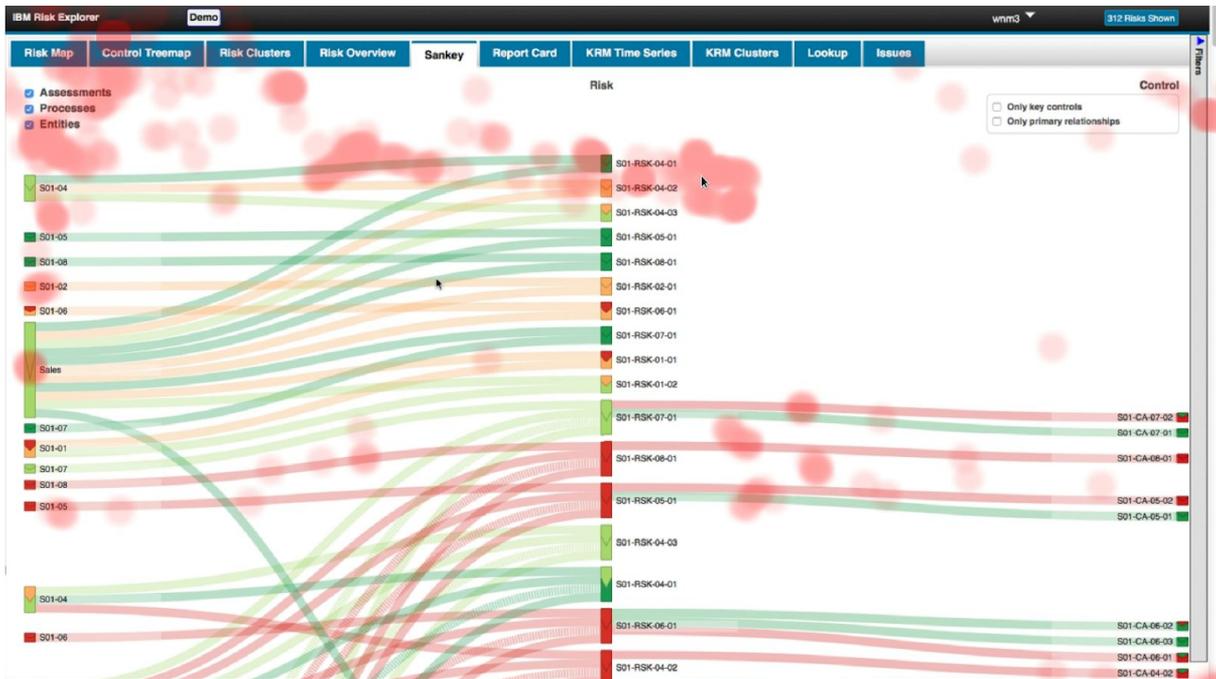


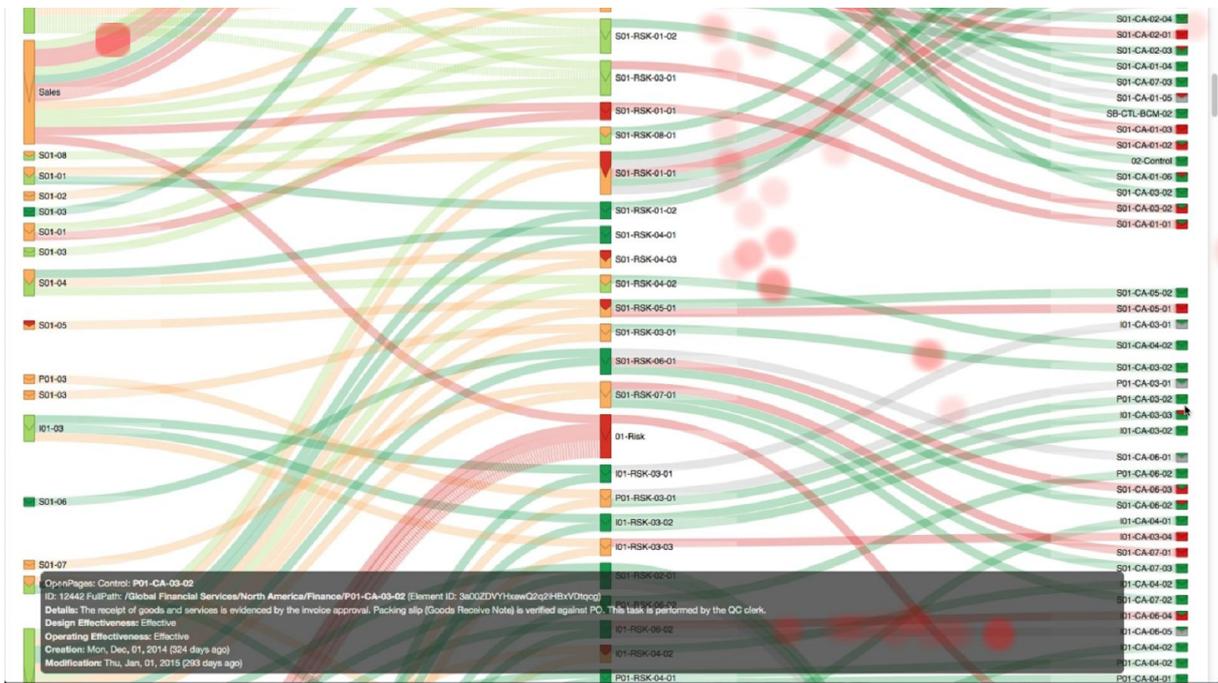
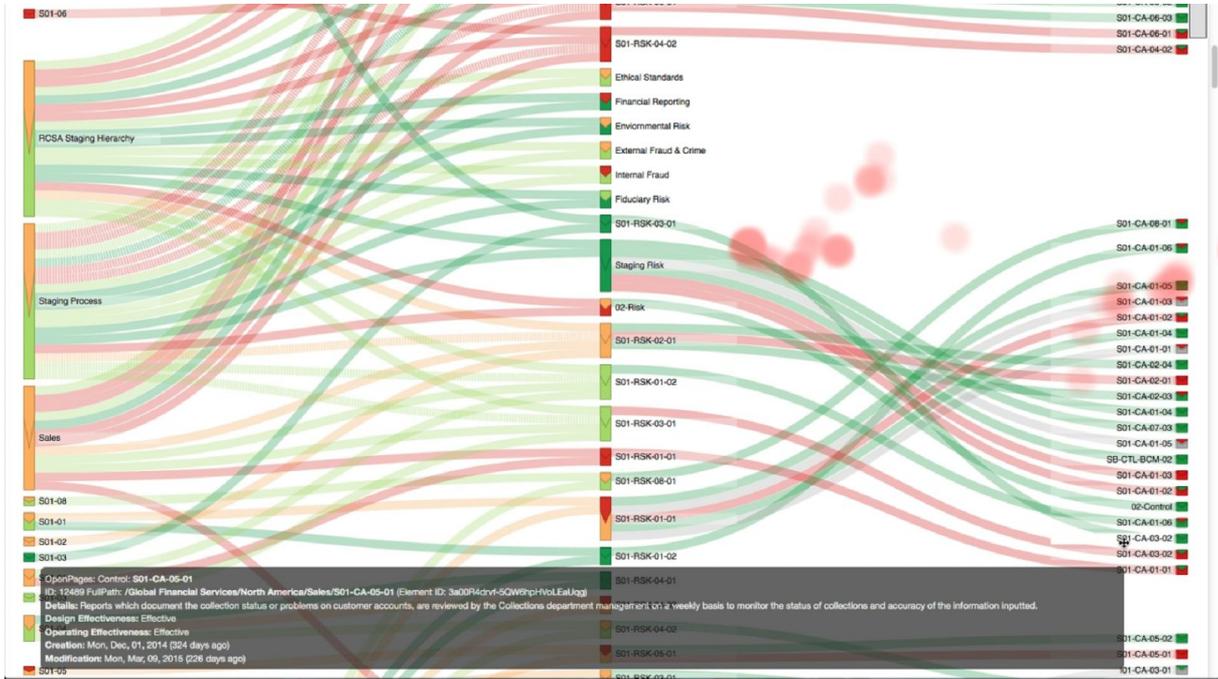


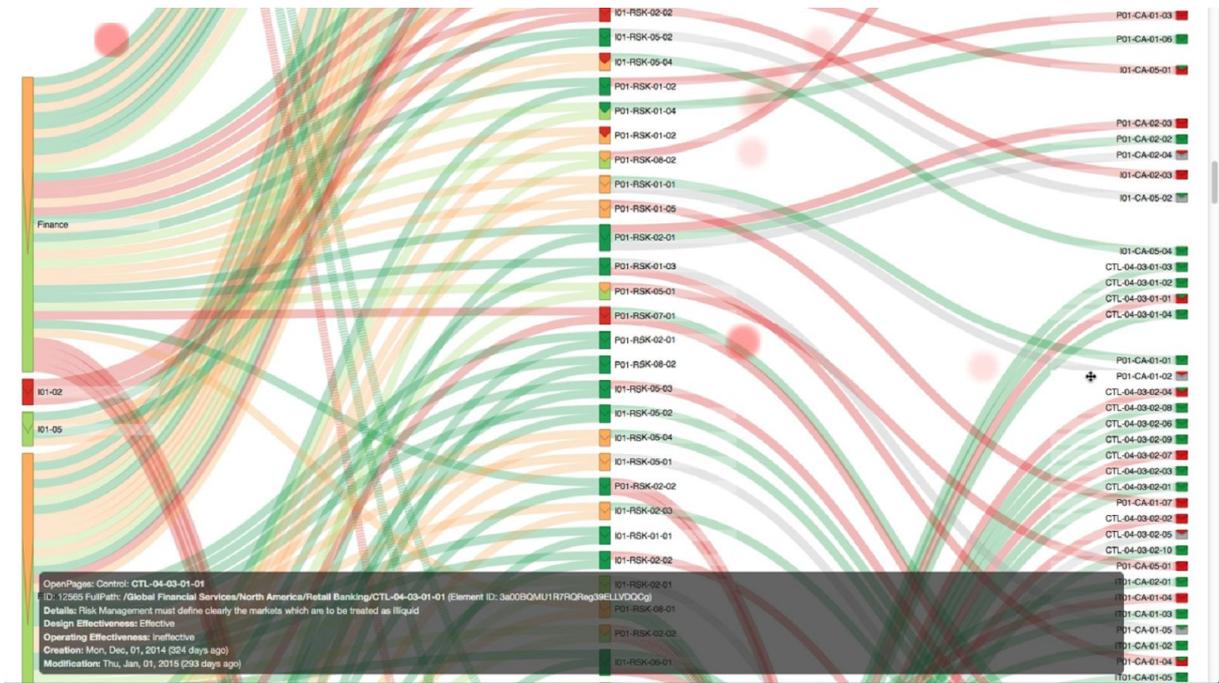
Antes de ir para a aba Sankey o participante procurou a resposta na aba risk Overview e não encontrou, na imagem vemos os pontos espalhados que mostram a busca pela resposta.

