

Relatório da Avaliação do Uso do Sistema de Repositório Institucional

Universidade Federal de Goiás

Faculdade de Informação e Comunicação
Laboratório de Políticas Públicas Participativas

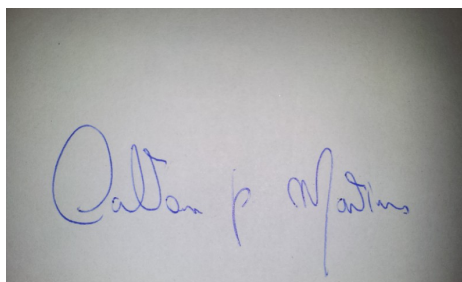
Ministério da Cultura

Coordenação Geral de Cultura Digital

**Projeto: “Laboratório de Políticas Públicas Participativas: foco em
acervos digitais”**

Junho/2016

Responsável pelo relatório



Prof. Dr. Dalton Lopes Martins

Coordenador do Laboratório de Políticas Públicas Participativas

Sumário

1Avaliação.....	3
2Google Analytics.....	5
3Teste de performance.....	7
4Referências.....	12

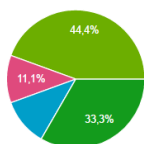
1 Avaliação

Ao longo deste primeiro ano de desenvolvimento do sistema Tainacan foram promovidas três oficinas que fazem parte de ações da política nacional para acervos digitais da cultura brasileira. Tal política é resultado da parceria entre a Secretaria de Políticas Culturais (SPC) do Ministério da Cultura e as universidades federais de Pernambuco (UFPE) e Goiás (UFG) tendo como objetivo incentivar a preservação e difusão da diversidade cultural brasileira na internet.

As oficinas contaram com um público selecionado por meio de editais específicos relacionados a preservação e acesso a acervos do patrimônio Afrodigital (Acervo da Cultura Afro Brasileira - MINC / UFPE / UFG, 2010). As oficinas ocorreram no período entre agosto e novembro de 2015, sendo realizadas respectivamente nas cidades de Goiânia, Porto Alegre e Recife. Dentre os resultados destas oficinas pode-se destacar a construção de um acervo digital na forma de repositório utilizando o sistema Tainacan batizado de Afro digital (Acervo da Cultura Afro Brasileira - MINC / UFPE / UFG, 2010).

O primeiro resultado a ser apresentado e discutido resultante de uma dessas oficinas é um questionário aplicado com o público participante da oficina realizada em Porto Alegre em meados do mês de outubro. Neste questionário foram aplicadas questões voltadas desde ao perfil do público (veja Figura 1).

Qual seu grau de instrução?



2º grau incompleto	0	0%
2º grau completo	0	0%
3º grau incompleto	0	0%
3º grau completo	3	33.3%
Mestrado incompleto	0	0%
Mestrado completo	1	11.1%
Doutorado incompleto	1	11.1%
Doutorado completo	4	44.4%

Qual é sua profissão?

Professora
Professor
Agente cultural
Professora
ARQUIVISTA
Antropólogo

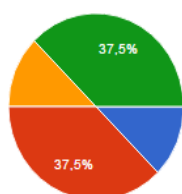
Figura 1. Perfil dos participantes da oficina ministrada em Porto Alegre em meados de outubro de 2015.

O primeiro aspecto é o perfil do público que participou que foi composto mais por pessoas que já possuíam uma formação em nível superior com maior número de pessoas com pós-graduação especialmente no nível doutorado. Isso ressalta o nível

acadêmico dos profissionais que participaram da oficina, sem falar na área de atuação que se concentra principalmente em áreas ligadas a acervos culturais e na educação.

O próximo aspecto deste questionário a ser destacado foi o questionamento se os participantes já haviam tido algum contato com outros sistemas de repositório antes de participarem da oficina com o Tainacan (veja a Figura 2). As respostas mostraram um equilíbrio entre pessoas que já tiveram algum contato nos últimos seis meses e outras que ainda não tiveram nenhum contato como mostra a figura abaixo extraída do formulário.

