



## FERRAMENTA PARA AVALIAÇÃO DE RISCO DE EVASÃO ACADÊMICA

**Fernando Souza Dias**

CI&T, Campinas

[fs.dias@outlook.com](mailto:fs.dias@outlook.com)

**Peter Jandl Junior**

Centro Universitário Padre Anchieta

[pjandl@anchieta.br](mailto:pjandl@anchieta.br)

### RESUMO

Atualmente, um dos grandes problemas encontrados para a conclusão dos cursos de graduação é a evasão escolar, considerada um problema social complexo que cresceu muito recentemente. Cada vez mais, dispomos de informações geradas nos cursos, as quais contêm um valioso histórico acadêmico. O objetivo deste trabalho foi desenvolver uma ferramenta que utiliza essas informações para estimar o risco de evasão, baseado no desempenho acadêmico. Em seu desenvolvimento, foram utilizadas as tecnologias PHP, Apache, MySql e Google Charts.

**Palavras-Chave:** desempenho acadêmico, desenvolvimento de software, evasão, perfil de evasão, risco de evasão.

### ABSTRACT

Nowadays one of the big problems found for term conclusion of undergraduate courses is evasion, which is a complex social problem that has grown in recent years. Even more we have the availability of data generated in courses, which has a valuable academic record. The objective of this project was to develop a tool that uses this information to estimate the risk of evasion based on academic performance. In its development were used technologies like PHP, Apache, MySql, and Google Charts.

**Keywords:** academic performance, software development, evasion, evasion profile, evasion risk.

## 1. INTRODUÇÃO

Na atual sociedade, são evidentes a importância e a necessidade da formação de pessoas qualificadas para o mercado de trabalho. Isso se deve porque, no mercado altamente competitivo de hoje, o maior capital que uma empresa ou um país pode possuir é seu capital intelectual, ou seja, as pessoas que trabalham nessa organização ou que moram nesse país e que



podem fazê-los melhores, “que fazem a diferença” (STEWART, 1998). O Brasil sofre nesse sentido, com falta de profissionais qualificados, como mostra o relatório sobre educação divulgado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), de 2011. Seus resultados mostram que, no país, somente 11% da população entre 25 e 64 anos de idade chegou à conclusão do ensino superior, quando o recomendável pelo próprio órgão é, ao menos, 31%. Nos últimos anos, com a ajuda de programas do governo, o número de inscrições nas Instituições de Ensino Superior (IES) cresceu bastante, como mostram as estatísticas divulgadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP, 2013), as quais mostram que, entre 2002 e 2012, o número de inscrições no ensino superior dobrou, superando a marca de 7 milhões de matrículas em 2012. No entanto, um dos grandes problemas encontrados para a conclusão dos cursos é a evasão escolar.

Para demonstrar a gravidade e o tamanho desse problema, podemos citar os dados apresentados por Lobo e Silva Filho (2007), os quais mostram que a média de evasão no ensino superior no Brasil, apurada no período de 2000 a 2005, foi de 22%. Um padrão semelhante é comentado por Lobo (2012, pg.32) em sua avaliação da situação da evasão no ensino superior brasileiro, apurada entre os anos de 2006 e 2009, na qual a média obtida também é de 22%. Considerando os números do Censo da Educação Superior, divulgados pelo Ministério da Educação (INEP, 2015, pg.20), no qual se verifica um total de 7.305.977 matrículas para o ano de 2013, conforme os dados apontados nas pesquisas citadas anteriormente, é possível estimar cerca de 1.607.314 evasões para esse mesmo ano.

A evasão no ensino superior é um problema social complexo (SCALI, 2009), que afeta diversos âmbitos de uma sociedade, como o econômico, já que a evasão gera prejuízos às instituições de ensino, além de causar problemas na formação de novos profissionais. No âmbito social, também verifica-se prejuízo, pois a pessoa passa a ter um menor número de oportunidades de trabalho sem a formação acadêmica e, normalmente, acaba por trabalhar em cargos com menor rentabilidade. Além disso, a evasão ainda afeta o âmbito financeiro pessoal, visto que é um péssimo investimento começar a graduação e evadir durante o curso, pois o valor pago é perdido, não se obtém o diploma para o exercício de uma profissão e nem o retorno que advém da formação acadêmica.

Por isso é necessário compreender os diversos fatores que levam um aluno a evadir. A principal razão apontada é a falta de recursos financeiros. Contudo, é importante atentar-se também para as questões de ordem acadêmica, como as expectativas do ingressante em relação



ao curso ou se a instituição de ensino pode apoiar e ajudar o discente a concluir seu curso, não somente com a oferta de alternativas de financiamento, mas também com ações pedagógicas eficazes. Outro fator comum para evasão no ensino superior é a má formação do aluno durante o ensino fundamental e médio, pois, despreparado, não consegue progredir no curso, abandonando-o.

Em cada um desses casos de evasão, os alunos deixam um histórico acadêmico valioso do ponto de vista das informações referentes a evasões anteriores nos diversos cursos, as quais podem formar um cenário para avaliação, alteração e melhoria do sistema acadêmico. Nesse caso, seria impossível fazer um levantamento manualmente, devido à grande quantidade de dados que precisa ser trabalhada.

O objetivo desse trabalho foi desenvolver uma ferramenta informatizada, que auxilie coordenadores, professores, gestores acadêmicos e demais interessados a verificar o risco de abandono, baseado em dados de desempenho escolar, na tentativa de estabelecer um padrão ou um perfil de aluno que tende à evasão no ensino superior, ofertando uma visão complementar para a análise da evasão dos cursos de graduação.

## **2. PROPOSTA DE FERRAMENTA**

A proposta da ferramenta desenvolvida baseou-se no fato de toda IES dispor de dados de desempenho acadêmico (disciplina, faltas, notas, situação de aprovação, porcentagem de presença, média das notas etc.). Também são típicos a obtenção e o cálculo de informações, como taxa de reprovação por disciplina, por semestre, e média de faltas, mas que não trazem correlação com evolução da evasão no curso e nem a correlação da reprovação nas disciplinas com a evasão.

A ferramenta foi desenvolvida tendo como base os dados do Centro Universitário Padre Anchieta, especificamente do curso de Sistemas de Informação no período de 2009 a 2012. Sua elaboração e seu desenvolvimento foram feitos de forma a tornar o software o mais genérico possível, podendo ser adaptado para aplicações em outros períodos, turmas, cursos ou instituições de ensino.

O seu desenvolvimento ocorreu utilizando-se o modelo de desenvolvimento cascata, no qual as ações acontecem de forma linear e sequencial, isto é, quando uma etapa é finalizada, ela é aprovada e, então, o desenvolvimento segue para uma nova etapa (SOMMERVILLE, 2003). Cada fase prossegue sem nenhuma sobreposição; os passos não são interativos entre si e seguem



uma ordem restrita (PRESSMAN, 2003). A ferramenta orienta-se pelos conceitos de desenvolvimento cascata, mas nem todos os procedimentos ou etapas do processo foram implementados ou aplicados.

