



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

ARQUITETURA DE CUBO DE DADOS PARA REPRESENTAÇÃO DE TELEMETRIA DE SATÉLITES

Yuri Matheus Dias Pereira
Mauricio Vieira Ferreira Gonçalves
Rodrigo Rocha Silva

Relatório Técnico resultado do
Exame de Proposta de Disserta-
ção do Curso de Pós-Graduação
em Engenharia e Gerenciamento de
Sistemas Espaciais.

URL do documento original:

[<http://urlib.net/>](http://urlib.net/)

INPE
São José dos Campos
2019

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6923/6921

Fax: (012) 3945-6919

E-mail: pubtc@sid.inpe.br

**COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO
DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):****Presidente:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Membros:

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Dr. Amauri Silva Montes - Coordenação Engenharia e Tecnologia Espaciais (ETE)

Dr. André de Castro Milone - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr. Joaquim José Barroso de Castro - Centro de Tecnologias Espaciais (CTE)

Dr. Manoel Alonso Gan - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Clayton Martins Pereira - Serviço de Informação e Documentação (SID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Marcelo de Castro Pazos - Serviço de Informação e Documentação (SID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

ARQUITETURA DE CUBO DE DADOS PARA REPRESENTAÇÃO DE TELEMETRIA DE SATÉLITES

Yuri Matheus Dias Pereira
Mauricio Vieira Ferreira Gonçalves
Rodrigo Rocha Silva

Relatório Técnico resultado do
Exame de Proposta de Disserta-
ção do Curso de Pós-Graduação
em Engenharia e Gerenciamento de
Sistemas Espaciais.

URL do documento original:

[<http://urlib.net/>](http://urlib.net/)

INPE
São José dos Campos
2019



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](#).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#).

Informar aqui sobre marca registrada (a modificação desta linha deve ser feita no arquivo publicacao.tex).

“But I try not to think with my gut. If I’m serious about understanding the world, thinking with anything besides my brain, as tempting as that might be, is likely to get me into trouble. It’s OK to reserve judgment until the evidence is in.”

CARL SAGAN E ANN DRUYAN
em “O Mundo Assombrado pelos Demônios:
A Ciência Vista Como Uma Vela no Escuro”, 1995

RESUMO

Satélites são monitorados pelas equipes de solo via pacotes de telemetria, que informam o estado atual dos equipamentos e permitem avaliar a capacidade do satélite de continuar a sua missão. Esses pacotes de telemetria constituem um corpo de dados de tamanho e complexidade significativa, sendo que satélites que funcionam por vários anos geram dados históricos de grande volume, ainda úteis para a operação. Neste artigo apresentamos uma arquitetura baseada em conceitos de Big Data e Business Intelligence para criar uma representação de dados de telemetria pronta para a análise por operadores e engenheiros de satélite no Instituto Nacional de Pesquisas Espaciais (INPE), bem como apresentamos o fluxo de dados utilizado pelos dados históricos de telemetria de um dos satélites operados pelo INPE.

Palavras-chave: Cubo de Dados. Big Data. Operação. Satélite. Data Warehouse.

DATA CUBE ARCHITECTURE FOR THE REPRESENTATION OF SATELLITE TELEMETRY

ABSTRACT

Abstract

Keywords: Atmospheric turbulence. WETAMC campaign. LBA project. Chaotic behavior. Chaotic attractor.

LISTA DE FIGURAS

	<u>Pág.</u>
4.1 Fluxo de dados em uma arquitetura de Big Data	8
4.2 Arquitetura de um cubo de dados	9

LISTA DE TABELAS

	<u>Pág.</u>
3.1 Operadores e Arquiteturas de Big Data	6
4.1 Dados de Operação	7
A.1 Cronograma de atividades	17
A.2 Publicações planejadas	17

LISTA DE ABREVIATURAS E SIGLAS

WETAMC	–	Campanha de Mesoescala Atmosférica na Estação Úmida
IBGE	–	Instituto Brasileiro de Geografia e Estatística
MC	–	Método das Covariâncias
EDO	–	Equações Diferenciais Ordinárias
EDP	–	Equações Diferenciais Parciais
ECT	–	Energia Cinética Turbulenta
FDP	–	Função de Distribuição de Probabilidade
PR	–	Plot de Recorrência
FFT	–	Fast Fourier Transform
tS1200	–	Temperatura medida no nível superior às 12 horas
tS2300	–	Temperatura medida no nível superior às 23 horas
tM1200	–	Temperatura medida no nível médio às 12 horas
tM2300	–	Temperatura medida no nível médio às 23 horas
tI1200	–	Temperatura medida no nível inferior às 12 horas
tI2300	–	Temperatura medida no nível inferior às 23 horas
wS1200	–	Velocidade vertical do vento medida no nível superior às 12 horas