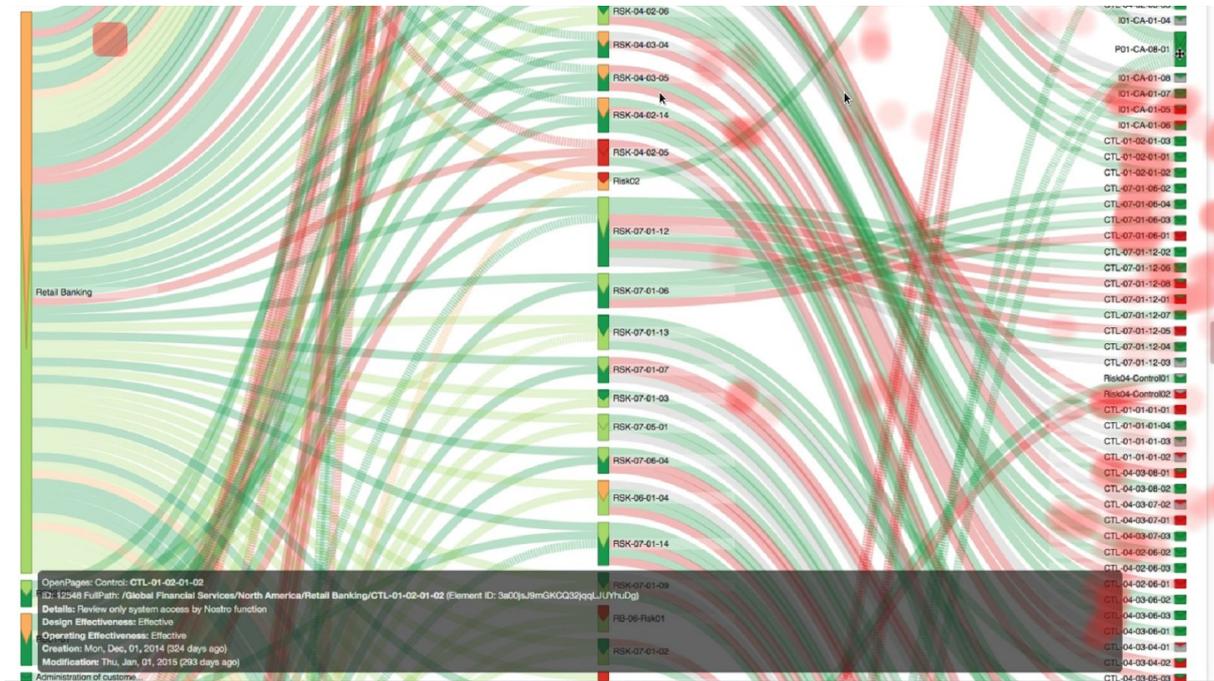
Depois o usuário mudou para aba Sankey. A maior parte dos pontos olhados está concentrada na coluna dos riscos, isso mostra que compreendeu esse gráfico, assim encontrou rapidamente a resposta.

Tarefa 6









Nas primeiras imagens vimos que o participante olhou mais para a coluna dos riscos em vez de olhar para a coluna dos controles. Logo depois ele perguntou qual era a coluna dos controles. O participante demorou um pouco para achar a resposta mas conseguiu achar e estava na região em destaque na imagem, na coluna da direita.

Tarefa 7



Procurou na aba Risk Overview e não achou. Teve dúvidas com relação à tarefa. Depois de tirar um dúvida com o facilitador, foi para a aba Risk Clusters e conseguiu encontrar rapidamente a resposta.

Tarefa 8

Procurou na aba Risk Clusters e não achou a resposta. Foi para a aba Control Treemap e utilizando filtros conseguiu encontrar a resposta rapidamente.

Preferências do usuário

O usuário gostou mais da aba Risk Clusters e em segundo lugar risk Overview e gostou menos da aba Risk Map.

Participante 2

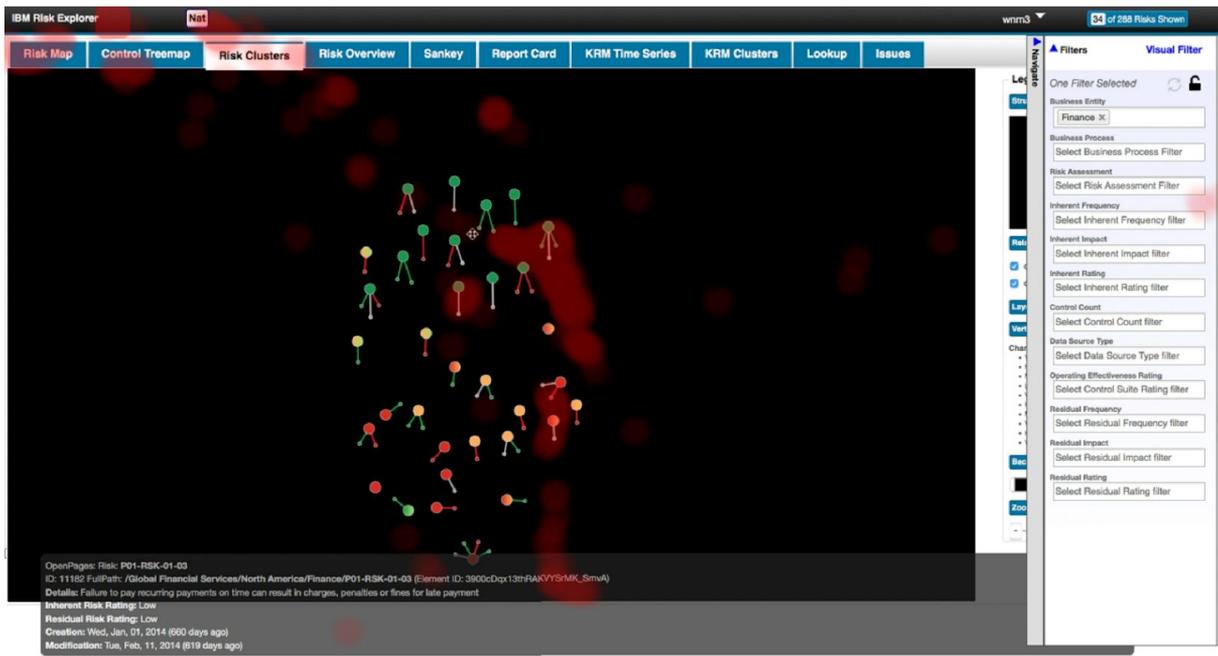
Tarefa 1

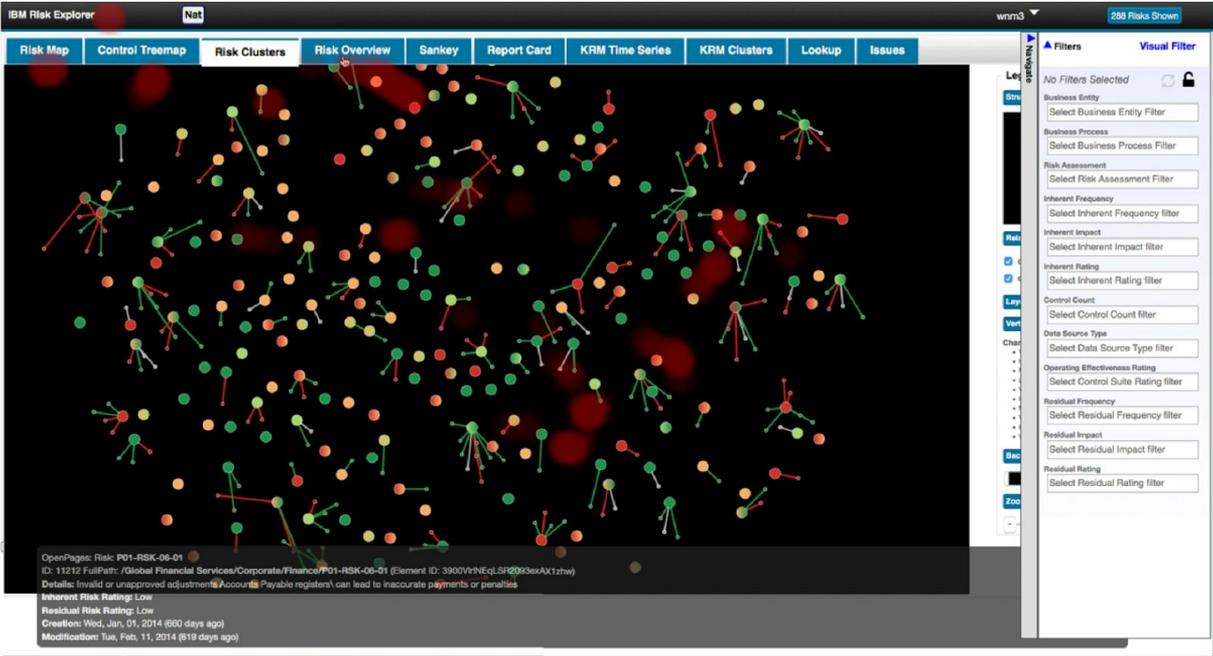


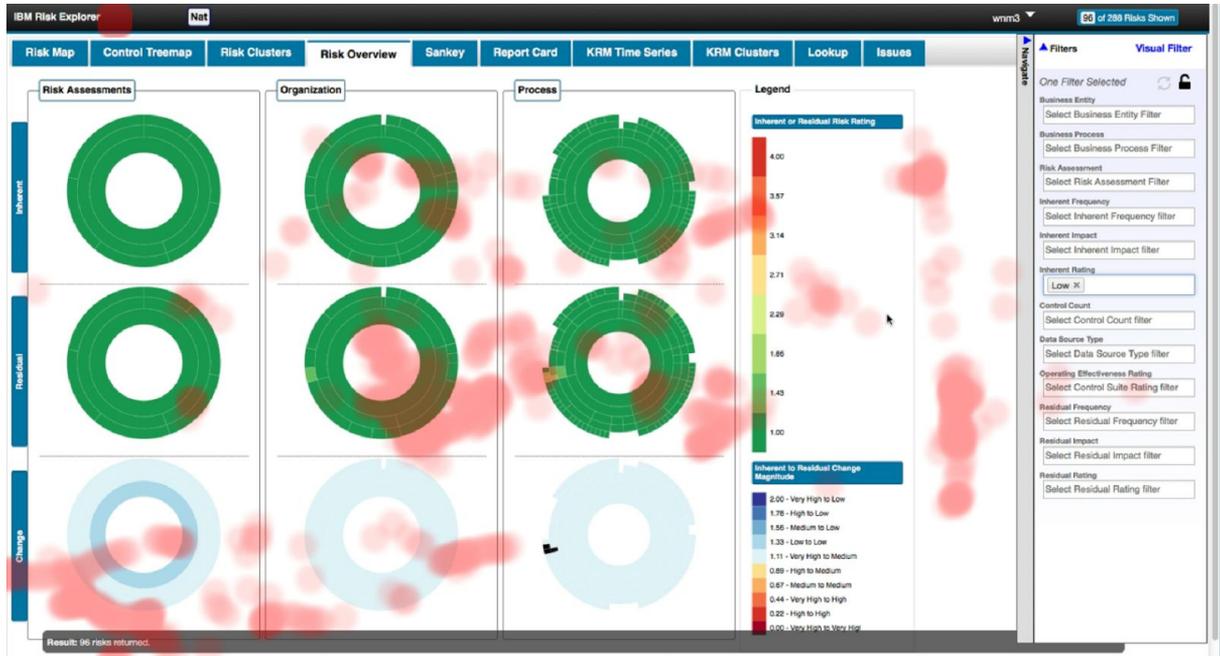


O olhar se concentrou na faixa dos riscos inerentes como esperado. O participante utilizou os filtros para encontrar a resposta. Depois de colocar os filtros deu para ver que o olhar se concentrou na região da resposta (primeira linha de sunbursts).

