



Universidade Federal do Rio Grande do Norte
Instituto Metr pole Digital

SmartMetropolis – Plataformas e Aplica es para Cidades Inteligentes

WP5 - Middleware

Portal de Dados Abertos da Prefeitura do Natal

Natal-RN, Brasil
Novembro de 2016

Equipe Técnica

Docentes

Prof.^a Dra. Thais Vasconcelos Batista (Coordenadora) - DIMAp-UFRN

Prof. Dr. Everton Ranielly de Sousa Cavalcante - DIMAp-UFRN

Prof. Dr. Frederico Araújo da Silva Lopes - IMD-UFRN

Prof. Dr. Nélío Alessandro Azevedo Cacho - DIMAp-UFRN

Prof. Dr. Gibeon Soares de Aquino Junior - DIMAp-UFRN

Discentes

Gusttavo Henrique Nascimento Sousa Silva

Sumário

1 Introdução	3
2 Objetivo	4
3 Desenvolvimento do Portal	5
3.1 Arquitetura do CKAN	5
3.2 Instalação e Configuração	5
3.3 Instalação de Extensões	6
3.4 Customização	7
3.5 Testes Realizados	8
4 Análise de Dados Candidatos à Publicação	12
5 Considerações Finais	14
Referências	

1 Introdução

De acordo com [1], “Dados Abertos são dados que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer pessoa – sujeitos, no máximo, à exigência de atribuição à fonte original e ao compartilhamento pelas mesmas licenças em que as informações foram apresentadas”. O conceito de Dados Abertos envolve não só a disponibilização de informação em documentos textuais, mas em uma série de formatos, que se estendem por imagens, planilhas, até tipos mais elaborados, como Linked Data (dados ligados entre si) [2].

A publicação de dados vem sendo uma prática cada vez mais adotada no mundo, por instituições públicas e empresas privadas. Caso sejam utilizados de maneira correta, podem gerar produtos e serviços úteis para a sociedade, podendo propiciar mecanismos de controle social, transparência, cidadania ativa, cooperação, melhorias na administração pública, criação de novas ferramentas, trazendo mais desenvolvimento e bem-estar social [3].

No Brasil, em 2011, foi instituída a Lei de Acesso à Informação (LAI), que entrou em vigor em 2012. A LAI regulamenta o direito constitucional de acesso às informações públicas dos órgãos e entidades do governo, que passaram a disponibilizá-las para qualquer pessoa, física ou jurídica, sem necessidade de apresentar um motivo [4]. Atualmente, vários órgãos do governo já disponibilizam seus dados.

No contexto de cidades inteligentes, um portal de dados abertos é indispensável, pois é nele que os dados coletados e gerados pelos sistemas que compõem a infraestrutura da cidade serão disponibilizados para a população. Portanto, além da publicação dos dados governamentais, o portal compartilhará informações relevantes do contexto de cidades inteligentes que poderão ser usadas pelos próprios cidadãos para criar ferramentas ofereçam melhorias na administração pública e para o bem estar social.

2 Objetivo

Apresentar a importância do portal de dados abertos no contexto de cidades inteligentes e mostrar os resultados dos estudos relacionados ao portal de dados da prefeitura do Natal. Descrevendo a experiência da implantação de uma instância de testes desse portal, mostrando os problemas enfrentados e as soluções encontradas para resolvê-los.

3 Desenvolvimento do Portal

Nesta seção serão apresentados estudos mais aprofundados relacionados ao desenvolvimento de um portal de dados abertos, descrevendo sucintamente suas etapas de instalação, configuração e customização.

3.1 Arquitetura do CKAN

A arquitetura da aplicação é composta pelos seguintes componentes:

- Interface WSGI;
- Apache (modwsgi), nginx, etc.;
- Os metadados são armazenados em um banco de dados PostgreSQL;
- O servidor Apache Solr é responsável pelo mecanismo de busca e indexação;
- Outros componentes: extensões complementares e DataStore para armazenamento independente de dados que compõem os recursos publicados. Nesse casos, é necessário sua instalação e/ou configuração.

3.2 Instalação e Configuração

A instalação de uma do CKAN pode ser feita de 3 formas: a partir de um pacote de sistema operacional (nesse caso, o SO deve ser Ubuntu 14.04 64 bits ou Ubuntu 12.04 64 bits), a partir do código fonte, ou a partir de uma imagem do container Docker.