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
2 FUNDAMENTAÇÃO	3
2.1 Operação	3
2.2 Big Data	3
2.3 Cubo de Dados	3
3 TRABALHOS CORRELATOS	5
3.1 Outros Operadores	5
4 PROPOSTA	7
4.1 Dados	7
4.2 Fluxo dos Dados de Operação	8
4.3 Arquitetura de um Cubo de Dados	8
5 IMPLEMENTAÇÃO	11
5.1 RFragCubing	11
5.2 Medida de Similaridade	11
6 CONCLUSÃO E TRABALHOS FUTUROS	13
REFERÊNCIAS BIBLIOGRÁFICAS	15
ANEXO A - CRONOGRAMA E PUBLICAÇÕES	17

1 INTRODUÇÃO

[Resto da introdução]

Os capítulos restantes desta dissertação estão organizados da seguinte maneira:

- Capítulo 2: Este capítulo apresenta os conceitos e fundamentos correlatos, como apresentando os conceitos do Cubo de Dados, as definições utilizadas de *Big Data*, a definição do problema para a operação e como as outras agências espaciais e operadores estão utilizando esses conceitos com base na literatura recente.
- Capítulo 3: Neste capítulo a arquitetura proposta é apresentada e seus conceitos principais explicados, bem como o fluxo de dados atual do CCS e como a nova arquitetura vai melhorá-lo.
- Capítulo 4: Esse capítulo apresenta alguns resultados já alcançados, demonstrando os softwares que já foram escritos e as análises que já foram executadas.
- Capítulo 5: Com base na arquitetura proposta e nos resultados intermediários alcançados, esse capítulo apresentará as conclusões obtidas bem como as direções de implementação para o resto do trabalho de mestrado.

2 FUNDAMENTAÇÃO

Neste capítulo, vamos apresentar os conceitos relevantes para este documento, bem como apresentação uma revisão da literatura na área de *Big Data*, terminando com uma visão geral de como o conceito e as tecnologias recentes estão sendo utilizadas por variadas agências espaciais.

2.1 Operação

2.2 Big Data

[5 Vs]

[Data Science?]

2.3 Cubo de Dados

[Copiar a estrutura da tese do Rodrigo]

3 TRABALHOS CORRELATOS

3.1 Outros Operadores

A tabela 3.1 mostra uma revisão feita em artigos recentes sobre os operadores de satélite e quais tecnologias eles estão utilizando para atingir objetivos semelhantes, principalmente com o uso de *Big Data*, como demonstrado pelos artigos publicados.

Tabela 3.1 - Operadores e Arquiteturas de Big Data

Referência	Operador	Ferramenta	Tecnologias
(ADAMSKI, 2016)	L3 (EUA)	InControl	Hadoop, Spark, HBase, MongoDB, Cassandra, Amazon AWS
(BOUSSOUF et al., 2018)	Airbus	Dynaworks	Hadoop, Spark, HDFS, HBase, PARQUET, HIVE
(SCHULSTER et al., 2018)	EUMETSAT	CHART	MATLAB, MySQL, Oracle
(ZHANG et al., 2017)	SISSET (China)	-	Hadoop, HDFS, PostgreSQL, MongoDB, Logstash, Kibana, ElasticSearch, Kafka, MapReduce
(YVERNES, 2018)	Telespazio France	PDGS	OLAP (DataCube), Saiku, Pentaho, Jaspersoft OLAP
(DISCHNER et al., 2016)	SwRI + NOAA	CYGNSS MOC	SFTP, -
(EDWARDS, 2018)	EUMETSAT	MASIF	FTP, RESTful service, JMS Message Queue, PostgreSQL
(EVANS et al., 2016)	S.A.T.E + ESA/ESOC	-	Java, CSV, algoritmos
(FEN et al., 2016)	CSMT& (China)	-	não menciona as tecnologias
(TROLLOPE et al., 2018)	EUMETSAT	CHART	algoritmos, estudo de caso
(GILLES, 2016)	L-3	InControl	Amazon EC2, LXC, Nagios, repetição do primeiro
(HIGHSMITH et al., 2015)	Boeing + NASA	-	lançadores, não é o foco da arquitetura
(HENNION, 2018)	Thales Alenia	AGYR	Logstash, Kafka, InfluxDB, ElasticSearch, Kibana, Grafana
(MATEIK et al., 2017)	Stinger, NASA	-	Logstash, ElasticSearch, Kibana, HDF5, CSV, R, Python, AWS, Excel
(FERNÁNDEZ et al., 2017)	NASA	MARTE	R, CSV, ad-hoc

4 PROPOSTA

[Que tal trocar isso para uma "definição do problema"?]