Tarefa 2



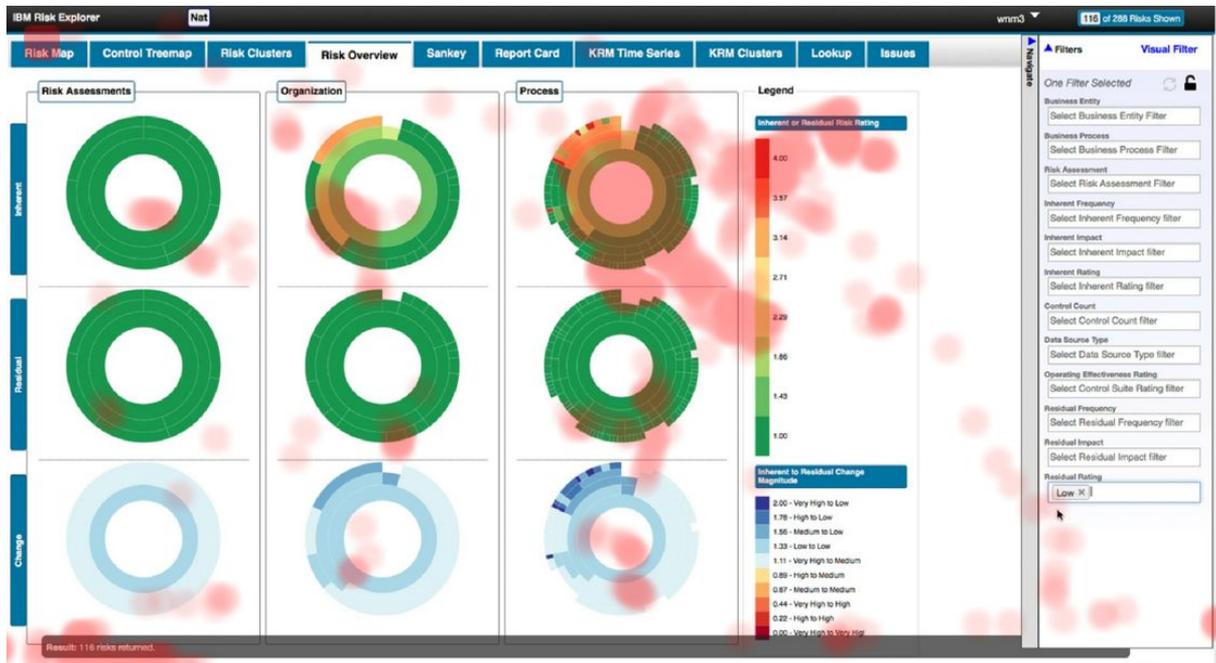




O participante utilizou os filtros. Olhou para áreas além das esperadas (a resposta estava na primeira linha de sunbursts), mas encontrou a resposta.

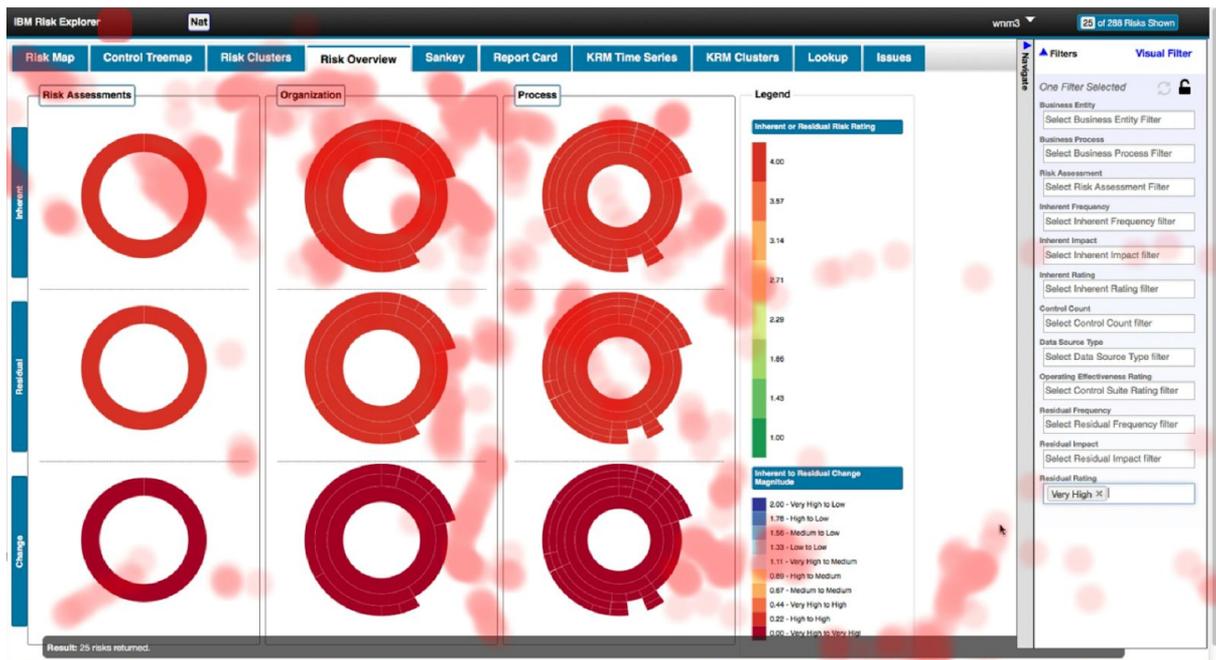
Tarefa 3





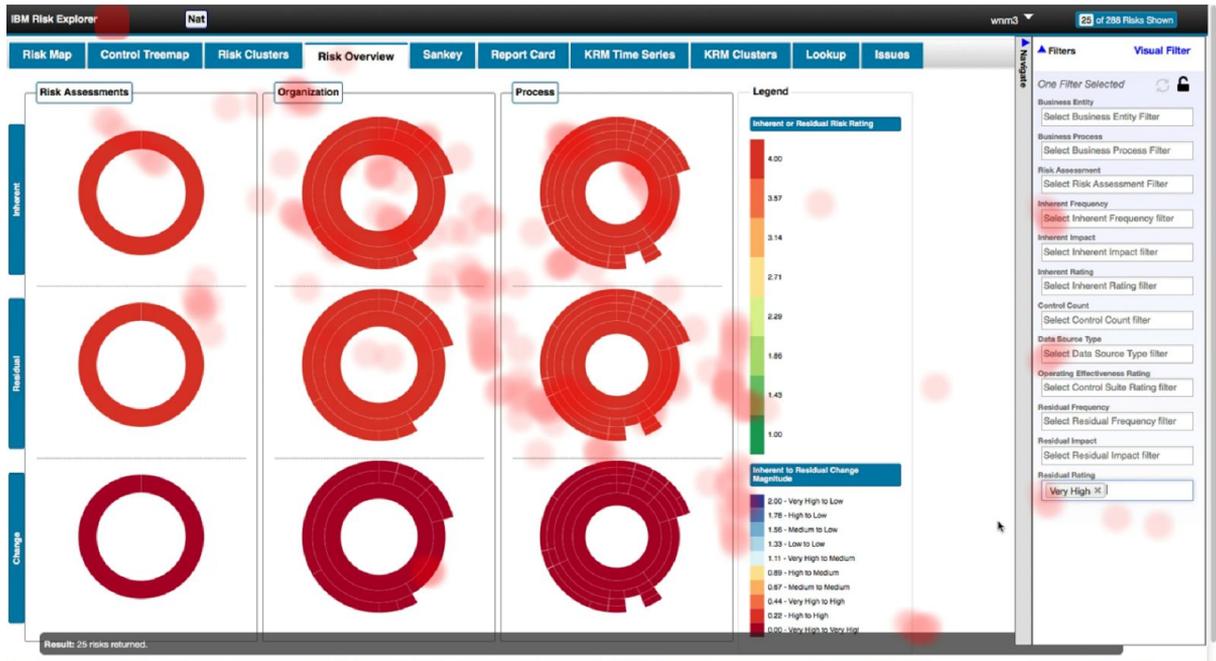
O olhar se concentrou na faixa dos riscos inerentes como esperado. O participante utilizou os filtros para encontrar a resposta. Depois de colocar os filtros observamos que o olhar se concentrou na região da resposta (segunda linha de sunbursts).

Tarefa 4



O participante utilizou os filtros para encontrar a resposta. Olhou para áreas além das esperadas (segunda linha de sunbursts), mas encontrou a resposta.