Você utilizou algum software para repositório digitais (dspace, fedora, etc) para administração de uma coleção?

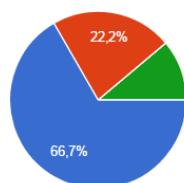


Sim, uso frequentemente (ao menos uma vez por semana)	1	12.5%
Sim, pouco uso (ao menos uma vez nos últimos 6 meses)	3	37.5%
Sim, a muito tempo atrás (mais de um ano)	1	12.5%
Não	3	37.5%

Figura 2. Utilização de algum software para repositório digital em meados de outubro de 2015.

O que se pode analisar por meio desses resultados é a possibilidade de interação com o Tainacan de dois perfis de usuários em condições diferentes em relação a experiência de uso com sistemas de repositório. De um lado um perfil marcado por pessoas que já tiveram alguma experiência recentemente (). Esse tipo de perfil traz consigo relatos interessantes do ponto de usabilidade do sistema ao ponto que já possuem uma carga de experiência e podem levantar possíveis dificuldades, melhoramentos e pontos positivos na execução de determinada atividade no Tainacan em relação a outros sistemas.

Você utiliza repositórios digitais para consulta e publicação de documentos digitais?



Sim, uso frequentemente (ao menos uma vez por semana)	6	66.7%
Sim, pouco uso (ao menos uma vez nos últimos 6 meses)	2	22.2%
Sim, a muito tempo atrás (mais de um ano)	0	0%
Não	1	11.1%

Figura 3. Gráfico com a frequência de utilização de sistema para repositórios digitais para consulta e publicação de documentos digitais.

De outro lado um perfil de participantes que ainda não tiveram nenhum contato recente, traz consigo um relato mais natural e desvinculado a outras experiências de uso com outros sistemas, o que possibilita avaliar a capacidade do sistema em promover um acolhimento e condução de um usuário durante a execução de determinada tarefa. Isso quer dizer, que o relato final da experiência do usuário será exclusivo ao uso do Tainacan, o que possibilita o levantamento de dificuldades, melhoramentos e pontos positivos no sistema como um todo que ao contrário do relato de um perfil de usuário

mais experiente pode contar com influências das suas experiências recentes com outros sistemas para repositório.

Outro resultado interessante deste questionário é o questionamento sobre a aplicabilidade da formação realizada na oficina no dia a dia dos participantes. A Figura 4 demonstra que os resultados se concentraram entre “boa aplicabilidade” e “excelente aplicabilidade”.

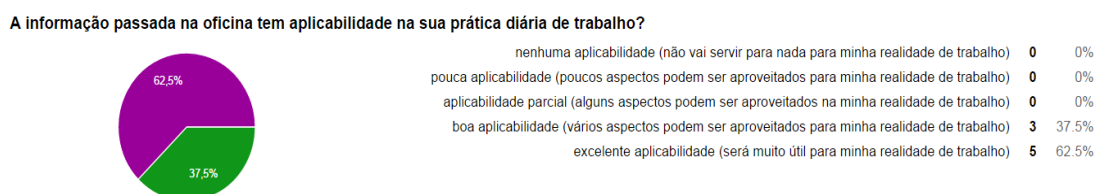


Figura 4. Aplicabilidade da oficina na sua prática diária de trabalho.

Essa concentração denotada no gráfico da figura pode ser entendida como a validação de todo o processo de desenvolvimento do Tainacan, desde o levantamento dos requisitos e funcionalidades a serem implementadas, aos aspectos de interface e usabilidade, ou seja, o sistema como um todo. Inicialmente isso se justifica pela característica do público participante da oficina se tratar de usuários que possuem acervos reais e passaram por um processo de seleção por meio de edital promovido Ministério da Cultura (Minc) para participação na oficina. Assim o aspecto positivo da avaliação dos usuários representada no gráfico da figura acima se torna bastante válido e os seus relatos e comentários ao longo da realização da oficina se tornaram uma fonte interessante para identificação de melhoramentos nos mais diversos aspectos do sistema, e também para a validação do processo de teste aplicado, e levantamento de novas funcionalidades para implementação.