4.1 Dados

A tabela 4.1 mostra os tipos de dados relevantes para a operação, a sua origem e o seu formato esperado, ignorando os dados provenientes da carga útil.

Tabela 4.1 - Dados de Operação

Tipo de Dado	Origem	Formato
Sensores de bordo	Equipamentos no satélite	Tabelas, CSV
Registros do Computador	Computador de Bordo	Texto (<i>Logs</i>)
Multimídia	Câmeras	MP4, JPG, RAW
Parâmetros orbitais	Operação, Rastreo	TLE, texto, tabelas
Documentação associada	Operadores, engenharia	Texto (Word, Excel)
Clima Espacial	Sensores no solo ou espaço	Texto, tabelas, avisos
<i>Situational Awareness</i>	Radares, US-STRACOM, etc	Texto, tabelas, avisos

Fonte: Expandido de (ZHANG et al., 2017)

Para este trabalho, apenas os dados vindos de sensores de bordo serão considerados. Os outros dados nesta tabela poderiam ser considerados para uma *Data Warehouse* mais completa, pois um cubo de dados pode ser formado sobre quaisquer um desses dados.

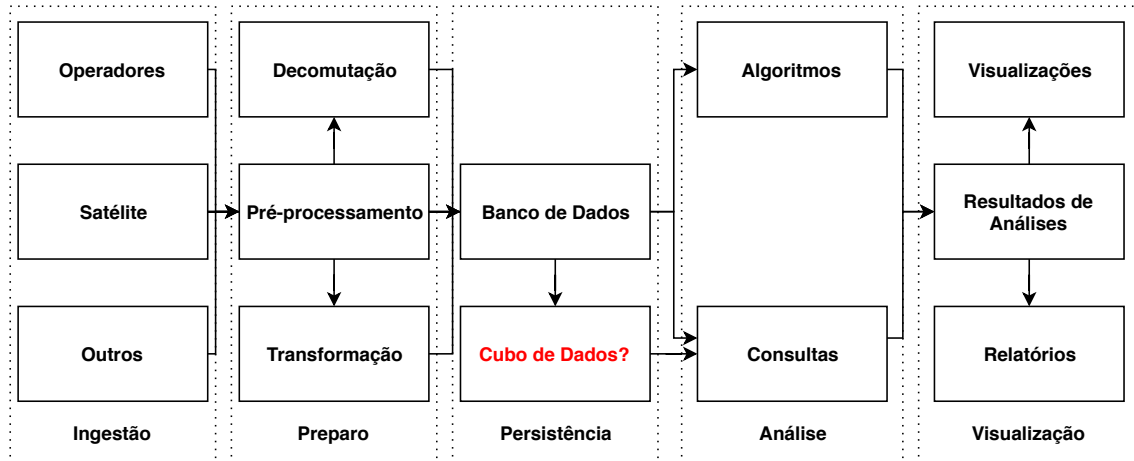
Por exemplo, um cubo de dados textual poderia ser feito sobre os documentos associados a operação, como o CONOPS, tabelas de telecomandos e documentação de engenharia de sistemas para facilitar a análise da documentação sendo gerada pelo satélite. Um cubo multimídia poderia ser gerado sobre os dados multimídia tirados pelas câmeras do satélite para correlacionar com os dados gerados pelos sensores, e assim em diante. Alguns exemplos de cubos possíveis de serem feitos estão em (SILVA, 2015).

Esta lista não é exaustiva, e pode incluir dados da carga útil caso sejam relevantes para a análise em questão, como ajudar na georeferênciação de imagens tiradas pelo satélite.

4.2 Fluxo dos Dados de Operação

A figura 4.1 demonstra o fluxo de dados esperado de uma arquitetura de *Big Data*.