Tarefa 5

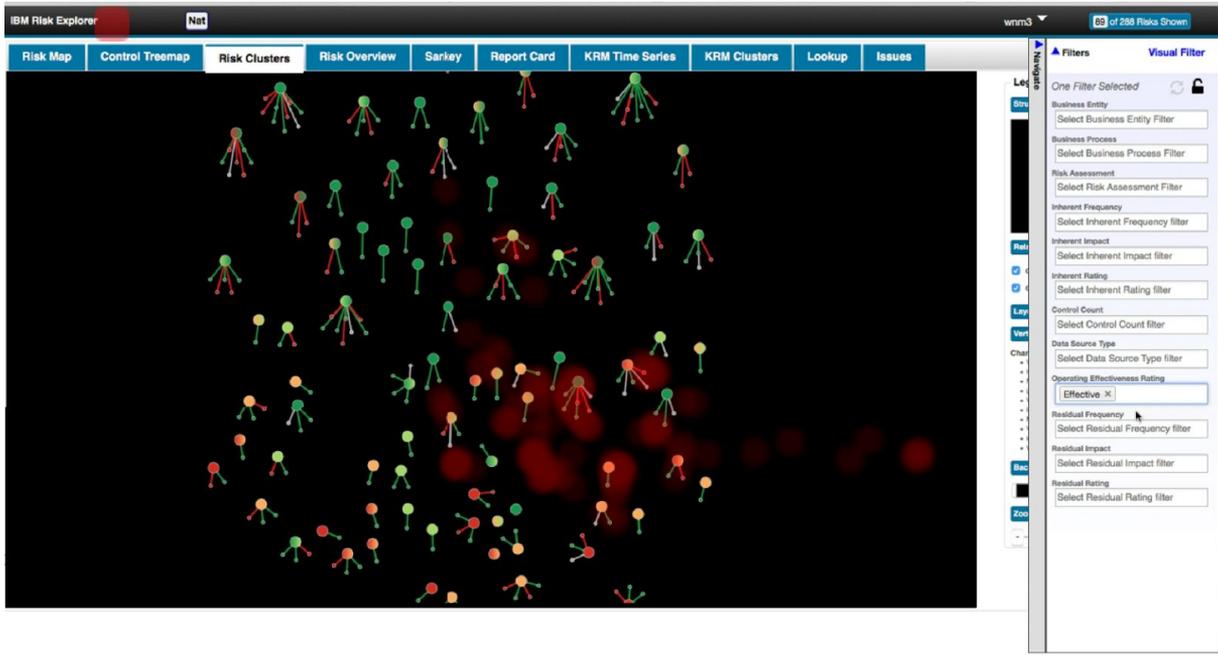
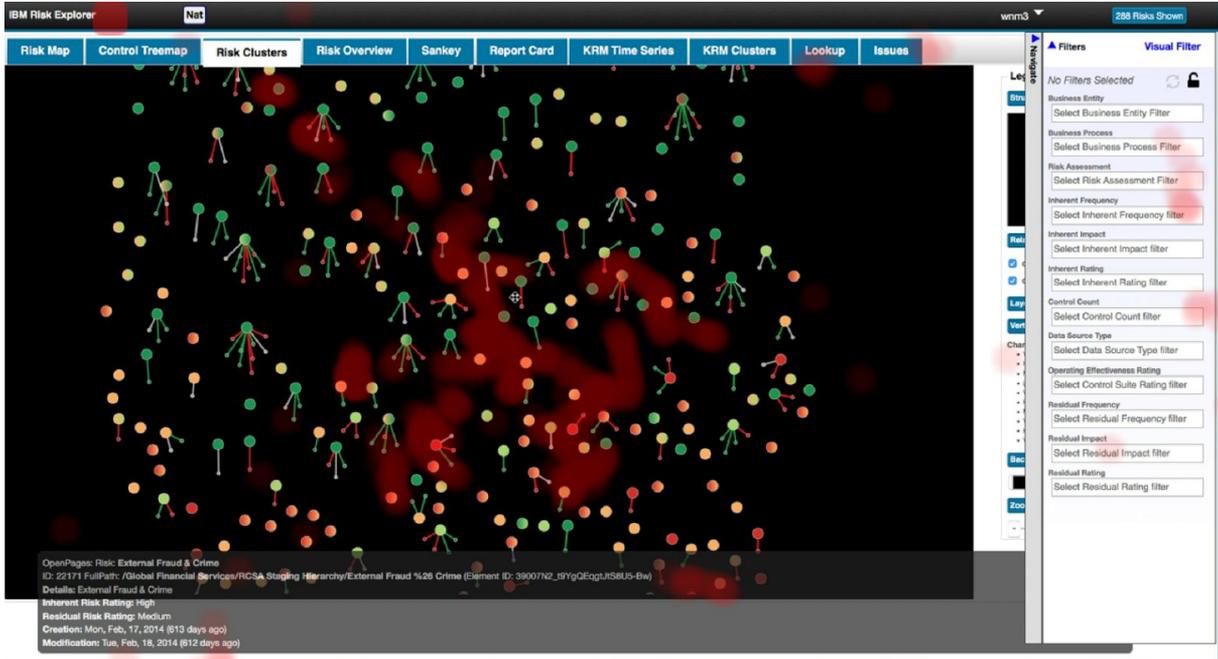


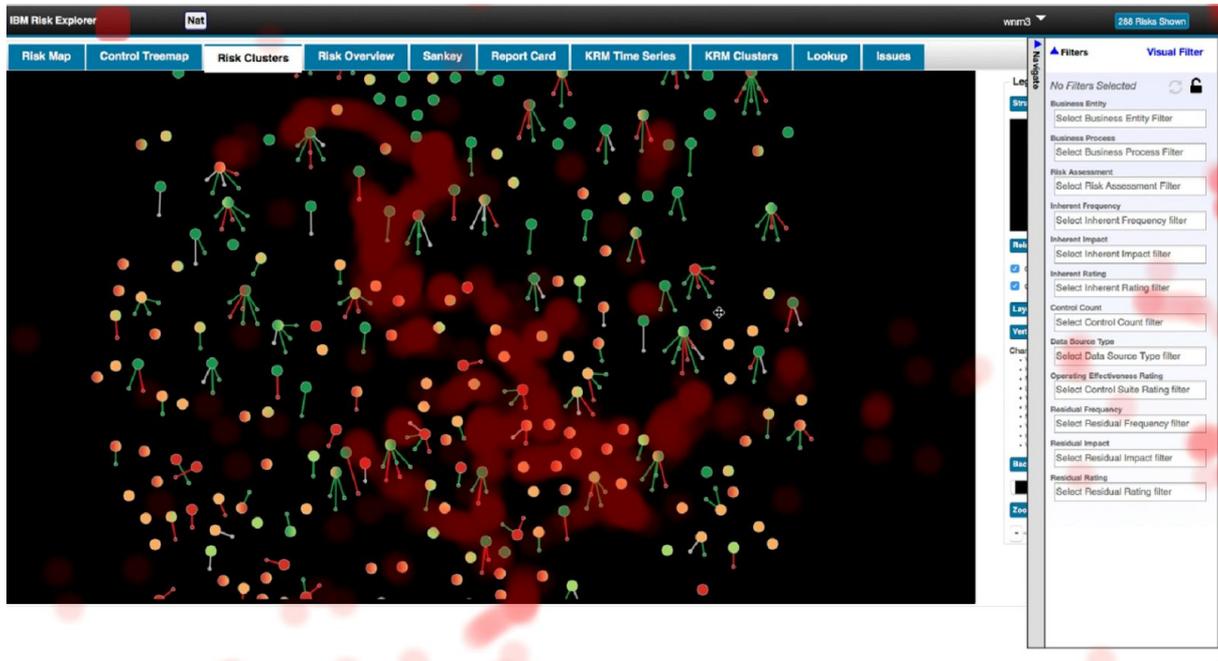


Tentou encontrar as respostas na aba risk Overview, depois buscou na aba Sankey. Não conseguiu responder à pergunta da tarefa. Pelas imagens conseguimos ver que temos pontos espalhados pelas telas o que indica a confusão para completar a tarefa.