A instância de testes do portal do Natal (ckan.imd.ufrn.br) foi instalada em uma máquina virtual no data center do IMD a partir do código fonte seguindo as instruções presentes na documentação do CKAN. As etapas seguidas para instalação foram as seguintes:

- instalar pacotes de dependências;
- instalar CKAN num ambiente virtual Python;
- configurar banco de dados Postgres;
- criar um arquivo de configurações do CKAN (production.ini);
- configurar o Solr;
- definir usuários e criar tabelas no banco de dados;

- configurar o DataStore (opcional);
- criar um link para o arquivo para who.ini
- por fim, foi necessário criar conta de usuário administrador do site.

3.3 Instalação de Extensões

Algumas extensões já vem instaladas e configuradas na aplicação, outras ainda precisam ser configuradas. Também existe a opção de instalação de plugins externos. Foi feito um estudo para analisar os principais plugins utilizados nos portais de dados abertos de outras instituições e decidir quais seriam úteis para o portal da prefeitura do Natal. Por fim, definiu-se quais seriam instalados/configurados em primeira instância.

Até o momento os seguintes plugins estão em funcionamento no portal:

- DataStore:** para armazenamento de dados contidos em arquivos;
- DataPusher:** para extrair os dados dos arquivos e inserir no DataStore;
- Multilingual:** para ver a interface web do CKAN traduzida.
- DataExplorer, DataStore Grid, DataStore Map, DataStore Graph, Text View, Image View, Web Page View:** para visualização de dados;
- Stats Extension:** extensão para estatísticas do CKAN.

Outros plugins que precisam ser instalados/configurados:

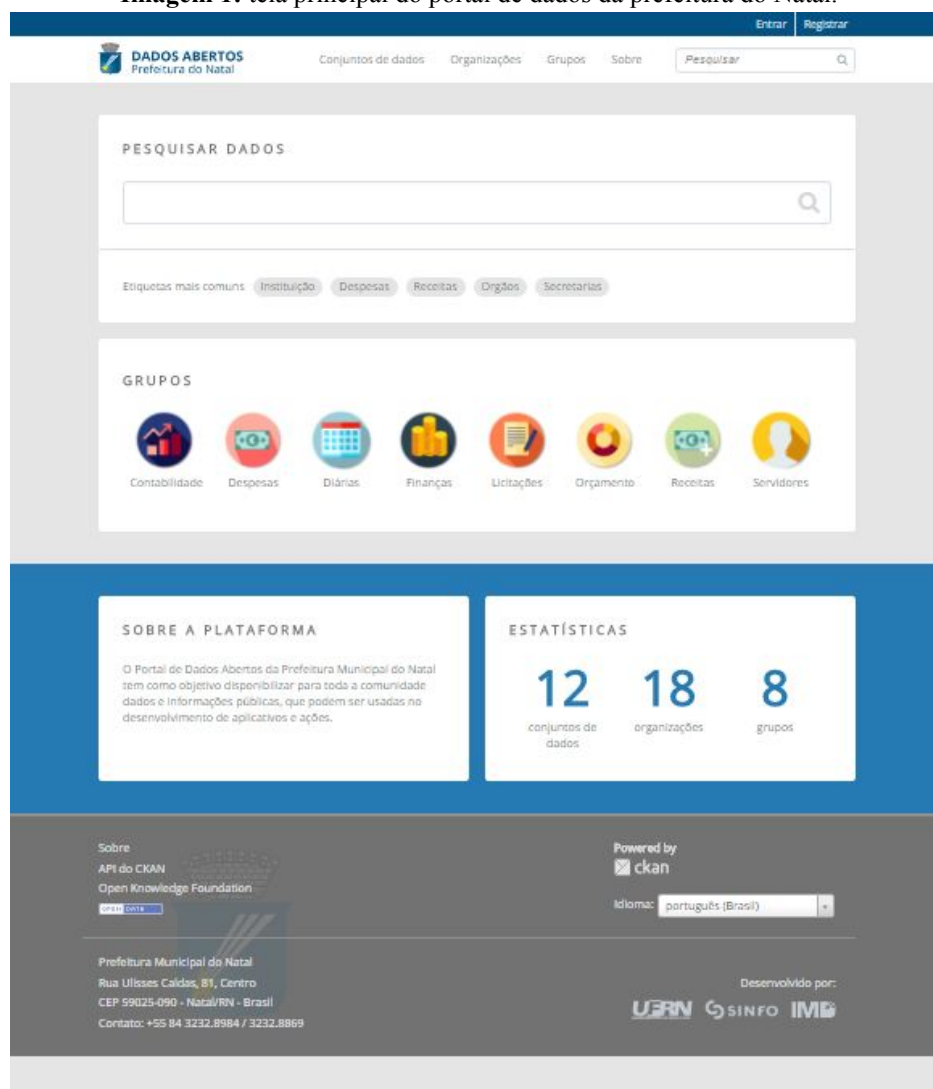
- Painel:** permite combinar vários pontos de vista em um único painel;
- Visualizador de PDF:** permite para processar arquivos PDF na página de recursos;
- GeoJSON mapa:** processa do GeoJSON arquivos em um mapa interativo;
- Gráficos Básicos:** fornece tipos de gráficos alternativos e as prestações;
- Discussões (DISQUS):** ferramenta para publicação de comentários de usuários em *datasets*.