### 3. ANÁLISE DE REQUISITOS

#### 3.1 Requisitos Funcionais

Segundo Sommerville (2003, p.83), os requisitos funcionais são “declarações de funções que o sistema deve fornecer, como o sistema deve reagir a entradas específicas e como deve se comportar em determinadas situações”. Também existem situações especiais, nas quais os requisitos funcionais podem estabelecer, de maneira explícita, o que o sistema não deve fazer.

Em suma, os requisitos funcionais (RF) descrevem os serviços que o sistema deve oferecer e como o software deve se comportar durante o uso. Os requisitos funcionais apresentados na Tabela 1 foram baseados nos objetivos preestabelecidos para a ferramenta.

**Tabela 1** – Requisitos Funcionais.

Identificador	Descrição do requisito funcional
RF01	A ferramenta deverá permitir a seleção de diferentes anos e semestres para avaliação.
RF02	A ferramenta deverá gerar um gráfico que possibilite a análise do desempenho do curso por turma.
RF03	A ferramenta deve calcular o perfil de um aluno frente a um semestre específico, através da entrada de suas reprovações
RF04	A ferramenta deverá efetuar a geração de gráficos e dos cálculos de forma dinâmica, por meio de acesso a um banco de dados.
RF05	O sistema não deve permitir a seleção de turmas, semestres ou anos não disponíveis no banco de dados.

#### 3.2 Requisitos Não Funcionais

Sommerville (2003, p.83) também define os requisitos denominados não funcionais, que são aqueles de caráter restritivo em relação aos serviços oferecidos ou às funções desempenhadas por um sistema, com ênfase nas restrições de tempo, do processo de

desenvolvimento do software, da existência de padrões, entre outras. O autor complementa da seguinte maneira:

Os requisitos não funcionais, como o nome sugere, são aqueles que não dizem respeito diretamente às funções específicas fornecidas pelo sistema. Eles podem estar relacionados a propriedades de sistema emergentes, como confiabilidade, tempo de resposta e espaço em disco. (SOMMERVILLE, 2003, p.85)

Os requisitos não funcionais (RNF) são limitações ou restrições gerais sobre o sistema, tais como limitações de tempo de processamento e restrições de custo, além de imposições tecnológicas ou de funções disponíveis. Tendem a delimitar as possibilidades de funcionamento do sistema. Os requisitos não funcionais identificados para a ferramenta proposta são listados na Tabela 2.

**Tabela 2** – Requisitos Não Funcionais.

Identificador	Descrição do requisito não funcional
RNF01	Para utilizar a ferramenta, é necessário um navegador.
RNF02	É necessário o acesso à internet para geração do gráfico de desempenho por turma.
RNF03	A ferramenta deve ser executada em um servidor Web.
RNF04	As ferramentas utilizadas para o desenvolvimento devem ser de custo financeiro reduzido, preferencialmente ferramentas livres.
RNF05	A ferramenta deve ser simples de ser utilizada.
RNF06	A ferramenta deve restringir as entradas a entradas válidas.
RNF07	Deve ser utilizado um banco de dados relacional para consulta dos dados.
RNF08	O custo e o tempo para colocar a ferramenta em funcionamento devem ser baixos.

#### 4. MODELAGEM DA FERRAMENTA

A Unified Modeling Language (UML) é uma linguagem universal para a especificação de softwares, constituída de uma série de símbolos e diagramas próprios, a qual é utilizada para especificar funções, ações e a classe de um sistema. Pode ser utilizada para documentação das fases de desenvolvimento, análise e levantamento de requisitos, até as partes finais, como manutenção do software. Ao descrever a UML, Medeiros (2004, p.10) complementa: “Percebemos [...] que a UML não nos indica como devemos fazer um software. Ela indica apenas as formas que podem ser utilizadas para representar um software em diversos estágios de desenvolvimento.”. A modelagem UML agrega também outros benefícios, como a

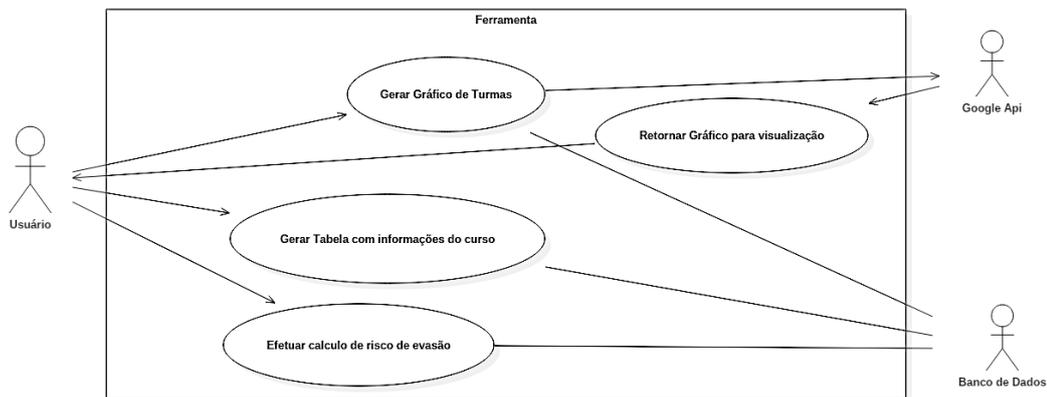
possibilidade de desenvolver um software de forma mais rápida e eficiente, reduzindo o desperdício e o retrabalho para se finalizar esse desenvolvimento (MATOS, 2003).

Dentre os vários artefatos existentes na UML, o diagrama de casos de uso é um dos mais importantes, pois pode ser compreendido por pessoas leigas e, ainda assim, é essencial para estabelecer as funcionalidades, comportamentos e interações desejados de um sistema.

#### 4.1 Casos de Uso

O diagrama de casos de uso é um dos diagramas da UML utilizados para especificar as ações do sistema do ponto de vista do usuário. São mostradas, através dele, as funções e interações do sistema, porém sem o aprofundamento técnico de como cada interação ocorre.

Na Figura 1 encontra-se o diagrama de casos de uso da ferramenta desenvolvida.