Tarefa 6

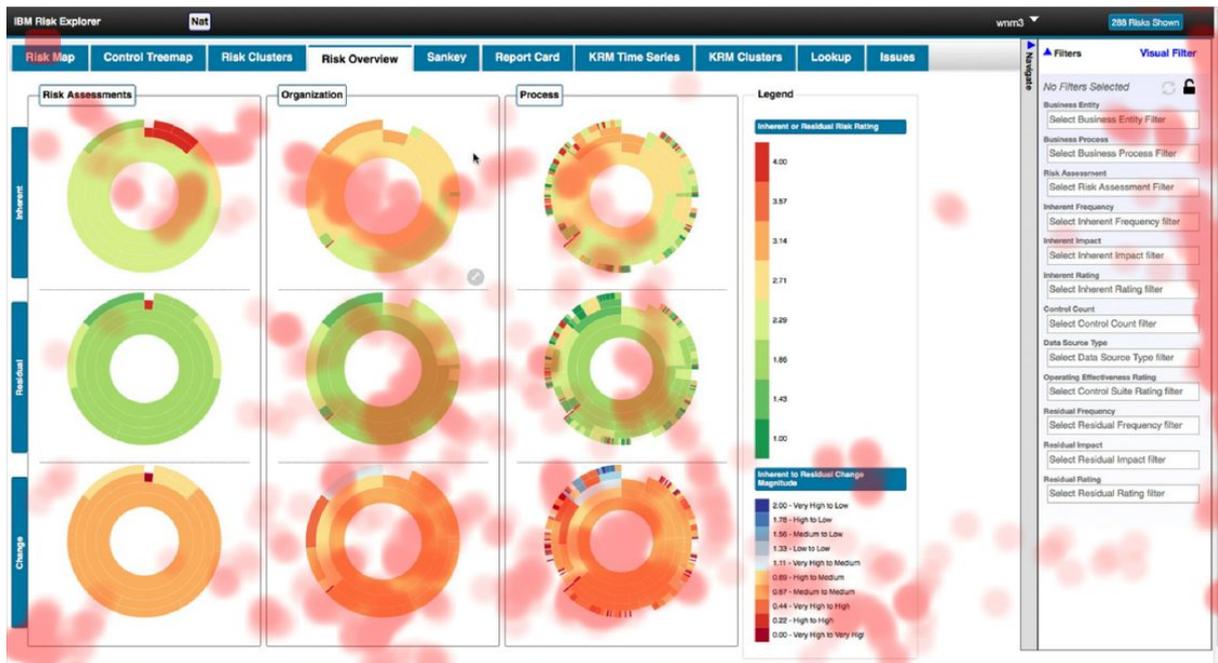


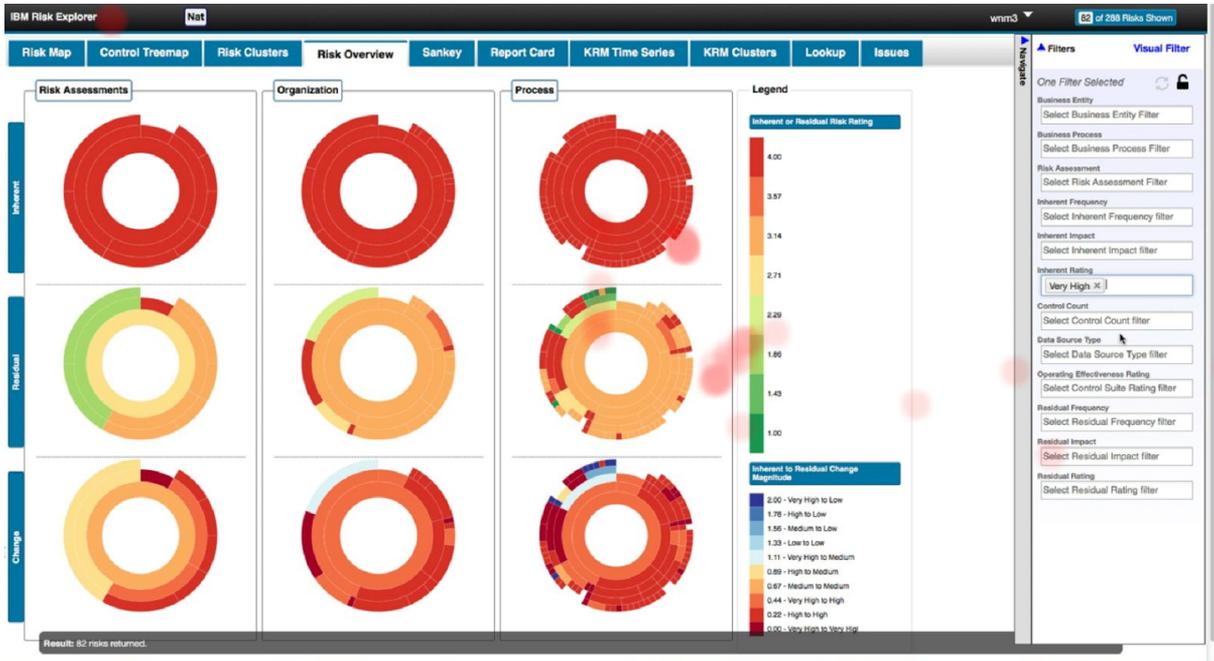


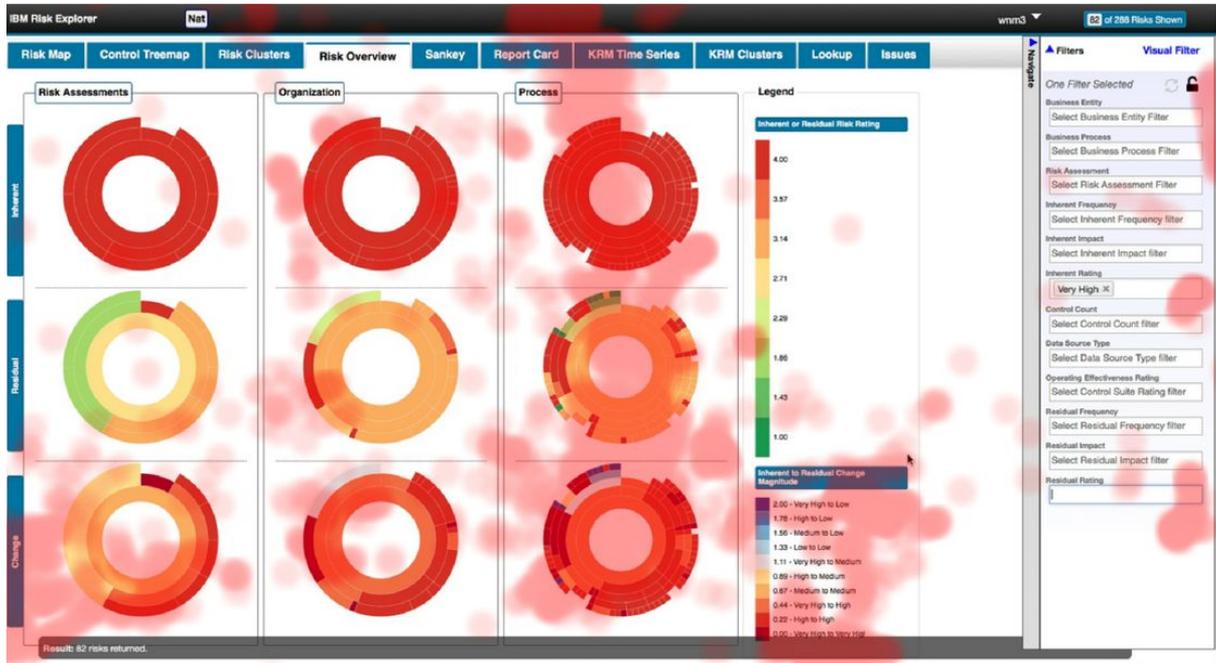


Buscou pela resposta nas abas Control Treemap, Sankey e Risk Clusters. Se concentrou na aba Risk Clusters. Utilizou um filtro e depois deixou de utilizar como vemos na sequência de imagens anteriores.

Tarefa 7

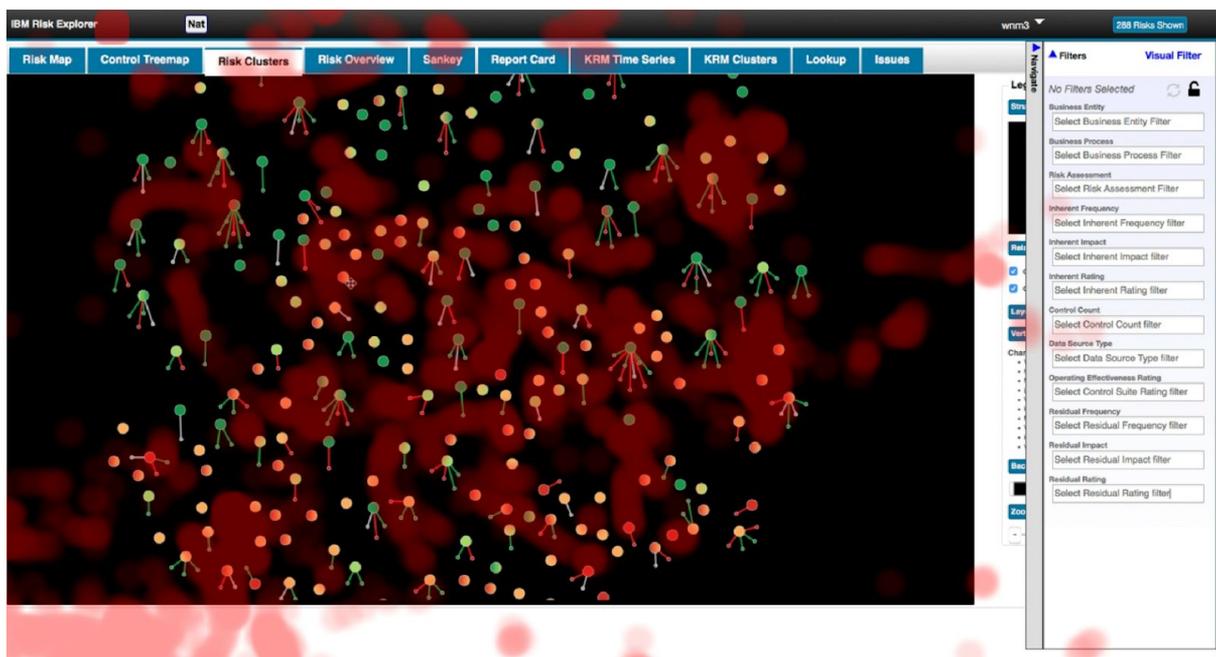






Não conseguiu encontrar a resposta, passeou pelas abas e não conseguiu encontrar uma visualização que ajudasse.

Tarefa 8





Demorou um pouco para concluir a tarefa mas conseguiu concluir. Se concentrou em procurar na aba do Risk Clusters. Quando o participante encontra a resposta para a tarefa vemos a quantidade de pontos concentrado na região da resposta.

Preferências do usuário

O usuário gostou mais das abas Risk Clusters e risk Overview.

Participante 3

Tarefa 1



Conseguiu concluir com rapidez a tarefa. Pelas imagens parece que observou a tela toda mas isso não influenciou na resposta.

Tarefa 2



Conseguiu concluir com rapidez a tarefa. Pelas imagens parece que observou a tela toda mas isso não influenciou na resposta.

Tarefa 3



Conseguiu concluir com rapidez a tarefa. Pelas imagens parece que observou a tela toda mas isso não influenciou na resposta.

Tarefa 4



Conseguiu concluir com rapidez a tarefa. Pelas imagens parece que observou a tela toda mas isso não influenciou na resposta.

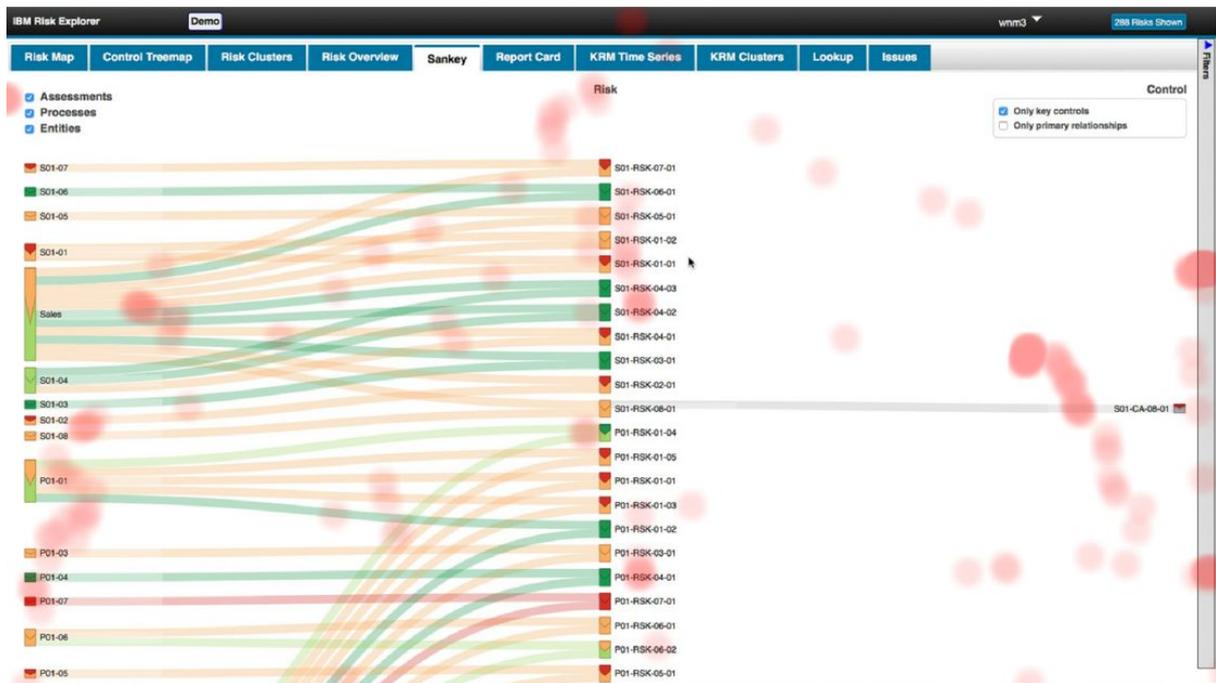
Tarefa 5



Respondeu através da aba risk Overview. Respondeu mas não corretamente, vemos que os pontos estão distribuídos pela tela, indicando a busca sem sucesso pela resposta.