3.4 Customização

A alteração do tema do portal pode ser feita pela própria interface WEB do CKAN, através de um formulário de configuração. Para isso, basta inserir o título do site, a logomarca, o código de estilo e outros atributos nos seus respectivos campos e submeter a alteração.

Foi feita uma customização beta para o portal da prefeitura do Natal (com a logo, nome da prefeitura e seguindo o padrão de as cores e estilos dos sites que a prefeitura possui atualmente). A imagem 1 mostra como a tela inicial do portal.

Imagem 1: tela principal do portal de dados da prefeitura do Natal.



Fonte: autoria própria (2016).

3.5 Testes Realizados

Foram realizados testes exploratórios no portal para avaliar a confiabilidade das operações de cadastro, alteração, remoção e listagens das entidades da aplicação (por exemplo, organizações, grupos, datasets, usuários etc.).

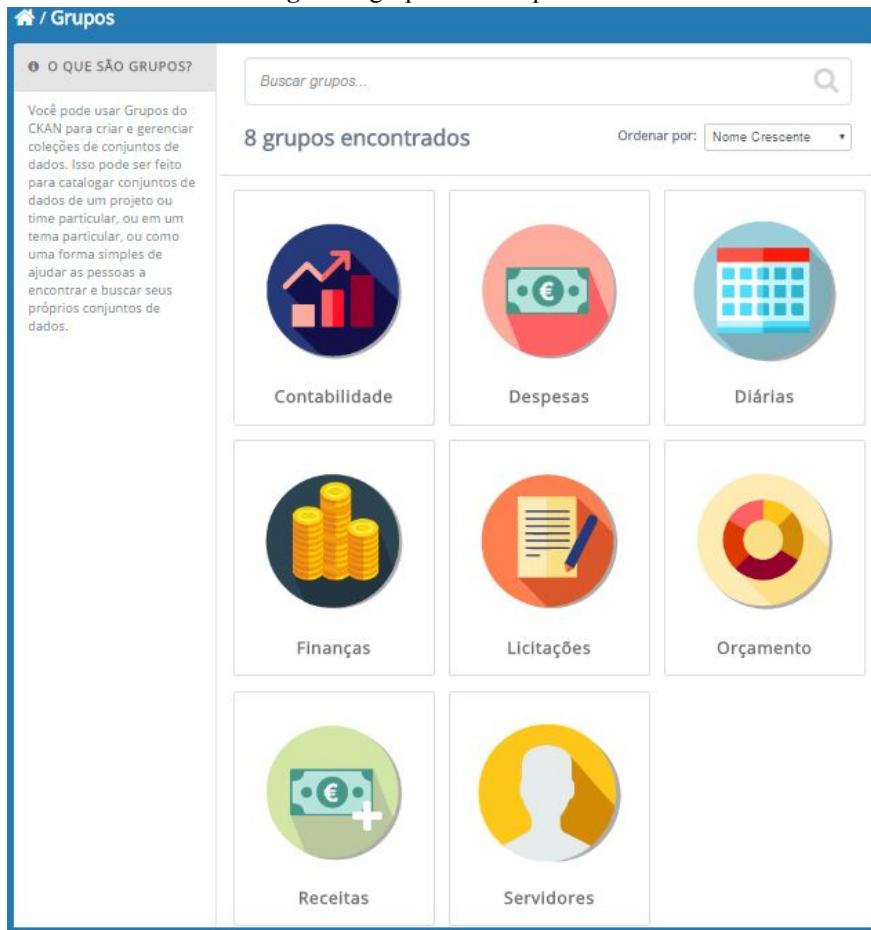
Na abordagem, criou-se organizações para as secretarias da prefeitura (imagem 2), adicionou-se diversos *datasets* (alguns *datasets* são exibidos na imagem 4) à essas organizações e agrupou-se os *datasets* entre as categorias criadas (os grupos criados são mostrados na imagem 3).