**Figura 1** – Casos de uso da ferramenta.

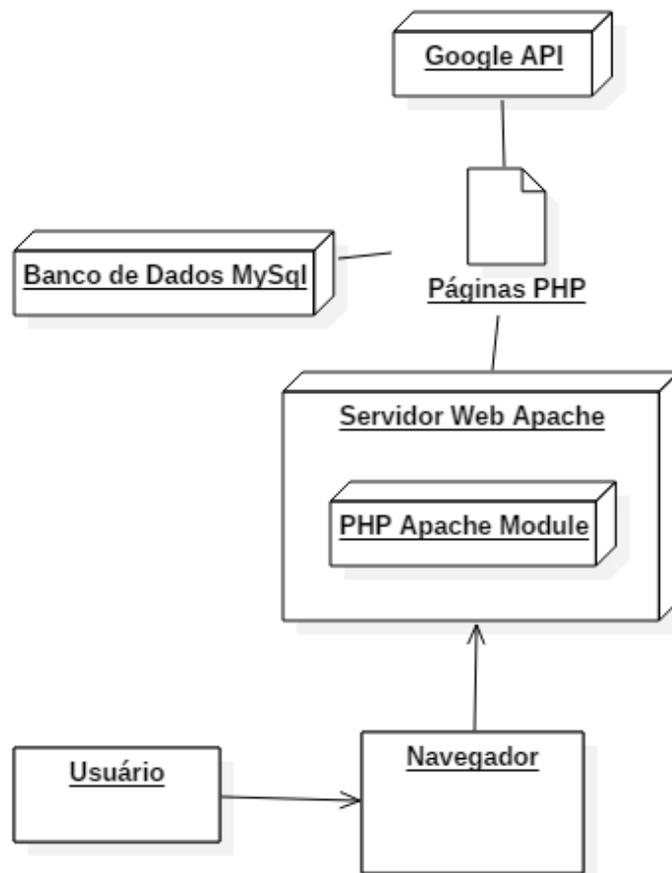
Pode-se notar que o caso de uso da ferramenta conta com três atores: o *usuário*, pessoa que utilizará a ferramenta, o *banco de dados* relacional, no qual as informações estão armazenadas e as consultas serão executadas para apresentação posterior na ferramenta, e o *Google API* (Application Programming Interface), que é o framework de gráficos do Google, o qual recebe as informações e configurações necessárias, e retorna para a ferramenta o gráfico finalizado.

Existem quatro casos de uso: (i) *gerar gráfico de turmas*, no qual é possível visualizar o gráfico de bolhas, montado a partir dos dados de uma turma específica; (ii) *retornar gráfico para visualização*, que ocorre após o processamento dos dados por parte da API Google Charts, na volta do gráfico para exibição ao usuário; (iii) *gerar tabela com informações do curso*, cujo uso ocorre quando o usuário seleciona o semestre do qual deseja visualizar as informações em

formato tabela do curso ao longo de um período; (iv) *efetuar cálculo de risco de evasão*, que permite ao usuário, a partir da seleção das reprovações, determinar o risco de evasão do aluno.

#### 4.2 Diagrama de Blocos

A Figura 2 ilustra o diagrama de blocos da ferramenta, mostrando a divisão dos principais blocos existentes e como são feitos os acessos. Percebe-se, pelo diagrama, que o usuário acessa o navegador; o navegador acessa o servidor de aplicação web Apache buscando a página solicitada pelo usuário; caso o servidor Apache encontre a página, ela é encaminhada ao PHP Apache Module, um módulo instalado no servidor Apache para interpretação do código PHP. Após essa interpretação, a página é aberta e, a partir das interações do usuário com as páginas PHP, ocorrem acessos – tanto ao banco de dados, como na API de gráficos do Google. Além disso, são feitos a montagem do conteúdo para exibição no navegador e os cálculos de funções dos demais fatores da ferramenta.



**Figura 2** – Diagrama de blocos da ferramenta.



## **5. TECNOLOGIAS UTILIZADAS**

### **5.1 PHP**

A linguagem de programação escolhida para o desenvolvimento do software foi o PHP, um acrônimo recursivo para PHP Hypertext Preprocessor. O PHP é uma linguagem interpretada, open source, muito utilizada para o desenvolvimento web, criada por Rasmus Lerdorf, engenheiro de software e membro da equipe Apache, com o objetivo de controlar o acesso às páginas web e retorno de páginas que seriam geradas de forma dinâmica (DAVIS e PHILLIPS, 2008; SOARES, 2010).

A sua escolha, neste projeto, ocorreu devido a algumas características da linguagem, como o baixo custo de aprendizagem. A linguagem PHP incorpora características de outras linguagens, como C++ e Java, por isso o custo de aprendizagem é baixo. Além disso, o PHP é simples de ser configurado, suas aplicações não ficam limitadas a uma plataforma específica e permitem acesso a diversos bancos de dados (DAVIS e PHILLIPS, 2008; SOARES, 2010). O PHP possui uma grande comunidade que auxilia na melhoria da linguagem, correções de defeitos, e documentação de funções e aplicações possíveis. Ainda é possível inserir os códigos PHP dentro do próprio código HTML. E como descreveu Soares: “Basicamente, qualquer coisa que pode ser feita por algum programa CGI, também pode ser feita com PHP” (2010, p.4).

### **5.2 Apache**

O Apache é o servidor web livre de maior uso no mundo; ele foi criado em 1995, por Rob McCool, e atualmente é mantido pela Apache Software Foundation e por uma comunidade de desenvolvedores voluntários, espalhados ao redor do mundo. O Apache funciona no modelo cliente servidor, no qual o cliente acessa as páginas alocadas no servidor Apache através de um navegador compatível com o protocolo Hyper Text Transfer Protocol (HTTP). O Apache é conhecido por ser um servidor web HTTP robusto, com serviços de grande abrangência, cobrindo desde problemáticas relacionadas à segurança até serviços de criação de domínios virtuais. Além do citado, o Apache possui suporte à instalação do PHP module, o que lhe possibilita efetuar a interpretação dos códigos PHP sem a necessidade de qualquer acesso externo, o que torna o processamento de suas páginas ainda mais rápido e com menor custo.

Sua escolha para o projeto baseia-se na sua velocidade de respostas, praticidade de configuração, utilização multiplataforma, em ser um produto open-source e possuir um



mecanismo que viabiliza a interpretação de páginas PHP de forma rápida, sem acesso externo para isso.

### 5.3 MySql

O sistema utiliza o MySql, um sistema gerenciador de banco de dados relacional (SGBDR), desenvolvido pela TCX, em 1996. Atualmente, é mantido pela Oracle. A motivação para o desenvolvimento do MySql se deu porque era necessário um SGBDR capaz de tratar grandes quantidades de dados e gerenciar informações de maneira inteligente (DAVIS e PHILLIPS, 2008), sem demandar um sistema computacional de alto desempenho e, dessa forma, que pudesse ser utilizado em máquinas de menor custo possível. O MySql é, atualmente, o banco de dados de código aberto mais popular do mundo, sendo empregado nas mais diversas aplicações, pois fornece ferramentas para consulta, armazenamento, manipulação e padronização dos dados através de linguagem SQL (Structured Query Language).