Tarefa 6







Buscou em três abas, risk Overview, Risk Map e Sankey. Respondeu através da aba Sankey, nas imagens vimos que confundiu a coluna de riscos pela de controles e a resposta foi um risco quando deveria ser um controle.

Tarefa 7



Buscou na aba Risk Overview, pontos espalhados pela imagem indicando confusão. Não respondeu exatamente o que foi pedido.

Tarefa 8



Buscou na aba risk Overview, se concentrou na área de change (linha inferior de sunbursts). Não respondeu exatamente o que foi pedido, respondeu o risco que teve mais mudança.

Preferências do usuário

O usuário gostou mais das abas risk Overview e Sankey e gostou menos da aba Risk Clusters.

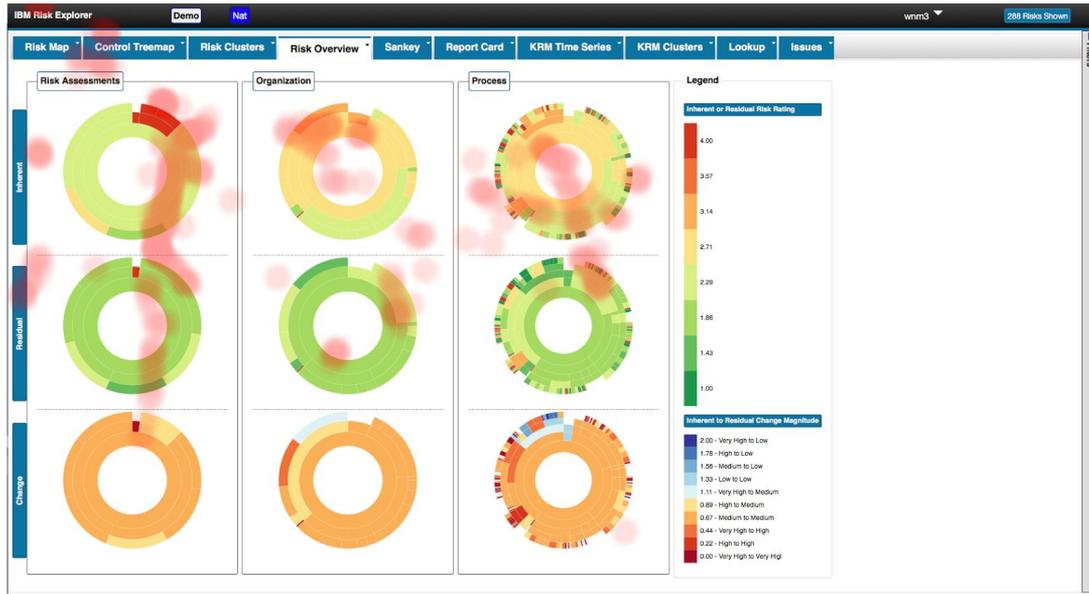
Participante 4

Tarefa 1



Nessa tarefa o usuário focou sua atenção na área em que estava a resposta (primeira linha de sunbursts) como vemos na figura.

Tarefa 2



O usuário focou sua atenção na área em que estava a resposta como vemos na figura.

Tarefa 3



O usuário focou sua atenção na área em que estava a resposta (segunda linha de sunbursts) e na área da resposta da tarefa anterior como vemos na figura, mas conseguiu encontrar a resposta.

Tarefa 4

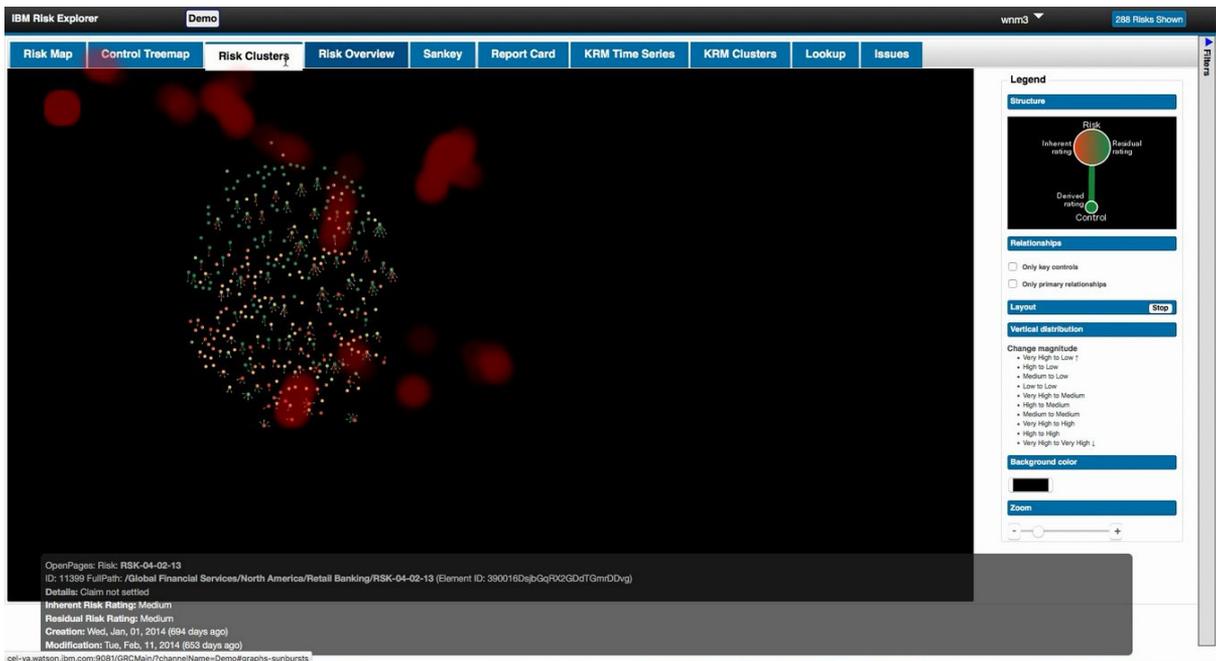


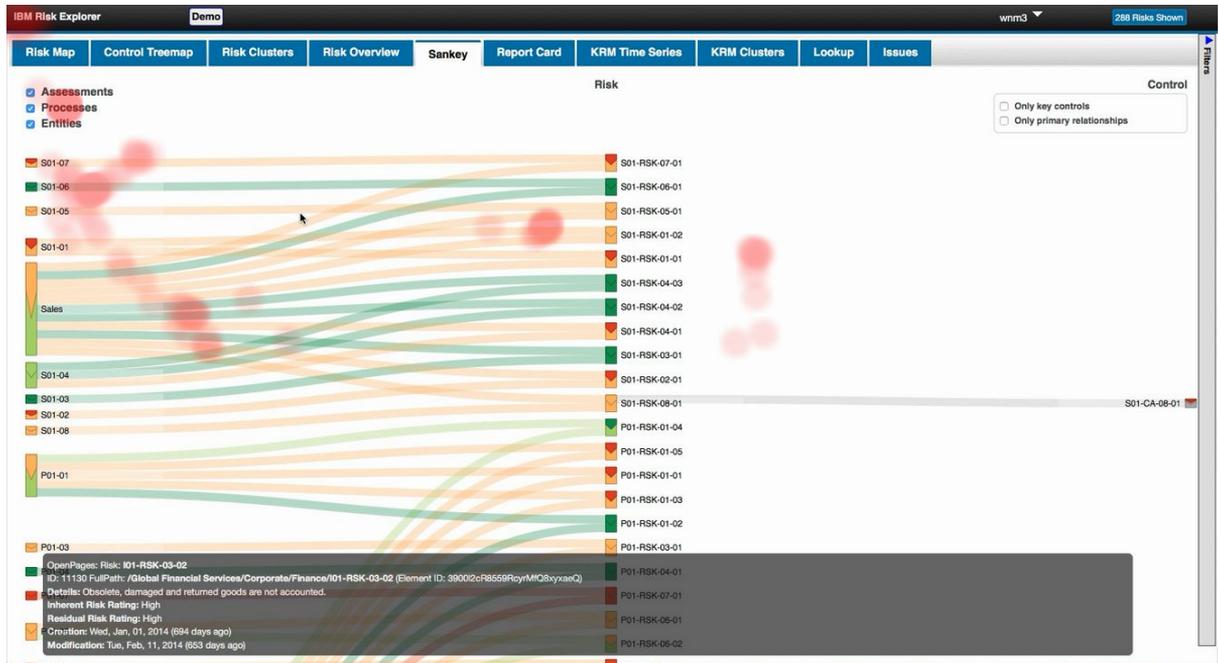
O usuário focou sua atenção na área em que estava a resposta como vemos na figura.

Tarefa 5



Tarefa 6





O usuário não conseguiu concluir a tarefa, ele procurou em várias abas mas não encontrou a resposta.