Imagem 2: organizações criadas para testes.



Fonte: autoria própria (2016).

Imagem 3: grupos criados para testes.



Fonte: autoria própria (2016).

Foram adicionados os grupos mostrados na imagem 3 para testes, porém ainda não estão definidos quais irão para produção e se algum desses deverá ser removido. Um estudo sobre a escolha dos dados que serão publicados no portal será detalhado na seção 4 deste relatório.

Imagem 4: *datasets* publicados para testes.

The image shows a web interface for a data portal. At the top, there is a search bar with the placeholder text "Pesquisar conjuntos de dados..." and a magnifying glass icon. Below the search bar, the text "12 conjuntos de dados encontrados" is displayed, along with a sorting dropdown menu set to "Relevância". The main content area is divided into several sections, each with a title and a brief description, followed by available file formats (CSV, PDF, TXT):

- TIPOS DE DESPESAS (ELEMENTOS)**: Apresenta uma relação de despesas por elementos. Formatos: CSV, PDF, TXT.
- TIPOS DE RECEITAS (CLASSIFICAÇÃO ECONÔMICA)**: Apresenta uma relação de receitas por instituição. Formatos: CSV, PDF, TXT.
- DESPESAS DE DIÁRIAS POR CREDOR/INSTITUIÇÃO**: Apresenta uma relação de despesas com diárias por credor. Formatos: CSV, TXT.
- ESCOLAS E CMEIS DO MUNICÍPIO**: Apresenta o quadro das escolas e CMEIs do Município do Natal. Formato: CSV.
- FINALIDADE DE RECEITA (RECURSO)**: Apresenta uma relação de receitas por instituição. Formatos: CSV, PDF, TXT.
- DESPESAS POR INSTITUIÇÃO/ORGÃO**: Apresenta uma relação de despesas por instituição. Formatos: CSV, PDF, TXT.
- EQUIPAMENTOS DE FISCALIZAÇÃO ELETRÔNICA**: Apresenta uma relação de equipamentos de fiscalização eletrônica em funcionamento na cidade do Natal.

On the left side, there is a sidebar menu with the following categories and counts:

- ORGANIZAÇÕES**
 - Prefeitura do Natal (9)
 - Atribuições da Secr... (2)
 - Secretaria Municipa... (1)
- GRUPOS**
 - Receitas (3)
 - Despesas (3)
 - Diárias (2)
 - Orçamento (1)
- ETIQUETAS**
 - Instituição (4)
 - Despesas (3)
 - Receitas (2)
 - Orgãos (2)
 - Secretarias (1)
 - Saúde (1)
 - Recurso (1)
 - Orçamento (1)
 - Líquidas (1)
 - Investimentos (1)
- FORMATOS**
 - CSV (12)

A link "Mostrar mais Etiquetas" is also visible in the sidebar.

Fonte: autoria própria (2016).

Foram realizados testes de inserção, alteração e remoção de recursos nos *datasets* com vários arquivos em diversos formatos extraídos de sites e portais da prefeitura já existentes. A imagem 5 mostra a visualização de um desses recursos no formato CSV.

Imagem 5: recurso do *dataset* de Despesas.

Organizações / Prefeitura de Natal / Tipos de Despesas (Elementos) / Despesas por Elemento - 2014

DESPESAS POR ELEMENTO - 2014 [Ir para recurso](#) [API de dados](#)

URL: <http://ckan.imd.ufrn.br/dataset/1aa667cb-5190-41bc-adb5-78169a6f5d61/resource/fa8bec12-a1eb-470b-9199-39b25ca5dc49/d...>
Despesas por elemento no ano de 2014.