O MySql foi escolhido como componente da ferramenta construída devido a ser um software open source, por suportar a manipulação de grandes quantidades de dados, como bancos com até 50 milhões de registros (MYSQL, 2013). Também é multiplataforma, pois pode ser utilizado em sistemas Unix, Linux e Windows, é compatível com a linguagem SQL, apresenta baixo custo de implementação e é considerado um SGBDR rápido e seguro.

### 5.4 Google Charts

O Google Charts é uma API do Google para construção de gráficos em aplicações web, cujo uso foi ampliado inclusive para fins comerciais. Com uma excelente documentação online, e uma grande quantidade de gráficos e configurações possíveis a cada gráfico (GOOGLE, 2015), o funcionamento dessa API se dá por meio do uso de JavaScript Object Notation (JSON).

Dessa maneira, o cliente pode ser desenvolvido em qualquer linguagem que gere um elemento JSON com as informações solicitadas para o gráfico escolhido, passando para a API do Google informações como o título do gráfico, as legendas, os dados para composição e outros dados necessários. Ao mesmo tempo, a aplicação deve fornecer uma divisão (com uso da tag HTML <div>) em sua página web, na qual o gráfico será exibido. Assim, as informações são enviadas à API do Google, na qual são processadas para gerar o gráfico desejado, como uma imagem a ser exibida na página web.





ou várias fontes, nas quais não é possível lidar com os dados, para um banco de dados centralizado, no qual possam ser utilizados em nível computacional.

A ferramenta ETL escolhida foi o Kettle, também conhecida como Pentaho Data Integration, da empresa de mesmo nome. Desenvolvido desde 2004, o Kettle é um conjunto de ferramentas e aplicativos, que permite manipulações de dados através de múltiplas fontes, como banco de dados, arquivo de texto e planilhas em formato Excel, entre muitos outros. Um dos focos do Kettle é a facilidade de utilização por parte dos usuários, portanto o software conta com diversos recursos visuais, e sua utilização e os aprendizados são relativamente rápidos e simples. O Kettle conta com duas versões: a Enterprise Edition, voltada para soluções corporativas, a qual, embora paga, possui uma utilização possível em modo Trial (gratuito para testes durante 30 dias); e a versão Community, que é um projeto Open Source de uso gratuito. A versão Community atendeu completamente os requisitos para a transformação dos dados necessários ao projeto. Dessa forma, o Kettle ainda passa a ter outra vantagem: seu uso gratuito.

Com o auxílio do Kettle, os dados foram transportados para o banco de dados MySQL. Após a importação, eles apresentavam alguns registros inconsistentes, fora do padrão e sem tipagem adequada.

Utilizando-se da linguagem SQL, foram excluídos os registros que apresentavam qualquer inconsistência, como notas maiores do que dez ou menores que zero, além de casos como registros do tipo texto em colunas de faltas ou notas. Após a exclusão desses registros, foram determinados os tipos corretos para cada coluna do banco de dados, determinando as colunas numéricas e colunas de texto. Ao concluir as padronizações e correções nos registros, os dados estavam prontos para uso da ferramenta.

## **7. DESENVOLVIMENTO DA FERRAMENTA**

Para iniciar o desenvolvimento da ferramenta, foi necessário decidir quais informações seriam utilizadas e quais precisariam ser geradas. As informações utilizadas na ferramenta foram: identificador do aluno (que não é exposto em nenhum momento), ano, semestre, disciplina, nome da disciplina e nota final. Com base nesses dados básicos, foram utilizadas consultas SQL para elaboração da lógica da ferramenta. A lógica da aplicação ficou toda nas consultas SQL, o que facilita a portabilidade do código, visto que a preocupação da programação, nesse caso, fica para exibição dos resultados.

A linguagem SQL foi criada para que fosse possível efetuar consultas de forma simples e padronizadas em bancos de dados. A simplicidade da linguagem fez com que ela rapidamente se tornasse um padrão para os bancos de dados e, atualmente, sua utilização cobre praticamente todos os bancos de dados relacionais (RAMALHO, 1999). Assim o SQL possibilita que os dados sejam consultados e, por meio de cruzamentos e lógicas, novos dados que não estavam na tabela sejam gerados nas consultas, e cálculos sejam efetuados.

Na Figura 4, pode-se ver um exemplo de uma das consultas utilizadas na ferramenta. Essa consulta extrai a informação do número de abandonos que ocorreram em uma turma, em um determinado período.

```

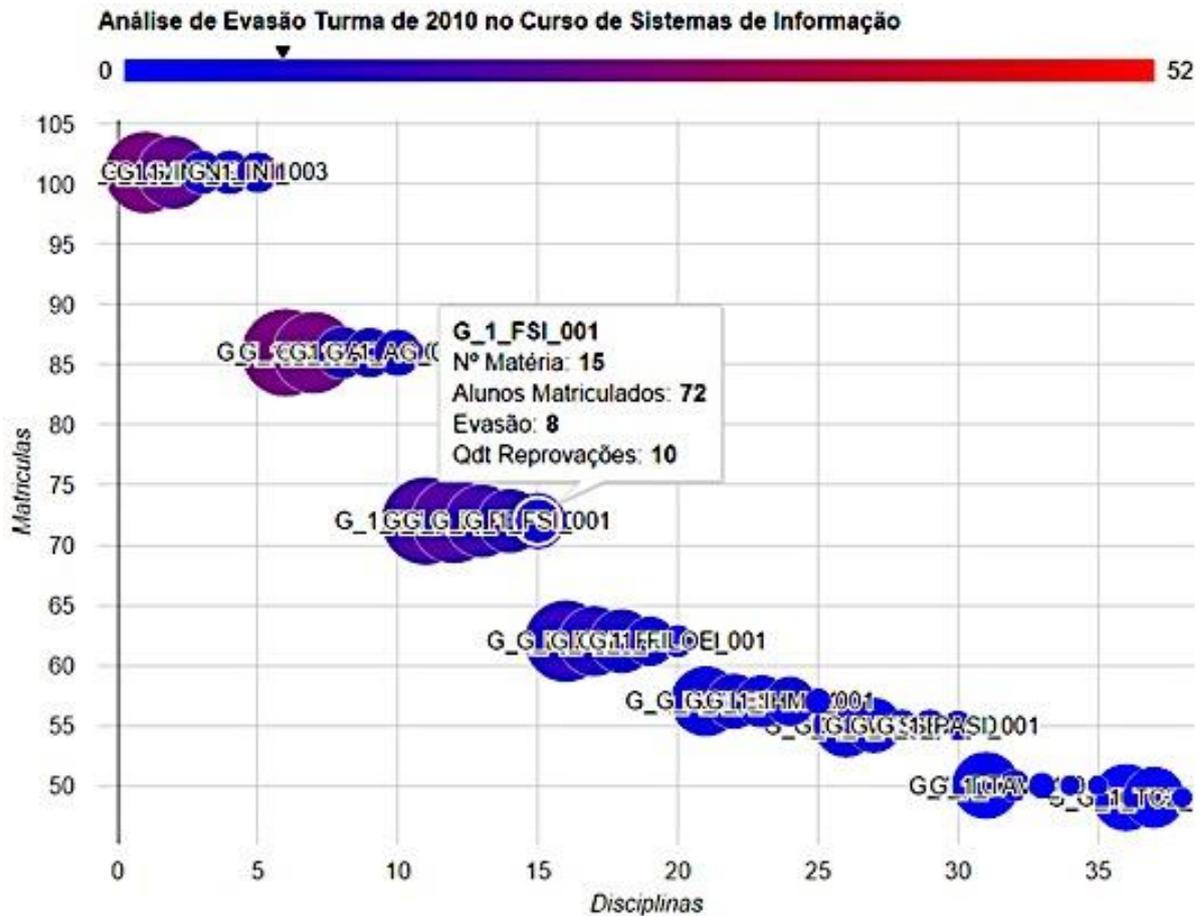
1 • SELECT COUNT(distinct aluno) as abandonos
2 FROM alunos_notas_faltas
3 WHERE aluno BETWEEN 1100000 AND 1199999 AND
4 ano = 2011 AND semestre = 1 AND aluno
5 NOT IN (SELECT DISTINCT aluno
6 FROM alunos_notas_faltas
7 WHERE aluno BETWEEN 1100000 AND 1199999 AND ano = 2014
8 AND semestre = 1);