Tarefa 7

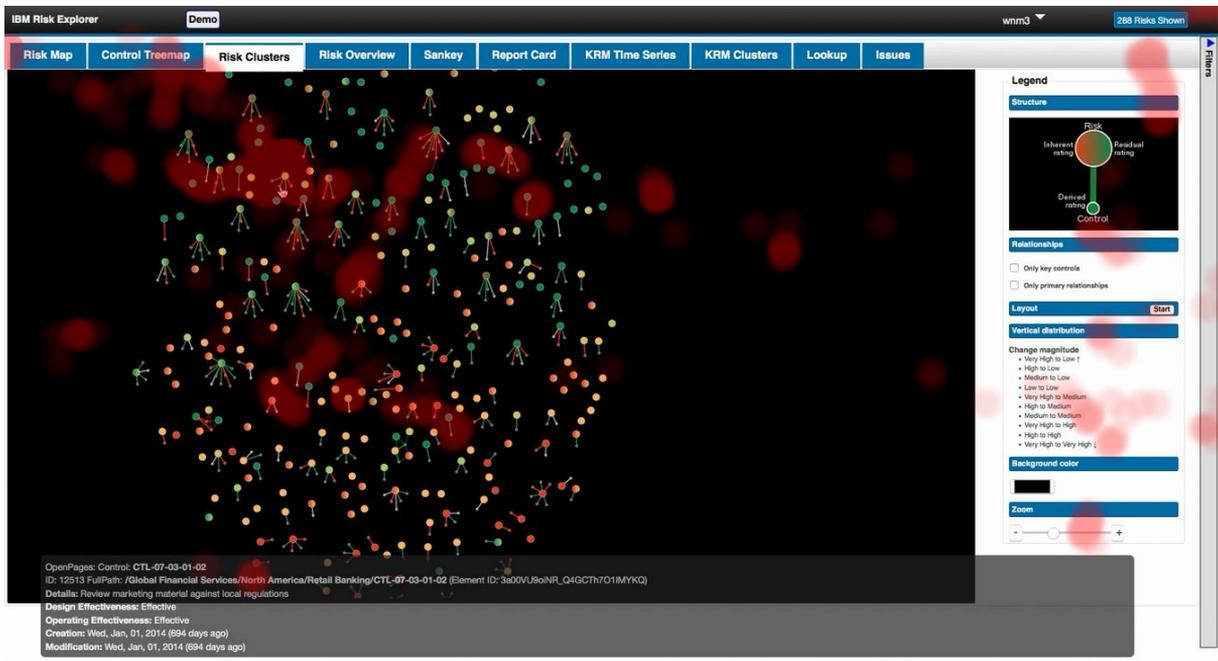
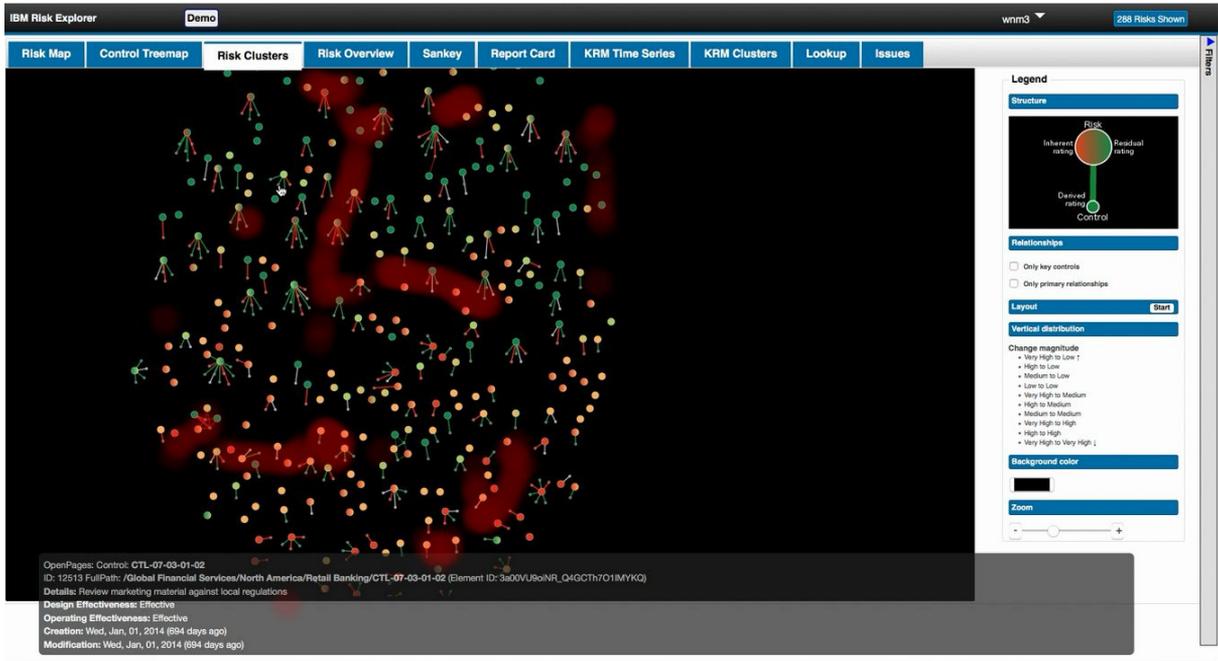


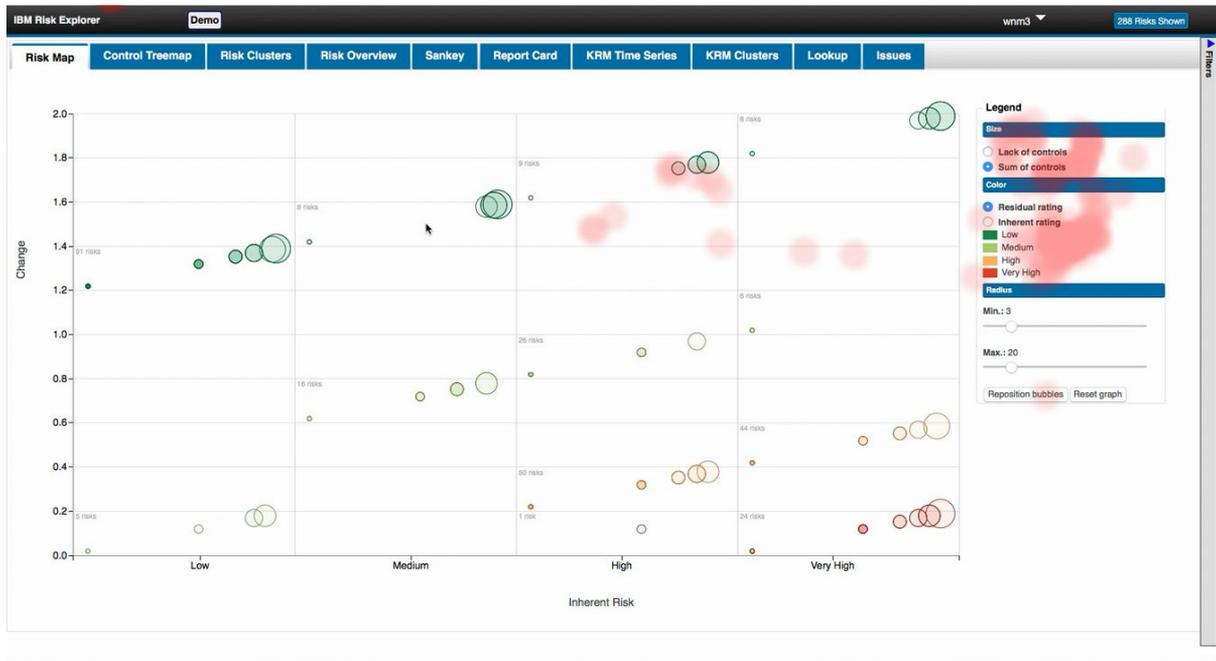


Procurou pelas telas e usou a aba Risk Map para responder. Não conseguiu responder corretamente.

Tarefa 8







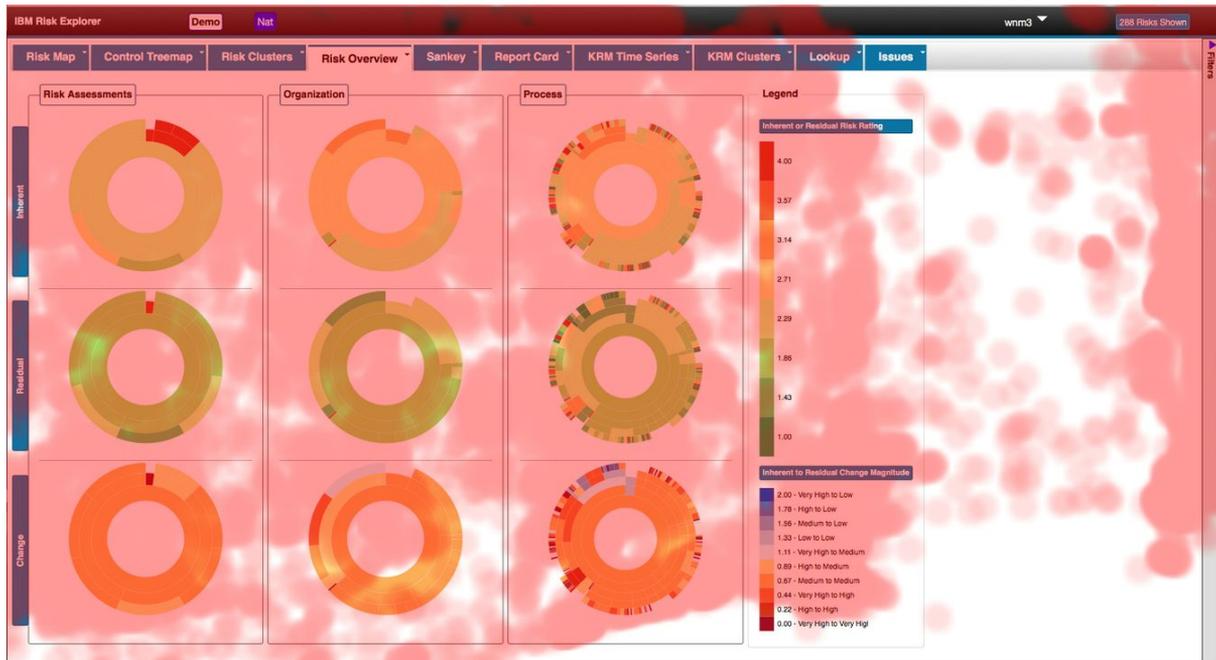
Procurou pela resposta em várias telas. Utilizou a tela Risk Map para encontrar a resposta, mas respondeu incorretamente.

Preferências do usuário

O usuário gostou mais das abas Risk Map e Risk Clusters e não gostou da aba Control Treemap, pois achou difícil entendê-la.