Figura 4.1 - Fluxo de dados em uma arquitetura de Big Data



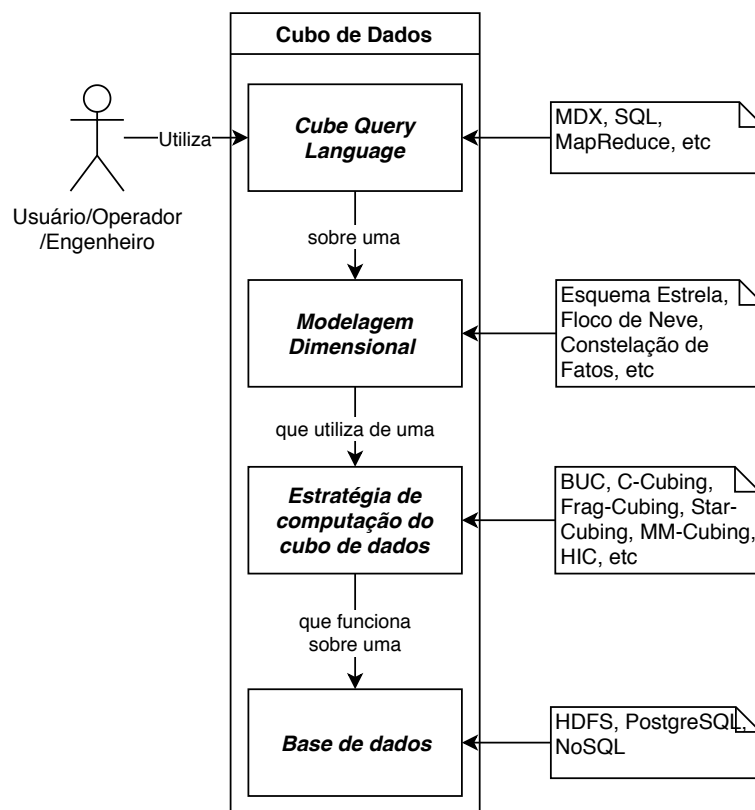
Fonte: Adaptado de (ZHANG et al., 2017)

4.3 Arquitetura de um Cubo de Dados

A figura 4.2 demonstra a divisão em 4 partes do que é entendido neste trabalho como um Cubo de Dados.

Esta proposta foca apenas na proposição de um algoritmo de computação do cubo de dados mais apropriado, utilizando das outras seções quando elas vão se tornando necessárias.

Figura 4.2 - Arquitetura de um cubo de dados



Fonte:

5 IMPLEMENTAÇÃO

[O que já foi feito aqui?]

Essa seção apresenta os trabalhos e os resultados feitos até o momento.

5.1 RFragCubing

[RFragCubing]

5.2 Medida de Similaridade

[SimilarityMeasure?]

6 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho apresenta uma abordagem de cubo de dados para executar operações de análise nos dados de telemetrias de satélites. Essa abordagem utiliza de conceitos de *Big Data* para orientar a execução de consultas em dados com muitas dimensões e de alta complexidade. Uma revisão da literatura de arquiteturas de *Big Data* é apresentada, demonstrando que tipos de tecnologias e abordagens estão em uso por outros operadores de satélite.

Também são apresentados resultados intermediários de análises e softwares feitos para a análise primária dos dados de telemetria.

Como essa arquitetura é melhor(diferente?) da utilizada por outros operadores? O que o Cubo de Dados traz de diferente?

Planos futuros, o que vai ser implementado daqui para frente

Queries interessantes dos operadores?

Implementação do cubo de dados?

REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMSKI, G. Data Analytics for Large Constellations. In: **SpaceOps 2016 Conference**. [S.l.]: American Institute of Aeronautics and Astronautics, 2016. (SpaceOps Conferences). 00000. 6
- BOUSSOUF, L.; BERGELIN, B.; SCUDELER, D.; GRAYDON, H.; STAMMINGER, J.; ROSNET, P.; TAILLEFER, E.; BARREYRE, C. Big Data Based Operations for Space Systems. In: **2018 SpaceOps Conference**. [S.l.]: American Institute of Aeronautics and Astronautics, 2018. 00000. 6
- DISCHNER, Z.; REDFERN, J.; ROSE, D.; ROSE, R.; RUF, C.; VINCENT, M. CYGNSS MOC; Meeting the challenge of constellation operations in a cost-constrained world. In: **2016 IEEE Aerospace Conference**. [S.l.: s.n.], 2016. p. 1–8. 00000. 6
- EDWARDS, T. Dealing with the Big Data - The Challenges for Modern Mission Monitoring and Reporting. In: **15th International Conference on Space Operations**. Marseille, France: American Institute of Aeronautics and Astronautics, 2018. ISBN 978-1-62410-562-3. 00000. 6
- EVANS, D. J.; MARTINEZ, J.; Korte-Stapff, M.; VANDENBUSSCHE, B.; ROYER, P.; RIDDER, J. D. Data Mining to Drastically Improve Spacecraft Telemetry Checking: A Scientist's Approach. In: **SpaceOps 2016 Conference**. [S.l.]: American Institute of Aeronautics and Astronautics, 2016, (SpaceOps Conferences). 00000. 6
- FEN, Z.; YANQIN, Z.; CHONG, C.; LING, S. Management and Operation of Communication Equipment Based on Big Data. In: **2016 International Conference on Robots Intelligent System (ICRIS)**. [S.l.: s.n.], 2016. p. 246–248. 00000. 6
- FERNÁNDEZ, M. M.; YUE, Y.; WEBER, R. Telemetry Anomaly Detection System Using Machine Learning to Streamline Mission Operations. In: **2017 6th International Conference on Space Mission Challenges for Information Technology (SMC-IT)**. [S.l.: s.n.], 2017. p. 70–75. 00003. 6
- GILLES, K. Flying Large Constellations Using Automation and Big Data. In: **SpaceOps 2016 Conference**. [S.l.]: American Institute of Aeronautics and Astronautics, 2016, (SpaceOps Conferences). 00000. 6