```

**Figura 4** – Query que retorna o número de abandonos em um período.

Pode-se perceber que são listados os discentes matriculados no primeiro semestre de 2011, com registro de aluno (RA) entre 1100000 e 1199999, e depois são comparados com os alunos de mesmo RA, porém, que estejam matriculados no primeiro semestre de 2014. Os alunos que estavam matriculados em 2011 e não estão matriculados em 2014 são os que abandonaram o curso nesse período.

Através de consultas, como a ilustrada na Figura 4, foi possível trazer todas as informações necessárias para elaboração do gráfico de bolhas, que possibilita a análise de evasão por turma, mostrando a curva de evasão semestre a semestre. Esse gráfico é visto na Figura 5.



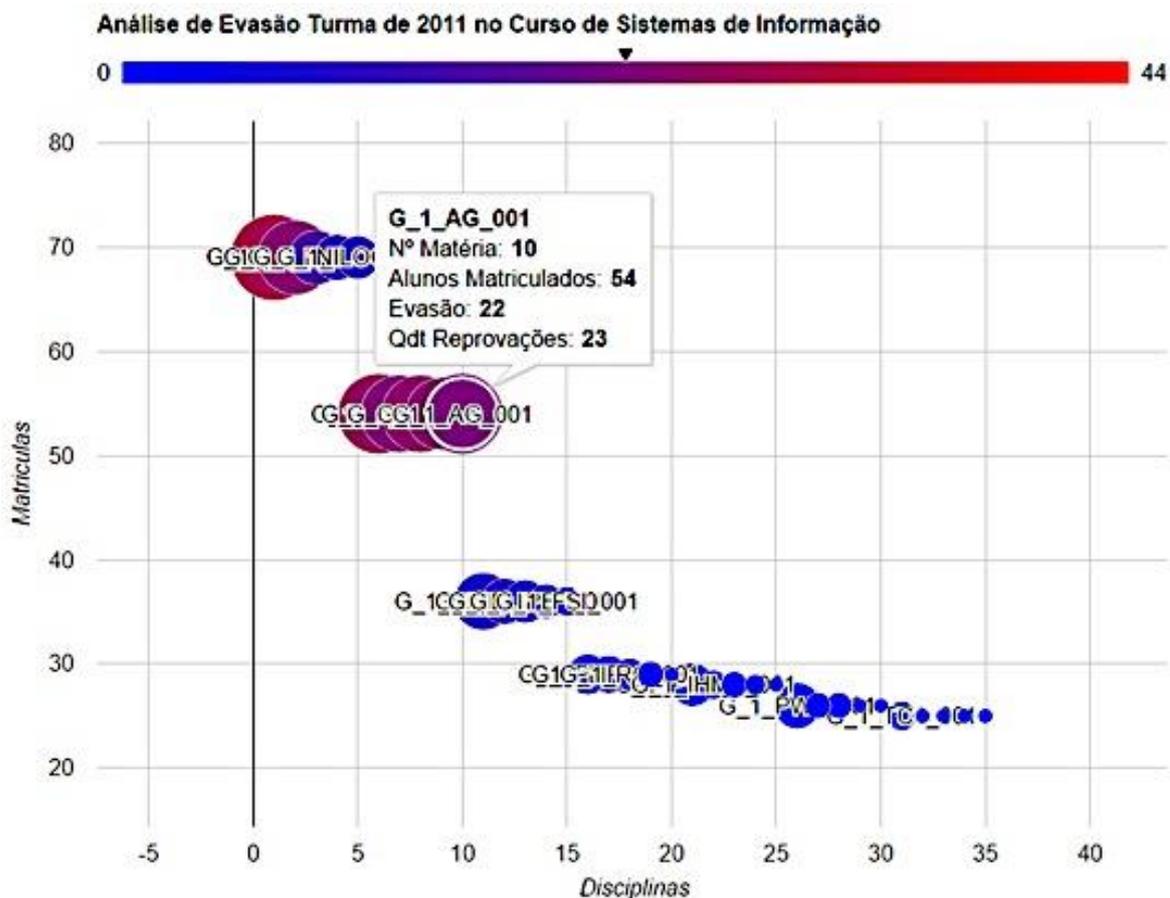
**Figura 5** – Gráfico de bolhas, análise de evasão por turma (2010).

O gráfico de bolhas da Figura 5 foi gerado pela ferramenta para a turma de 2010, do curso de Sistemas de Informação, do Centro Universitário Padre Anchieta, utilizando a API do Google Charts. Em seu eixo de domínio, figura o número de matrículas em cada disciplina. As disciplinas (bolhas individuais) ficam agrupadas por semestre, permitindo observar, ao final de cada período, que há uma diminuição no número de matrículas, o que corresponde à evasão ocorrida naquele semestre. O eixo das abscissas corresponde a uma numeração sequencial das disciplinas, para que fiquem organizadas, facilitando a visualização. Como é possível que as bolhas se sobreponham, dificultando a leitura dos códigos das disciplinas, ao posicionar o mouse sobre uma bolha, é exibida uma janela auxiliar com seu código e outras informações associadas.