A próxima seção apresentará uma descrição do monitoramento da instância do Tainacan (Afrodigital, 2016) utilizada nas oficinas realizadas ao longo deste primeiro ciclo de desenvolvimento.

2 Google Analytics

Nesta seção serão apresentadas algumas estatísticas dentro do período de outubro de 2015 a maio de 2016 oriundas do Google Analytics (Google Analytics, 2016) acerca da instância do Tainacan (Afrodigital, 2016) utilizada durante as oficinas e que recebeu os mais diversos acervos culturais. Assim dentro dessas estatísticas estarão informações sobre a quantidade de visualizações de páginas, usuários ativos, regiões com maior número de acessos, entre outras métricas.

A primeira estatística a ser apresentada é o número de visualizações de páginas (veja a Figura 5). Neste período foram realizadas cerca de 12.312 visualizações de página, com aproximadamente 4.375 sessões estabelecidas, com duração média de 4 minutos e meio. O pico de visualizações foi registrado no dia 25 de novembro de 2015 com

aproximadamente 982 visualizações, o que coincide com o período de realização da oficina em Recife-PE. Outro dado interessante é o percentual de novas sessões (61%) que reflete o número de interações pela primeira com o site. Além disso, a métrica página/sessão (2,81) representa o número médio de páginas visitadas no site.

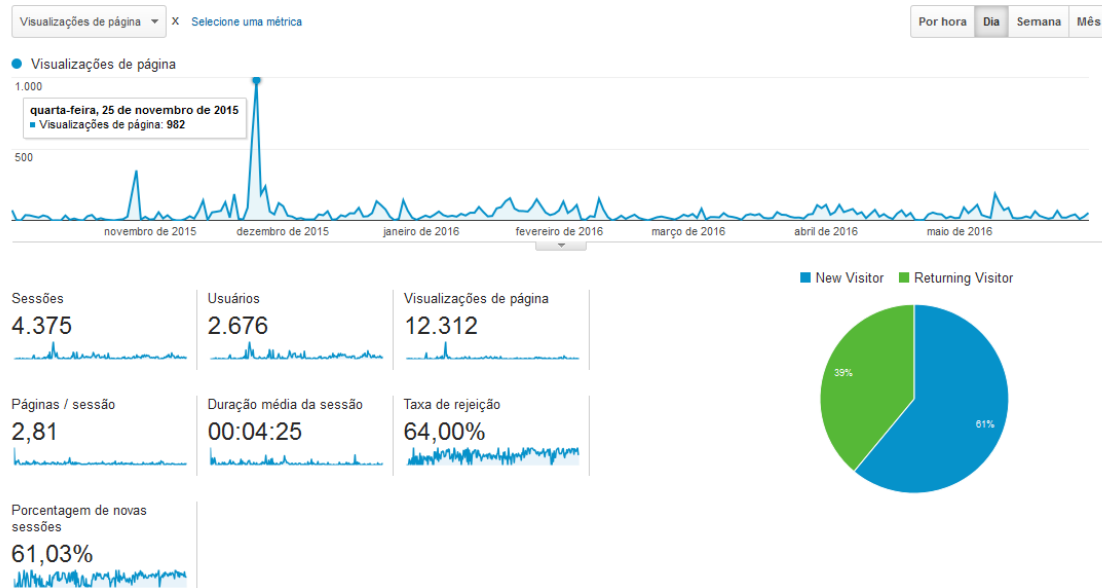


Figura 5. Número de visualizações de página na instância Afrodigital no período de outubro de 2015 a maio de 2016.

Dando sequência a apresentação das estatísticas chegamos a quantidade de usuários ativos no período avaliado (veja a Figura 6). Neste período foram contabilizados uma média de 17 usuários por dia, 88 usuários por sete dias, 194 usuários por 14 dias e 543 usuários por 30 dias.