Explorador de Dados Embutir

Add Filter

Grid Graph Map 53 records 1 - 53 Search data ... Go Filters

_id	Grupo ...	Element...	Empenh...	Anulado	Liquida...	Pago
1	PESSOA...	VENCIM...	620.734...	83.546.8...	557.183...	553.049...
2	OUTRA...	Subveny...	950.000,00	0,00	0,00	0,00
23	OUTRA...	OUTRO...	173.990,88	0,00	173.990,88	173.029,87

RECURSOS

- Despesas por Elemento - ...
- Despesas por Elemento - ...
- Despesas por Elemento - ...
- Despesas por Elemento - ...
- Despesas por Elemento - ...
- Despesas por Elemento - ...

SOCIAL

- Google+
- Twitter
- Facebook

INFORMAÇÕES ADICIONAIS

Campo	Valor
Ultima atualização	25/Agosto/2016
Criado	25/Agosto/2016
Formato	CSV
Licença	Outra (Aberta)
created	há 2 meses
datastore active	1
format	CSV
has views	1
id	fa8bec12-a1eb-470b-9199-39b25ca5dc49
last modified	há 2 meses
on same domain	1
package id	1aa667cb-5190-41bc-adb5-78169a6f5d61
position	2
revision id	b01ec18e-8f25-4a98-9ce3-06f96bdff386
state	active
url type	upload

Ocultar

Fonte: autoria própria (2016).

4 Análise de Dados Candidatos à Publicação

Antes de compartilhar qualquer informação, é necessário avaliar, entre outros aspectos, sua relevância para a sociedade. Nesse sentido, foi feito um comparativo entre os portais das prefeituras do Recife, São Paulo, Fortaleza, Rio de Janeiro e Porto Alegre, observando as categorias e dados que estes portais já publicaram. O objetivo era fazer um apanhado para saber quais dados comuns estes portais municipais estão publicando atualmente.

O resultado do estudo pode ser visto na tabela 1 e mostra que dados relacionados às áreas de finanças e orçamentos, assim como os das áreas de turismo, cultura, esporte e lazer são os mais publicados dentre os portais analisados.

Tabela 1: apanhado das publicações dos portais municipais.

Portais de Dados Abertos das Prefeituras							
Temas / Grupos	Conjuntos de Dados	#	Recife	São Paulo	Fortaleza	Rio de Janeiro	Porto Alegre
Educação	Censo Escolar	4	x	x	x	x	
	Rede de Educação	3	x		x		x
	Dados de Servidores	2	x	x			
Finanças e Orçamentos	Despesas	5	x	x	x	x	x
	Obras	4	x	x		x	x
	Contratos	4		x	x	x	x
	Empresas da Cidade	3	x	x	x		
	Receitas	3	x		x	x	
	Convênios	3		x		x	x
	Diárias	2			x		x
	Impostos e Taxas	3		x		x	x
	Imóveis da Administração Pública Municipal	2	x			x	
	Gastos com Passagens	1					x
	Bens	1			x		
Infraestrutura e Urbanismo	Área urbana	3	x		x	x	
	Cemitérios e Funerárias	3	x	x		x	
	Aeroportos e Rodoviárias	1					x
	Acessibilidade	2				x	x
	Feiras e Mercados	4	x	x		x	x

Meio Ambiente	Recursos Hídricos	4	x	x		x	x
	Coleta de Lixo	4	x	x		x	x
	Áreas Verdes	3	x		x	x	
	Dados Meteorológicos	2				x	x
Saúde e Bem Estar	Casos de Epidemias e Doenças	4	x		x	x	x
	Unidades de Saúde	5	x	x	x	x	x
	Academias da Cidade	2	x			x	
	Serviço de Atendimento Móvel de Urgência	1	x				
	Testes e Vacinação	2	x			x	
Segurança e Defesa Civil	Equipamentos de segurança urbana	3		x		x	
	Monitoramento das Áreas de Riscos	2	x			x	
Transporte e Mobilidade	Ciclovias, Ciclofaixas, Estações de Aluguel de Bikes e Rotas	4	x	x		x	x
	Ocorrências de Trânsito	3	x			x	x
	Fluxo de veículos	1	x				
	Equipamentos de Monitoramento e Fiscalização de Trânsito	4	x		x	x	x
	Transporte Público	3			x	x	x
	Calçadas dos Corredores de Transporte Metropolitano	1	x				
Turismo, Cultura, Esporte e Lazer	Equipamentos Culturais	5	x	x	x	x	x
	Equipamentos de Esporte	5	x	x	x	x	x
	Equipamentos de Lazer	5	x	x	x	x	x
	Eventos	4	x		x	x	x
	Hotéis	3	x			x	x
	Transparência no turismo, esporte, cultura e lazer	2		x			x
	Bares e Restaurantes	1	x				