O tamanho de cada bolha é determinado pelo número de reprovações que aconteceram na disciplina. Existe, na parte superior do gráfico, uma legenda de cor da bolha: na Figura 5 vai de 0 a 52 (evasão máxima aferida em todo período avaliado). Assim as cores das bolhas, tal

como seu tamanho, refletem quantos foram os alunos reprovados na disciplina e evadidos do curso, mostrando, assim, o risco que essa disciplina gera em relação à evasão total, pois as reprovações que não constituem evasão são desconsideradas aqui.

Selecionando uma das bolhas, outras informações tornam-se disponíveis: o código da disciplina, o número sequencial da disciplina, a quantidade de alunos que se matricularam na disciplina, o número de alunos que reprovaram na disciplina e evadiram o curso, e também a quantidade de reprovações na disciplina. Na Figura 6 temos o mesmo gráfico gerado pela ferramenta para a turma de 2011.



**Figura 6** – Gráfico de bolhas, análise de evasão por turma (2011).

Analisando as Figuras 5 e 6, é possível notar um padrão de curva de evasão ao longo do curso, no qual é visível que o número de evasões é muito maior nos primeiros semestres, tendendo a uma estabilização nos últimos semestres do curso; o número de reprovações, que determina o tamanho da bolha, também é drasticamente reduzido com o passar dos semestres, e as cores das bolhas se tornam cada vez mais próximas da marca 0 (zero), mostrando que o

risco de evasão por reprovar nas últimas disciplinas do curso, cruzado com a evasão total, é relativamente baixo.

A outra parte da ferramenta diz respeito ao cálculo de risco de evasão. A forma de desenvolvimento aplicada é a mesma que foi utilizada na análise de evasão por turma; a lógica novamente é constituída através de consultas SQL, e a ferramenta efetua a formatação e exibição dos dados.

No caso do cálculo de evasão, não é gerado um gráfico, mas, sim, uma tabela que possibilita visualizar informações sobre o curso, contemplando todos os dados armazenados no banco de dados e não apenas os de um ano específico do curso. É possível a seleção do semestre que deseja visualizar e, posteriormente, a seleção de reprovações de um aluno para efetuar o cálculo de risco de evasão.

A Tabela 3 foi gerada pela ferramenta, e nela, são vistos os dados do curso de Sistemas de Informação, do Centro Universitário Padre Anchieta, com intervalo de apuração entre os anos de 2009 e 2012, para o 2º semestre do curso.

**Tabela 3** – Dados de 2009 a 2012, do 2º semestre do curso de Sistemas de Informação.

**Dados do Período de 2009 a 2012 do Curso de Sistemas de Informação 2º Semestre**

Disciplina	Total Reprovações	Rep Disciplina abandonaram	Evadiram e reprov Total	Total evasão	Evasão Própria	Evasão Total
ALGORITMOS II	112	42	86	94	37.5%	44.68%
ALGORITMOS I	97	57	86	94	58.76%	60.64%
CÁLCULO DIFERENCIAL E INTEGRAL	85	42	86	94	49.41%	44.68%
MATEMÁTICA DISCRETA	67	47	86	94	70.15%	50%
ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES	64	39	86	94	60.94%	41.49%
INTRODUÇÃO À PROGRAMAÇÃO	62	39	86	94	62.9%	41.49%
ADMINISTRAÇÃO GERAL	47	35	86	94	74.47%	37.23%
INTRODUÇÃO À INFORMÁTICA	40	37	86	94	62.5%	39.36%
INGLÊS INSTRUMENTAL	36	33	86	94	61.67%	35.11%
INTRODUÇÃO À LÓGICA	35	32	86	94	61.43%	34.04%

Nível de relação com a evasão ■ Crítico ■ Alto ■ Médio ■ Baixo

Para que a ferramenta gere esses dados, é necessário apenas selecionar o semestre desejado. No caso da Tabela 3, o semestre selecionado foi o 2º semestre.

As informações disponíveis são: o nome da disciplina, o total de reprovação na disciplina, o número de alunos que reprovaram na disciplina e abandonaram o curso, o total de alunos que reprovaram em qualquer disciplina e evadiram o curso, a evasão total no período contando alunos que reprovaram em alguma disciplina e os que não reprovaram, a porcentagem de evasão própria da disciplina e a porcentagem de evasão total.

A evasão própria é o resultado de um cálculo, que leva em conta o número de reprovações na disciplina e o número de alunos que reprovaram na disciplina e abandonaram. Explicando melhor, por exemplo, pode-se ver, na Tabela 3, que o curso obteve um total de 97 reprovações em Algoritmos I, e que 57 alunos que reprovaram em Algoritmos I não se matricularam no 3º semestre do curso. Vale citar que os dados da tabela para evasão sempre levam em conta a evasão com relação ao semestre seguinte. Assim, observa-se que o risco de evasão, considerando-se apenas a própria disciplina, é a relação dos abandonos com reprovação na disciplina pelas reprovações na mesma disciplina, ou seja:  $(\text{Evasão com reprovação na disciplina})/(\text{Reprovações na disciplina}) \times 100$  no exemplo:  $(57/97) \times 100 = 58,76\%$

Já a evasão total é o resultado do cálculo da quantidade de alunos que reprovaram na disciplina e abandonaram o curso dividido pela evasão total no período; assim, usando novamente o exemplo de Algoritmos I, vê-se que houve 94 evasões no total e 57 abandonos com reprovação na disciplina. Dessa forma, tem-se o seguinte cálculo:  $(\text{Evasão com reprovação na disciplina})/(\text{Total de evasão no período}) \times 100$  no exemplo:  $(57/94) \times 100 = 60,64\%$

**Tabela 4** – Dados de 2009 a 2012, do 4º semestre do curso de Sistemas de Informação.

**Dados do Período de 2009 a 20012 do Curso de Sistemas de Informação 4º Semestre**

Disciplina	Total Reprovações	Rep Disciplina abandonaram	Evadiram e reprov Total	Total evasão	Evasão Própria	Evasão Total
ALGORITMOS II	113	62	122	129	54,87%	48,06%
ALGORITMOS I	99	71	122	129	71,72%	55,04%
CÁLCULO DIFERENCIAL E INTEGRAL	85	60	122	129	70,59%	46,51%
SISTEMAS OPERACIONAIS	76	31	122	129	40,79%	24,03%
ESTRUTURAS DE DADOS	73	30	122	129	41,1%	23,26%
MATEMÁTICA DISCRETA	68	54	122	129	79,41%	41,86%
ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES	65	46	122	129	70,77%	35,66%
INTRODUÇÃO À PROGRAMAÇÃO	64	46	122	129	71,85%	35,66%
LINGUAGEM DE PROGRAMAÇÃO II	63	11	122	129	17,46%	8,53%
LINGUAGEM DE PROGRAMAÇÃO I	60	29	122	129	48,33%	22,48%
PROBABILIDADE E ESTATÍSTICA	53	23	122	129	43,4%	17,83%
PLATAFORMAS DE DESENVOLVIMENTO	49	11	122	129	22,45%	8,53%
ADMINISTRAÇÃO GERAL	48	42	122	129	87,5%	32,56%
INTRODUÇÃO À INFORMÁTICA	42	39	122	129	82,86%	30,23%
INGLÊS INSTRUMENTAL	37	35	122	129	94,59%	27,13%
INTRODUÇÃO À LÓGICA	36	33	122	129	91,67%	25,58%
BANCO DE DADOS I	35	9	122	129	25,71%	6,98%
REDES DE COMPUTADORES	32	10	122	129	31,25%	7,75%
FUNDAMENTOS DE SISTEMAS DE INFORMAÇÃO	28	19	122	129	87,86%	14,73%
FILOSOFIA E ÉTICA	13	7	122	129	53,85%	5,43%