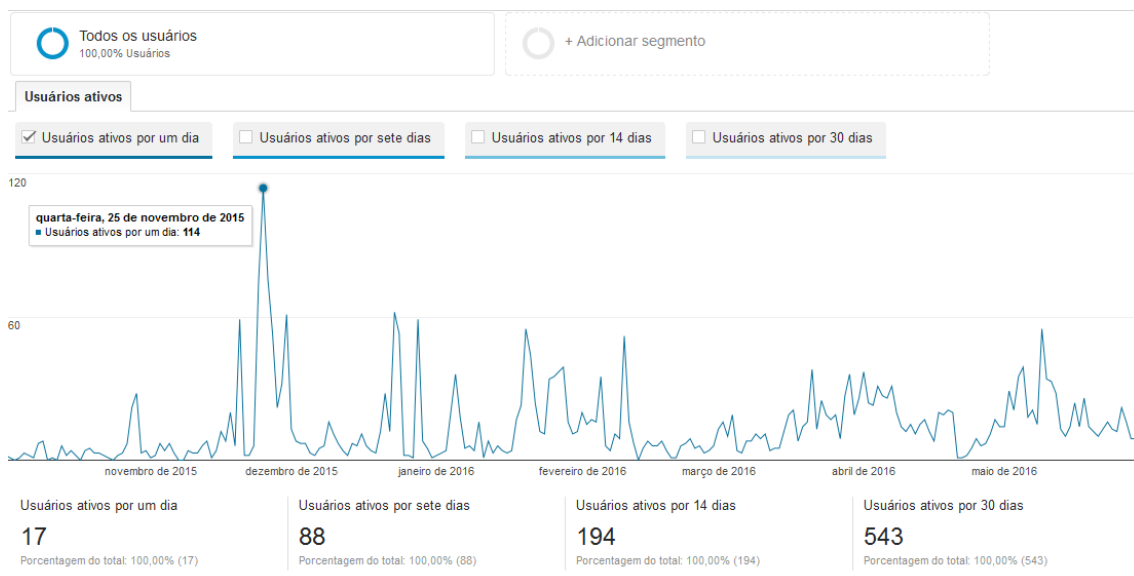


Figura 6. Número de usuários ativos ao site Afrodigital no período de outubro de 2015 a maio de 2016.

Em seguida, a próxima métrica a ser apresentada é a localização da concentração de sessões estabelecidas ao site Afrodigital (veja Figura 7). Nesta métrica fica visível a maior concentração de sessões ao site está aqui no Brasil (58,72%), dado justificável pela visibilidade e divulgação do Tainacan estar mais concentrada no âmbito regional e nacional, apesar que iniciativas a nível internacional já estejam em curso.

Um possível reflexo dessas iniciativas a nível internacional pode ser identificado nos dados provenientes do país Portugal, onde apesar de um baixo número de sessões, apresenta uma taxa de rejeição bem abaixo dos outros países e próxima a taxa brasileira, o que pode ser interpretado que a maioria dessas sessões foi estabelecida com o objetivo de acessar o site e seu conteúdo e não por algum engano ou outro motivo.