Fonte: autoria própria (2016).

5 Considerações Finais

Foi identificado um erro ao realizar a instalação do CKAN a partir da imagem do Docker. Não foi encontrada uma solução para o problema, por isso utilizou-se o método de instalação a partir do código fonte.

Alguns erros também foram encontrados ao realizar os testes do portal, dentre eles:

- a. **Após publicar arquivo com tamanho maior que o limite disponibilizado:** o tamanho padrão para upload de arquivos no portal é de 10mb, ao submeter arquivos maiores que esse limite o portal apresentou um erro;
- b. **Após publicar arquivo em um formato que não seja seu formato original:** ao tentar visualizar o recurso, o CKAN tenta abrir o arquivo no formato que foi informado e não consegue;
- c. **Após publicar arquivo com extensão desconhecida pelo CKAN:** nesse caso, o plugin de visualização desse arquivo não está instalado/configurado ou a instalação/configuração do plugin foi feita de forma incorreta;
- d. **Após publicar um recurso cujos dados não foram persistidos no DataStore:** Ao tentar visualizar o recurso, o CKAN tenta usar o dataProxy para processar o arquivo e exibir seus dados no dataExplorer, porém nem sempre o dataProxy consegue processar o arquivo;
- e. **Após publicar um CSV estruturado de um forma que o dataExplorer não consiga processá-lo:** caso hajam “,” ou “;” dentro das células, ocorre uma quebra de linha e a tabela é exibida de forma errada. Para evitar o erro, deve-se criar arquivos CSV colocando aspas duplas em todas as células;
- f. **Após publicar dados em momentos em que o sistema está atendendo à muitas requisições:** ele para de persistir os dados publicados. O sistema permite a publicação, mas, ao sair da tela, o arquivo desaparece.

Em todos os erros citados, a mensagem de erro só é exibida ao visualizar o recurso e não no momento em que é realizado o upload. Nesse caso, se o mantenedor do dataset não perceber a ocorrência do erro, o mesmo poderá ser ignorado.

Apesar dos problemas enfrentados e dificuldades observadas, o CKAN provou ser uma excelente plataforma para catalogação de dados. O portal da prefeitura, que está em fase de testes, ainda sofrerá algumas mudanças, principalmente no que se refere aos dados que serão publicados. Essas e outras questões serão estudadas mais adiante.

Além dos estudos tratados neste relatório, foi produzido um artigo que trata das estratégias de desenvolvimento do portal de dados abertos da cidade do Natal. O trabalho será apresentado no III Workshop sobre Cidades inteligentes - WCID 2016 que será realizado em Natal no mês de novembro.

Referências

- [1] Open Knowledge Foundation: The Open Data Handbook (2012), <http://opendatahandbook.org/>
- [2] Linked Open Data: The Essentials. A Quick Start Guide for Decision Makers, Vienna, 2012. <http://5stardata.info/>
- [3] KUCERA, Jan et al. Methodologies and Best Practices for Open Data Publication. In: ANNUAL INTERNATIONAL WORKSHOP ON DATABASES, TEXTS, SPECIFICATIONS AND OBJECTS, 2015, Jicin, p. 52 - 64.
- [4] RIBEIRO, Claudio Jose Silva; ALMEIDA, Reinaldo de Figueiredo. Towards Open Government Data: the publication of the Brazilian Social Security System's Statistical Data. In: INTERNATIONAL MULTI-CONFERENCE ON SOCIETY, CYBERNETICS AND INFORMATICS, 6., 2012, Orlando. Proceedings. Orlando: Imsci, 2012. p. 156 - 160.
- [5] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (*references*)
- [6] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [7] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.