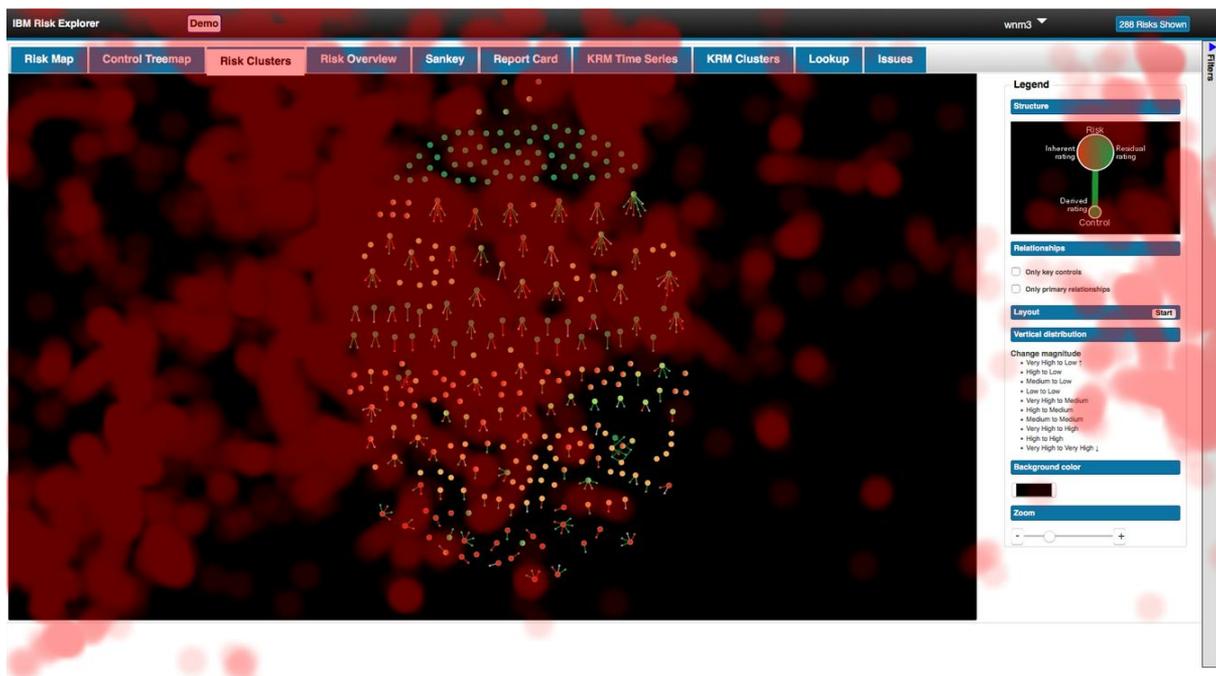
Junção dos dados de todos os participantes por tela

Risk Overview



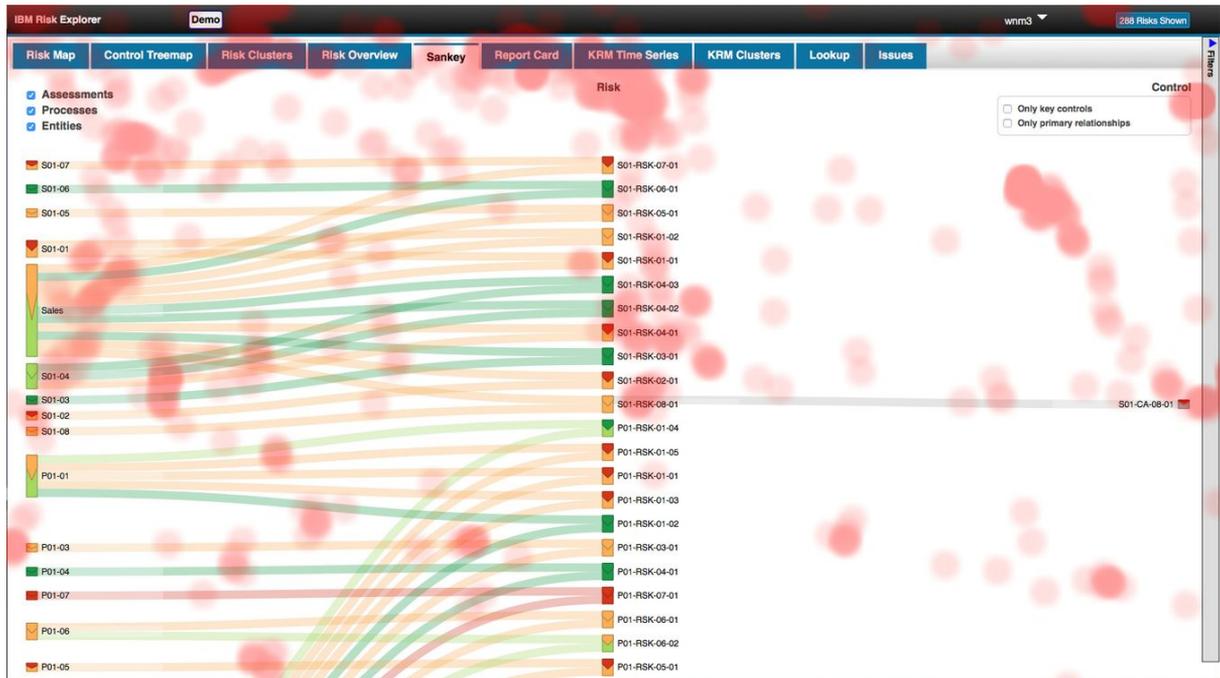
Observamos a grande utilização dessa tela, conforme a grade concentração de pontos nela.

Risk Clusters



Observamos a concentração dos pontos na região central da tela, onde se encontram a maior parte dos riscos e controles.

Sankey



Observamos a grande concentração nos rótulos e na coluna de riscos.

Control Treemap



Essa tela foi pouco utilizada como vemos através da imagem.

Risk Map



Na primeira figura do Risk Map o heat map representa a situação em que a opção da legenda seleccionada é o "lack of controls" e na segunda figura a opção seleccionada é o "sum of controls". Vemos uma grande concentração de pontos na legenda dessa aba.

6. Conclusões e trabalhos futuros

As primeiras tarefas (de 1 a 4) foram rapidamente executadas, podemos ver através das respostas dos participantes e através dos *heat maps*. Todos os quatro participantes utilizaram a aba risk Overview para realizar essas tarefas. A escala de cores na área *change*, na parte inferior da visualização, confundiu os participantes. O motivo é que o mesmo tom de laranja é usado em diferentes pontos das duas escalas usadas na visualização. No geral, a aba risk Overview foi bem interpretada pelos usuários e uma das mais utilizadas durante a avaliação.

A aba Sankey também foi bastante utilizada, os participantes se confundiram entre a coluna de riscos (ao centro) e a coluna representando os nós folha das taxonomia (à esquerda). Realçar as diferenças entre essas estruturas potencialmente resolve esta questão.

A aba Risk Map foi pouco utilizada. Os participantes demoraram para entender a finalidade dessa visualização. Ela combina a escala de mudança com o *rating* inerente. Uma possível melhora seria trocar a escala de mudança pelo *rating* residual, tornando mais natural para os especialistas o posicionamento nos eixos X e Y.

A aba Control Treemap foi vista como confusa e foi indicada como a menos útil para alguns participantes.

A aba Risk Clusters foi no geral bem recebida, os participantes entenderam seu propósito e conseguiram identificar agrupamentos de riscos, controles e seus relacionamentos. O participante que não gostou indicou que havia muita informação e a animação que o gráfico possui gerada pelo algoritmo de posicionamento dos nós baseado em força também pode ter influenciado este ponto.

As tarefas 7 e 8 eram as mais complexas e foram executadas com maior dificuldade. Na tarefa 7 quatro participantes responderam à pergunta, mas apenas um respondeu corretamente. Na tarefa 8 os quatro responderam, mas apenas dois responderam corretamente. Através das imagens foi possível ver os pontos bem espalhados pela tela indicando que os participantes tentaram buscar informações explorando todas as telas que foram utilizadas para tentar concluir a tarefa.

As tarefas 5 e 6 foram executadas com dificuldade também. Na tarefa 5 três conseguiram responder, mas apenas um respondeu corretamente. Na tarefa 6 dois conseguiram responder, mas apenas um corretamente. Muitos acharam que a resposta seria facilmente encontrada na aba risk Overview, mas não é possível localizar a resposta nessa aba, mas quando foram para aba Sankey ou Risk Clusters conseguiram localizar.

Esse trabalho mostrou que para algumas tarefas o sistema é útil e tem uma boa usabilidade, já para outras a inclusão de legendas que expliquem melhor o uso de cada visualização pode melhorar a usabilidade.

Os testes de usabilidade permitem saber se o sistema que será usado é bem recebido pelos usuários, assim seus desenvolvedores podem fazer melhorias e correções de acordo com o retorno obtido. O uso de *eye tracking* nesses testes é uma técnica interessante que

acrescenta mais uma camada de dados para estudo, especialmente quando as visualizações representam um ponto central no sistema, como é o caso deste estudo.

Como trabalhos futuros planejamos envolver mais especialistas em futuras sessões de testes com usuários, analisar em mais detalhes as visualizações preferidas pelos participantes e avaliar soluções identificadas neste estudo.

7. Referências

- Barr, D. J. 2008. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of memory and language*, 59(4), 457-474.
- BWISE GRC. <http://www.bwise.com/>
- Compliance 360. <http://www.compliance360.com/>
- Cybis, W., Holtz, A., & Faust, R. (2010). *Ergonomia e usabilidade*. São Paulo: Novatec.
- Goldberg, J. H., Wichansky, A. M., 2003. *Eye Tracking in Usability Evaluation: A Practitioner's Guide*. The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements, Elsevier Science, Oxford, 2002.
- Granka, L. A., Joachims, T., & Gay, G. , 2004. Eye-tracking analysis of user behavior in WWW search. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 478-479). ACM.
- IBM Open Pages. <http://www-03.ibm.com/software/products/pt/openpages-grc-platform>
- International Organization for Standardization, 1998. ISO 9241 Part 11: Guidance on usability.
- Jacob, R. J., & Karn, K. S. , 2003. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind*, 2(3), 4.
- Li, D., Babcock, J., & Parkhurst, D. J. ,2006. openEyes: a low-cost head-mounted eye-tracking solution. In *Proceedings of the 2006 symposium on Eye tracking research & applications* (pp. 95-100). ACM.
- Matsunaga, C. *Eyetracking: interação baseada nos olhos*. III COMA - Coletivo da pós-graduação em arte.
- MetricStream GRC. <http://www.metricstream.com/>
- Nielsen, J., 1993. *Usability Engineering*.
- Pernice, K. e Nielsen, J., 2009. *How to Conduct Eyetracking Studies*.
- Poole, A., Ball, L. J. *Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects*
- Salvucci, D. D., & Goldberg, J. H. , 2000. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications* (pp. 71-78). ACM.
- Shneiderman, 1996. *Visual Design Seeking Mantra*.
- Tarantino, A., 2008. *Governance, Risk, and Compliance Handbook: Technology, Finance, Environmental, and International Guidance and Best Practices*.
- U.S. Department of Health & Human Services. *Usability.gov: Improving the User Experience*. <http://www.usability.gov/>

8. Parte Subjetiva

O principal desafio foi envolver no teste pessoas especializadas em plataformas GRC e que tenham experiência em interagir com dados GRC. Dentro da empresa em que foram realizados os testes recebemos uma lista com 18 profissionais especializados na área. Contatamos todos, mas apenas 6 pessoas aceitaram, 5 agendaram um horário para a realização, mas no final, apenas 4 compareceram. Fizemos um teste piloto com uma pessoa que não era da área para validação do plano de teste e percebemos a importância de envolver o público-alvo do sistema para realizar/lidar com o sistema e participar do teste.

Com esse trabalho aprendi que quando fazemos teste de sistemas computacionais envolvendo pessoas devemos estimar que parte deles não aceitará o convite ou não irá comparecer ao teste. Esse trabalho me fez aprender muito sobre a área da IHC e como filtrar, processar e interpretar todos os dados que capturamos durante a interação.