Pais ?	Sessões ? ↓	Porcentagem de novas sessões ?	Novos usuários ?	Taxa de rejeição ?	Páginas / sessão ?	Duração média da sessão ?
	4.375 Porcentagem do total: 100,00% (4.375)	61,03% Média de visualizações: 61,03% (0,00%)	2.670 Porcentagem do total: 100,00% (2.670)	64,00% Média de visualizações: 64,00% (0,00%)	2,81 Média de visualizações: 2,81 (0,00%)	00:04:25 Média de visualizações: 00:04:25 (0,00%)
1. Brazil	2.569 (58,72%)	51,03%	1.311 (49,10%)	44,80%	3,96	00:06:25
2. Russia	929 (21,23%)	60,06%	558 (20,90%)	89,88%	1,19	00:02:33
3. United States	239 (5,46%)	96,23%	230 (8,61%)	96,65%	1,07	00:00:22
4. (not set)	211 (4,82%)	97,16%	205 (7,68%)	98,58%	1,01	00:00:05
5. United Kingdom	50 (1,14%)	88,00%	44 (1,65%)	94,00%	1,10	00:00:05
6. China	46 (1,05%)	91,30%	42 (1,57%)	91,30%	1,09	00:00:53
7. Kyrgyzstan	36 (0,82%)	86,11%	31 (1,16%)	100,00%	1,00	00:00:00
8. Portugal	35 (0,80%)	77,14%	27 (1,01%)	42,86%	3,00	00:03:44
9. Netherlands	30 (0,69%)	90,00%	27 (1,01%)	80,00%	1,20	00:00:29
10. Germany	17 (0,39%)	88,24%	15 (0,56%)	100,00%	1,00	00:00:00

Figura 7. Número de sessões por país estabelecidas ao site Afrodigital no período de outubro de 2015 a maio de 2016.

Na próxima seção serão apresentados os resultados de testes de performance aplicados ao Tainacan em ambientes que simulavam repositórios digitais.

3 Teste de performance

Nesta seção serão apresentados os resultados de alguns testes realizados utilizando o Tainacan. Nestes testes se buscou verificar o seu desempenho no carregamento de coleções que já possuíam determinadas quantidades de itens, como 100 itens, 1000 itens (1k), 10000 itens (10k) e 100000 itens (100 k), onde foi testado seu tempo de carregamento em três instâncias do Tainacan com perfis diferentes de repositório que simulavam quantidades de itens diferentes. Todas as coleções desses três perfis de

repositórios possuíam 2 níveis de categorias com dez subcategorias e cada item com duas classificações.

Antes de descrever o perfil dos repositórios utilizados nos testes, as configurações da máquina que está rodando o Tainacan são:

- ! Modelo: Intel(R) Xeon(R) CPU E5-2609 0 @2.40GHz, 16GB RAM, 5.4TB HD;
- ! Sistema Operacional: Debian GNU/Linux 8;
- ! Versão do Apache Web Server: 2.4.10;
- ! Versão do PHP: 5.6.17;
- ! Versão do Wordpress: 4.5.2;
- ! Versão do banco de dados Mysql: 5.5.47;

Em seguida as três instâncias do Tainacan utilizadas foram definidas em “Vazio”, 11 milhões de itens e 62 milhões de itens. No perfil “Vazio” a fase de preenchimento do repositório foi feita gradualmente, ou seja, para cada tipo de coleção testada, por exemplo, coleção de 100 itens, foi criada e populada no repositório e posteriormente foi realizado o teste de carregamento da coleção. Assim, após a finalização dos testes, as demais coleções foram adicionadas ao repositório e testadas, ou seja, o tamanho do repositório sofria um crescimento gradual.

Os perfis 11 milhões e 62 milhões foram preenchidos antes da realização dos testes de carregamento, utilizando um script específico desenvolvido para esta finalidade, que cria diversas coleções com quantidades iguais. Assim nestes perfis foram criadas e populadas todos os tipos de coleções testadas, em seguida os repositórios foram populados até a quantidade de 11 milhões e 62 milhões.

O script de população das coleções desenvolvido para o Tainacan foi desenvolvido em duas versões, sendo a versão inicial já incorporada ao Tainacan e outra desenvolvida na forma de uma aplicação web independente que conecta diretamente a base de dados utilizada pela Tainacan.

A versão incorporada ao Tainacan permite ao administrador do repositório popular determinada coleção definindo a quantidade de categorias por nível, número de níveis, itens por categoria e quantidade de classificações por item (veja Figura 8).

Popular coleção

Coleção

Categorias por nível

Número de níveis:

Total de Categorias :

Itens por categoria

Total de itens :

Classificação

Total de Classificações :

Figura 8. Página da funcionalidade que popula coleções no Tainacan.