Nível de relação com a evasão: ■ Crítico ■ Alto ■ Médio ■ Baixo

A Tabela 4 contém os dados do 4º semestre, na qual se nota que, com o passar do curso, os números apresentados sofrem alterações. A disciplina usada no exemplo, Algoritmos I, possui 99 reprovações registradas, um acréscimo de duas reprovações em relação à Tabela 3, e o número de reprovação com evasão nessa disciplina aumentou bastante: agora, são 71 alunos

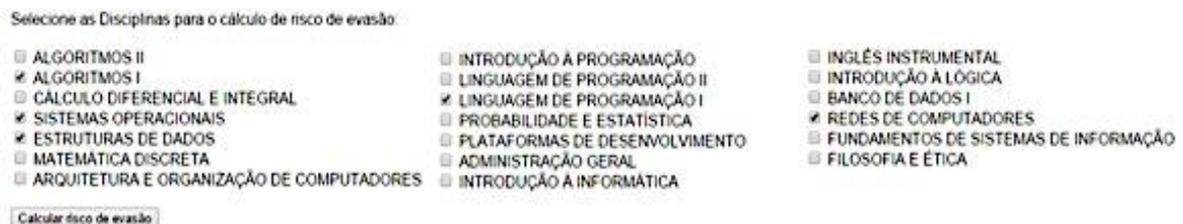
reprovados na disciplina e que abandonaram o curso, e a evasão total também aumentou; pode-se ver que 129 alunos abandonaram o curso no período avaliado, antes do 5º semestre.

Por causa dessas alterações, os índices de evasão própria e evasão total também foram alterados. Agora, percebe-se que os índices de evasão própria das disciplinas do 1º e 2º semestres sofreram um aumento, devido ao fato de a quantidade de reprovações ter se mantido muito próxima à apresentada anteriormente, porém a quantidade de alunos que reprovaram na disciplina mesmo nos primeiros semestres e evadiram o curso aumentou muito. Isso se deve ao fato de uma reprovação, mesmo nos semestres iniciais, influenciar o aluno na evasão do curso em semestres posteriores.

Já os índices de evasão total caíram, visto que, a cada semestre, há mais evasões totais e o peso do aumento das evasões totais contra o aumento da evasão na disciplina acaba sendo maior. Dessa forma, a evasão total, que verifica o impacto da evasão em uma disciplina em relação a todo o período avaliado, torna-se menor quanto mais semestres do curso são considerados, pois o peso de uma disciplina na evasão se torna menor. No caso da evasão própria, isso não ocorre pelo fato de o peso ser calculado em relação à própria disciplina.

Utilizando-se essas informações apresentadas, foi gerado um cálculo para averiguar o risco de evasão de um aluno dado a suas reprovações.

O usuário seleciona as disciplinas que o aluno reprovou ou, no caso de uma simulação futura, as disciplinas que o aluno tende a reprovar, e clica em calcular risco de evasão, conforme a Figura 7.



Selecione as Disciplinas para o cálculo de risco de evasão:

<input type="checkbox"/> ALGORITMOS II	<input type="checkbox"/> INTRODUÇÃO À PROGRAMAÇÃO	<input type="checkbox"/> INGLÊS INSTRUMENTAL
<input checked="" type="checkbox"/> ALGORITMOS I	<input type="checkbox"/> LINGUAGEM DE PROGRAMAÇÃO II	<input type="checkbox"/> INTRODUÇÃO À LÓGICA
<input type="checkbox"/> CÁLCULO DIFERENCIAL E INTEGRAL	<input checked="" type="checkbox"/> LINGUAGEM DE PROGRAMAÇÃO I	<input type="checkbox"/> BANCO DE DADOS I
<input checked="" type="checkbox"/> SISTEMAS OPERACIONAIS	<input type="checkbox"/> PROBABILIDADE E ESTATÍSTICA	<input checked="" type="checkbox"/> REDES DE COMPUTADORES
<input checked="" type="checkbox"/> ESTRUTURAS DE DADOS	<input type="checkbox"/> PLATAFORMAS DE DESENVOLVIMENTO	<input type="checkbox"/> FUNDAMENTOS DE SISTEMAS DE INFORMAÇÃO
<input type="checkbox"/> MATEMÁTICA DISCRETA	<input type="checkbox"/> ADMINISTRAÇÃO GERAL	<input type="checkbox"/> FILOSOFIA E ÉTICA
<input type="checkbox"/> ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES	<input type="checkbox"/> INTRODUÇÃO À INFORMÁTICA	

**Figura 7** – Seleção de disciplinas para cálculo de evasão, 4º semestre.

A Figura 7 ilustra um exemplo de marcação (seleção) possível de disciplinas. Após clicar no botão de calcular risco de evasão, a ferramenta efetuará o cálculo e apresentará o resultado conforme Figura 8.



## Propenso à Evasão

**Risco Calculado: 58%**

**Figura 8** – Resultado da ferramenta para cálculo da Figura 7.

O resultado do risco de evasão é apresentado em formato de porcentagem junto com uma classificação de risco, nesse exemplo, a classificação foi de aluno propenso à evasão. Esse cálculo é efetuado a partir da soma de todos índices de evasão próprias das disciplinas selecionadas dividido pela quantidade de semestres cursados, ou seja:  $(Evasão\_Própria\_1 + Evasão\_Própria\_2 + Evasão\_Própria\_N) / (Quantidade\ de\ semestre\ cursados)$  no exemplo:  $(71,72+40,79 +41,1+48.33+31,25)/4=58\%$

Dessa forma, as reprovações individualmente perdem peso a cada semestre, pois o valor do divisor é aumentado a cada semestre que o aluno cursa, porém, novas reprovações aumentam o risco de evasão, pois seus valores de evasão própria são adicionados ao dividendo. Vale ressaltar que esse cálculo busca mostrar o risco de evasão de um aluno para o semestre seguinte; no exemplo, esse é o risco de o aluno não voltar a se matricular no 5º semestre do curso.