HENNION, N. Big-data for satellite yearly reports generation. In: **2018 SpaceOps Conference**. [S.l.]: American Institute of Aeronautics and Astronautics, 2018. 00000. [6](#)

HIGHSMITH, H.; BROCK, J. E.; STEPHENS, D. E. Space Launch System (SLS) data acquisition and sensor system for human space flight. In: **2015 IEEE Aerospace Conference**. [S.l.: s.n.], 2015. p. 1–9. 00006. [6](#)

MATEIK, D.; MITAL, R.; BUONAIUTO, N. L.; LOUIE, M.; KIEF, C.; AARESTAD, J. Using Big Data Technologies for Satellite Data Analytics. In: . [S.l.]: American Institute of Aeronautics and Astronautics, 2017. ISBN 978-1-62410-483-1. 00001. [6](#)

SCHULSTER, J.; EVILL, R.; PHILLIPS, S.; FELDMANN, N.; ROGISSART, J.; DYER, R.; ARGEMANDY, A. CHARTing the Future – An offline data analysis and reporting toolkit to support automated decision-making in flight-operations. In: **15th International Conference on Space Operations**. Marseille, France: American Institute of Aeronautics and Astronautics, 2018. ISBN 978-1-62410-562-3. 00001. [6](#)

SILVA, R. R. **Abordagens para Cubo de Dados Massivos com Alta Dimensionalidade Baseadas em Memória Principal e Memória Externa: HIC e BCubing**. 00000. Tese (Doutorado) — Instituto Tecnológico de Aeronáutica, São José dos Campos, 2015. Acesso em: 01 ago. 2018. [7](#)

TROLLOPE, E.; DYER, R.; FRANCISCO, T.; MILLER, J.; GRISO, M. P.; ARGEMANDY, A. Analysis of automated techniques for routine monitoring and contingency detection of in-flight LEO operations at EUMETSAT. In: **2018 SpaceOps Conference**. Marseille, France: American Institute of Aeronautics and Astronautics, 2018. ISBN 978-1-62410-562-3. 00001. [6](#)

YVERNES, A. Copernicus Ground Segment as a Service: From Data Monitoring to Performance Analysis. In: **15th International Conference on Space Operations**. Marseille, France: American Institute of Aeronautics and Astronautics, 2018. ISBN 978-1-62410-562-3. 00000. [6](#)

ZHANG, X.; WU, P.; TAN, C. A big data framework for spacecraft prognostics and health monitoring. In: **2017 Prognostics and System Health Management Conference (PHM-Harbin)**. [S.l.: s.n.], 2017. p. 1–7. 00000. [6](#), [7](#), [8](#)

ANEXO A - CRONOGRAMA E PUBLICAÇÕES

A tabela mostra o cronograma esperado para as próximas atividades do mestrado.

Tabela A.1 - Cronograma de atividades

Atividade	maio	jun.	jul.	ago.	set.	out.	nov.	dec.	jan.	fev.
Exame de Proposta	X									
Submissão Artigo Periódico					X					
Apresentação Conferência						X				
Defesa final							X			X

A tabela mostra os veículos de publicação planejados e os já publicados/em processo de publicação.

Tabela A.2 - Publicações planejadas

Nome	Qualis	Prazo	Notas
WETE 2018	Conferência		Publicado
IAC 2019	Conferência		Artigo aceito, porém feito pela matéria do Prof. Geilson (não é nessa área), porém sou o 1 autor
BDCAT	Conferência	29/08/2019	Nova Zelândia...
WETE 2019	Conferência		?
IEEE América Latina	B2		?
International Journal of Data Warehousing and Mining	B1 (CC)		Não tem qualis para ENG-III, caro, JCR 0,66

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.