A segunda versão desenvolvida (veja a Figura 9) é mais avançada e permite ao usuário popular uma coleção diretamente passando o seu endereço URL, ou definir um número de coleções que devem ser criadas no repositório seguindo as mesmas configurações presentes na primeira versão, no que se refere a número de subcategorias, número de níveis e classificações. A diferença fundamental desta versão é a possibilidade de inserir coleções no repositório, o que não era possível na primeira.

Popular Tainacan

Inserindo valores no repositório do Tainacan!

OU

Subcategories per level

Number of levels

Total of Categories:

Items per category

Total of Items:

Classification

Total of Classifications:

POPULAR!

Figura 9. Página do script de população desenvolvido para a realização dos testes no Tainacan.

Um aspecto importante a ser ressaltado no desenvolvimento das versões deste script de população foi a possibilidade de testar o Tainacan em duas vias. A primeira via mais diretamente ligada à inserção de grandes quantidades de itens e coleções em uma instância, verificando o comportamento do sistema em cada cenário de teste. A segunda via mais indiretamente, o script viabilizou a criação de coleções e repositórios com grandes quantidades de itens de forma automática e relativamente rápida o que permitiu a execução dos testes nos cenários propostos.

A metodologia utilizada nos testes para a coleta dos dados e estimativa do tempo de carregamento foi a utilização do plugin Firebug (Firebug, 2016) no navegador Firefox. O Firebug por disponibilizar recursos para monitoramento CSS, HTML e JavaScript em qualquer página da Web. Dentro esses recursos há uma aba intitulada “Rede” que traz a informação de quanto tempo se levou para o carregamento da página (veja Figura 10).

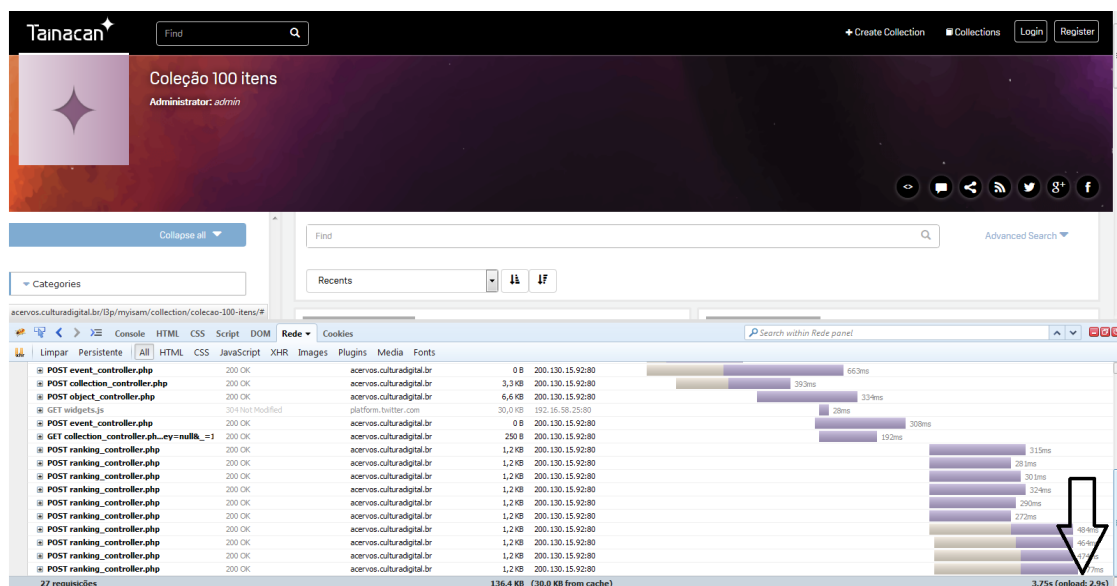
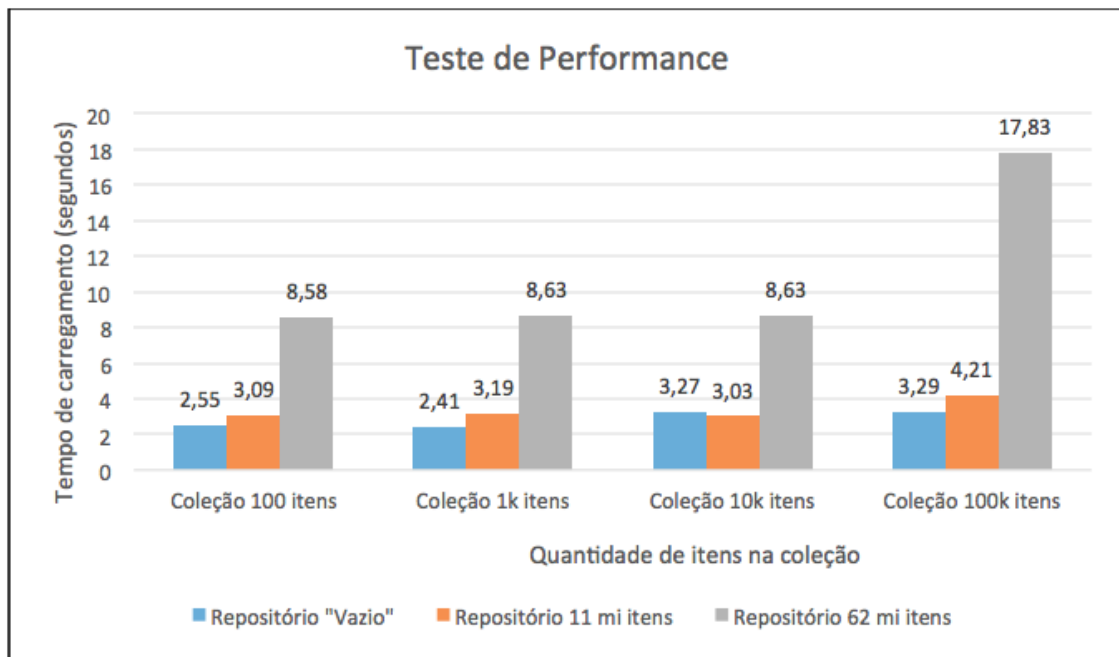