## 8. CONCLUSÕES

Com o desenvolvimento da ferramenta apresentada, podemos efetuar a análise de desempenho e o cálculo de risco de evasão de forma satisfatória, podendo ser projetado o risco de evasão de um aluno, baseado em possíveis reprovações. A ferramenta pode ser utilizada por coordenadores, professores ou gestores acadêmicos, para auxiliar na tomada de decisões em relação a alunos que apresentem reprovações, antes de evadirem dos cursos.

As tecnologias escolhidas para o desenvolvimento da ferramenta apresentaram um ótimo desempenho, tanto nos requisitos mais abstratos, como tempo de aprendizado *versus* utilização dos recursos com eficiência, quanto em requisitos diretos, como velocidade de execução das funções da ferramenta, custo financeiro das tecnologias e hardware necessário para execução.

A pesquisa e ferramenta desenvolvidas têm valor devido à sua capacidade de auxiliar os cursos na obtenção de um risco de evasão e na avaliação de disciplinas que causam maior impacto no curso, podendo ser usado inclusive para sua remodelação ou alterações nas disciplinas que apresentem maior dificuldade aos alunos. Além disso, há contribuição da



pesquisa para a construção de novas ferramentas que possam ser criadas a partir das tecnologias e dos cálculos aqui apresentados, servindo, assim, como referência para desenvolvimentos posteriores.

A ferramenta apresentada ainda possui algumas características que podem ser refinadas, para que sua aplicação e utilização se tornem mais amplas. Dentre algumas melhorias possíveis num futuro, estão a implementação de uma função para associação de disciplinas de diferentes grades curriculares do mesmo curso, apontando quais são semelhantes, apesar da diferença de nomenclatura. Também poderiam ser acrescentadas algumas parametrizações, tais como quantidade de disciplinas no semestre e seleção de período desejado, baseado no desenvolvimento já efetuado.

Por fim, é visível a relação das evasões com as reprovações nas disciplinas, e notório também que nem todas as evasões são oriundas dessa situação de reprovação, porém, de forma inegável, quanto maior a quantidade de reprovações, maior o risco de o curso não ser concluído. É possível afirmar que os objetivos anteriormente propostos foram atingidos e, ao final do trabalho, vê-se uma série de padrões e relações interessantes, que não eram perceptíveis sem a aplicação e utilização da ferramenta desenvolvida.

## 9. REFERÊNCIAS BIBLIOGRÁFICAS

**APACHE.** *Documentação do Servidor HTTP Apache Versão 2.4.* Disponível em: <<https://httpd.apache.org/docs/2.4/>>. Acesso em: 27 Setembro 2015.

**COREY, Michael et al.** *Oracle 8i Data Warehouse.* Tradução de João Tortello. Rio de Janeiro: Campus, 2001.

**DAVIS, Michele E.; PHILLIPS Jon A.** *Aprendendo PHP & MySQL.* 2ª Ed. Rio de Janeiro: Alta Books, 2008.

**GOOGLE.** *Using Google Charts.* Disponível em: <<https://developers.google.com/chart/interactive/docs/>>. Acesso em: 25 Setembro 2015.

**INEP.** *Censo da educação superior 2013: coletiva. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira,* 2014. Disponível em <[http://download.inep.gov.br/educacao\\_superior/censo\\_superior/apresentacao/2014/coletiva\\_censo\\_superior\\_2013.pdf](http://download.inep.gov.br/educacao_superior/censo_superior/apresentacao/2014/coletiva_censo_superior_2013.pdf)>. Acesso em 10 Outubro 2015.



**INEP.** *Censo da educação superior 2013: resumo técnico.* Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2015. Disponível em: <[http://download.inep.gov.br/download/superior/censo/2013/resumo\\_tecnico\\_censo\\_educacao\\_superior\\_2013.pdf](http://download.inep.gov.br/download/superior/censo/2013/resumo_tecnico_censo_educacao_superior_2013.pdf)>. Acesso em 10 Outubro 2015.

**LOBO E SILVA FILHO, Roberto Leal et al.** *A evasão no ensino superior brasileiro* IN Cadernos de Pesquisa. Instituto Lobo para o Desenvolvimento da Educação, da Ciência e da Tecnologia. São Paulo, v. 37, n. 132, p. 641-659, 2007. Disponível em: <<http://dx.doi.org/10.1590/S0100-15742007000300007>>. Acesso em: 10 Outubro 2015.

**MATOS, Alexandre Veloso de.** *UML: Prático e Descomplicado.* 3ª Ed. São Paulo: Érica, 2003.

**MEDEIROS, Ernani S.** *Desenvolvendo software com UML 2.0 definitivo.* 1ª Ed. São Paulo: Pearson Markron Books, 2004.

**MYSQL.** *MySQL 5.6 Reference Manual.* Disponível em: <<http://dev.mysql.com/doc/refman/5.6/en/>>. Acesso em: 02 Outubro 2015.

**OECD** (2011). *Education at a Glance 2011: OECD Indicators.* OECD Publishing, Paris. Disponível em: <<http://dx.doi.org/10.1787/eag-2011-en>> Acesso em: 12 Outubro 2015.

**PHP.** *Manual do PHP.* Disponível em: <[https://secure.php.net/manual/pt\\_BR/index.php](https://secure.php.net/manual/pt_BR/index.php)>. Acesso em: 02 Outubro 2015.

**PRESSMAN, Roger S.** *Engenharia de Software.* 5ª Ed. São Paulo: MCGRAW-HILL, 2003.

**RAMALHO, José A. A.** *SQL: A linguagem dos Bancos de Dados.* 1ª Ed. São Paulo: Berkeley Brasil, 1999.

**ROLDÁN, María C.** *Pentaho Data Integration (Kettle) Tutorial.* Disponível em: <[http://wiki.pentaho.com/display/EAI/Pentaho+Data+Integration+\(Kettle\)+Tutorial](http://wiki.pentaho.com/display/EAI/Pentaho+Data+Integration+(Kettle)+Tutorial)>. Acesso em: 06 Setembro 2015.

**SCALI, Danyelle Freitas.** *Evasão nos Cursos Superiores de Tecnologia: a Percepção dos Estudantes sobre seus Determinantes.* 2009. 140f. Dissertação (Mestrado em Educação) Faculdade de Educação. São Paulo: Universidade Estadual de Campinas, 2009.

**SOARES, Bruno A. L.** *Aprendendo a linguagem PHP.* 1ª Ed. São Paulo: Novatec Editora, 2010.

**SOMMERVILLE, Ian.** *Engenharia de Software.* 6ª Ed. São Paulo: Pearson Education, 2003.



**STEWART, Tomas A.** *Capital intelectual: a nova vantagem competitiva das empresas.* 2ª Ed. Rio de Janeiro: Campus, 1998.