Figura 10. Aba "Rede" do plugin Firebug que permite determinar o tempo de carregamento de uma página.

Dessa forma para o levantamento do tempo de carregamento de cada coleção foram realizadas 4 leituras, onde a primeira foi descartada e as outras três restantes foram utilizadas para a cálculo de um tempo médio de carregamento. Assim os dados apresentados no Gráfico 1 trazem o tempo médio de carregamento de cada coleção em cada perfil de repositório utilizado na execução dos testes.

Gráfico 1. Gráfico comparativo com os tempos de carregamento de coleções de acordo com o tamanho do repositório.



Os resultados (veja o *Gráfico 1*) apontam que o Tainacan suporta um repositório com uma quantidade de itens aproximada de 10 milhões de itens com coleções com 2 níveis de hierarquia com 10 subcategorias e que possuam aproximadamente 100 mil itens. Outro aspecto detectado durante a execução dos testes foi uma sensível lentidão no carregamento da coleção 100k no repositório com 62 milhões de itens, com um aumento de 8,63 s para 17,83 s.

Dessa forma, os próximos desdobramentos acerca do desenvolvimento objetivam aumentar a capacidade do Tainacan em suportar coleções maiores com um repositório com uma quantidade maior de itens, sendo num patamar de coleções com 500 mil itens e repositórios com uma quantidade de 100 milhões de itens.

4 Referências

1. *Afrodigital*. (31 de Maio de 2016). Fonte: Acervo da Cultura Afro Brasileira: <http://afro.culturadigital.br/>
2. *Firebug*. (31 de Maio de 2016). Fonte: Firebug Web Development Evolved: <http://getfirebug.com/>
3. *Google Analytics*. (31 de Maio de 2016). Fonte: Google Analytics: <https://analytics.google.com/analytics/web/>